The Foundations of Wealth-Enhancing Democracy: Aristotle, Lindahl, and Institutional Design in Ancient Greece

Robert K. Fleck
The John E. Walker Department of Economics
Clemson University
Clemson, SC 29634
phone: (864) 656-3964
e-mail: rfleck@clemson.edu

F. Andrew Hanssen
The John E. Walker Department of Economics
Clemson University
Clemson, SC 29634
phone: (864) 656-5474
e-mail: fhansse@clemson.edu

December 31, 2015

Abstract: This paper examines the circumstances under which majority rule decision-making will support, rather than impede, wealth creation. We develop a model – inspired by Aristotle and Lindahl – in which voting on whether to provide a public good can work well or badly, depending on whether voters share the public good's costs and benefits in a manner close to or far from the Lindahl prescription. Under some conditions, it is feasible to design political institutions that complement the economy's exogenous features (specifically, the opportunities for wealth creation), so that when citizens vote on the basis of what they stand individually to gain or lose from a policy decision, they will vote as if they are seeking to maximize net social benefits. Yet under other (i.e., non-ideal) conditions, the best feasible policy under majority-rule governance will require facing a tradeoff between two objectives: the maintenance of the majority's support for wealth-enhancing policies versus obtaining benefits from specialization and trade. We use the model to guide our analysis of how two pioneers of majority-rule systems – Athens and Sparta – designed (and redesigned) their institutions in response to changes in their opportunities for wealth creation. We also discuss lessons for the modern world, including polarization in American politics.

For helpful comments, we thank Dan Benjamin, Bill Dougan, Dino Falaschetti, Brooks Kaiser, Carl Kitchens, Dean Lueck, Bryan McCannon, Josh Ober, Paul Pecorino, Athanassios Pitsoulis, Dave Teegarden, Matt Van Essen, Dan Wood, seminar participants at Clemson University, Florida State University, the Property and Environment Research Center (PERC), the University of Alabama, the 2013 meetings of the International Society for New Institutional Economics (ISNIE),the 2014 meetings of the Southern Economic Association, the 2015 meetings of the Association of Ancient Historians, and the 2015 World Economic History Congress. Much of the work on this paper was conducted while the authors were visiting PERC as Lone Mountain Fellows. We thank PERC for its financial support and many valuable discussions.

[O]f the private land, part should be near the border, and the other near the city, so that, each citizen having two lots, they may all of them have land in both places; there is justice and fairness in such a division, and it tends to inspire unanimity among the people in their border wars. Where there is not this arrangement some of them are too ready to come to blows with their neighbors, while others are so cautious that they quite lose the sense of honor. Wherefore there is a law in some places which forbids those who dwell near the border to take part in public deliberations about wars with neighbors, on the ground that their interests will pervert their judgment. For the reasons already mentioned, then, the land should be divided in the manner described.

-Aristotle, *Politics*, Part X of Book Seven, translated by Benjamin Jowett

I. INTRODUCTION

Although the overwhelming majority of the modern world's wealthy countries have democratic forms of government, the path to becoming a wealthy democracy has often proven elusive. One reason is that rather than inspiring economic development, voting may focus the policymaking process on the redistribution of wealth.[1] Indeed, even when dealing with the provision of public goods such as defense, the policy-making process can degenerate into pork barrel politics.[2] Yet if such problems are widely understood, we should expect farsighted voters to support steps (even if costly) to mitigate the resulting harm. And voters have done so, in a variety of ways and with mixed success, as research by scholars including Buchanan and Tullock (1962), Olson (1982), Riker (1982), and Acemoglu et al. (2014) demonstrates.

We build on such work by developing a model that focuses on the relationship between the

[1]This point has long been examined by scholars and, indeed, might be described as common knowledge. One well known comment on this is George Bernard Shaw's (1944): "A government which rob's Peter to pay Paul can always depend on the support of Paul."

[2]For example, many politicians oppose the closing of home-district military bases even though the bases have minimal national security value; see, e.g., Defense Base Closure and Realignment Commission (2005).

quality of democratic decision-making and the presence (or absence) of complementarities between political institutions and features of the economy. The model allows us to analyze how the nature of policy decisions, combined with the pattern of asset ownership, determine whether majority rule decision-making will support wealth-enhancing policies. The model highlights a key tradeoff. On the one hand, keeping voters' incentives aligned with respect to policy decisions may require forgoing potential gains from market-based exchange, because such gains can reduce the homogeneity of policy preferences among voters. On the other hand, to obtain mutual gains from exchange, it may be necessary to incur costs related to voters' incentives; specifically, as a consequence of trade and specialization, some wealth-creating policies may lack sufficient voter support. Thus, designing successful institutions depends on recognizing when the best feasible democratic decisions are "second-best" – because a system based on majority rule decisions cannot always support the "first-best" set of policies.

We examine these tradeoffs in the context of the institutions established by the first large-scale democracies, those of ancient Greece. At a time when Greece's rivals were ruled by kings or emperors, Greek *poleis* (city-states) experienced remarkable economic growth over a period of several centuries.[3] We argue that an important reason for this is that – as the quotation from Aristotle with which this paper begins illustrates – the Greeks made extraordinary efforts to design institutions that aligned incentives for important collective and private decisions. That said, the

[3]Neither women nor subordinate populations (slaves or serfs) could vote. However, among the male citizenry, rights were distributed with remarkable equality. Because our paper focuses on the functioning of collective decision-making among the enfranchised, it makes good sense to consider Spartan institutions highly "democratic" in the sense that, among the set of voters (i.e., the citizen men of Sparta), Spartan political institutions allocated power in an unusually egalitarian manner. Why only a relatively small fraction of the population ruled by the Spartan government could vote is a substantially different question (e.g., Fleck and Hanssen 2006, 2009).

specific institutions differed substantially among the various *poleis*—in ways that illustrate the tradeoff highlighted in our model (and in Aristotle's quote). Most famously, although Athens and Sparta took very different approaches to governance (because of different opportunities for generating and protecting wealth), each established systems under which individuals paid for the provision of public goods roughly in proportion to the marginal benefits they could expect to receive – in the parlance of modern public economics, they employed logic akin to Lindahl taxation.[4] Sparta established institutions that generated homogeneity in policy preferences by suppressing heterogeneity among citizens' human capital investments and property holdings. In Athens, where the importance of commercial activities would have rendered enforced homogeneity insuperably costly, the response was to engineer institutions that linked payments to individuals in a manner that provided incentives to perform critical tasks (e.g., rowing in the navy) and distributed collectively generated wealth to a broad segment of the citizenry.[5] But in each case, by keeping incentives well aligned for activities related to the creation and protection of wealth, the two *poleis* were largely able to avoid the types of political conflict that can cripple democracies.[6]

Furthermore, the process of design appears to have been quite conscious. Both Athens and

---

[4]Our point is not that the Greeks implemented a actual Lindahl mechanisms, but rather that their institutions generated incentives in line with the logic of Lindahl taxation – and our historical analysis will make this clear. Note that the difficulty of achieving a Lindahl equilibrium, given the extreme information revelation requirements, has received much attention in the literature on mechanism design; see, e.g., Vickrey (1961), Clarke (1971), and Groves (1973).

[5]Sparta established its famous military-oriented institutions, while Athens established a direct democracy under which assembly votes trumped pre-existing laws. Athens later altered its institutions in a number of ways in order to reduce assembly influence and create a de facto rule of law.

[6]Each of the two was remarkably stable in its system of government. Athens broader democracy lasted for nearly two centuries (until the independence of the Greek world was brought to a close by the successors to Alexander the Great), while Sparta's military-oriented constitution endured for more than two centuries (until a catastrophic military defeat). See Sections III and IV.

Sparta experienced major and abrupt changes in their sources of wealth – Sparta in its conquest of a neighboring territory, Athens in first uniting and later losing an empire. Thus, for both Athens and Sparta, institutional redesign helped to align voters' incentives on issues necessary for the survival and economic performance of the *polis*. When viewed in the context of our model, the lessons from Athens and Sparta provide broader insight into the successes and failures of democratic institutions.

Our paper thus contributes to the literature on the importance of "good" (wealth-enhancing) institutions for economic development, and our conclusions bolster the increasing recognition that the quality of an institution depends on the extent to which it aligns private incentives with public interests. In doing so, we draw on ideas expressed in the foundations of the public choice literature (e.g., Buchanan and Tullock 1962) and a large recent literature.[7] Naturally, we also rely heavily on the voluminous literature on Greek history, and we seek to contribute to the emerging literature examining institutional design as part of the foundation for the extraordinary economic performance of ancient Greece.[8]

---

[7]See, e.g., Lipset (1959), Fernandez and Rodrik (1991), Barro (1997), Acemoglu and Robinson (2000, 2001), Fleck (2000), Boix (2003), Jack and Lagunoff (2006), Acemolgu, Johnson, Robinson, and Yared (2008, 2009), Falaschetti (2008), Treisman (2012). Our analysis draws on the literature that examines the conditions under which a broadening or narrowing of the group with policy-making power can move policy in a beneficial direction. For example, North and Weingast (1989) argue that the broadened distribution of power in 17th century Britain allowed the government to commit credibly to uphold private property rights. Justman and Gradstein (1999) explain Britain's 19th century expansion of rights as a self-interested response by ruling elites that garnered increasing support as time passed. Lizzeri and Persico (2004) and Llavador and Oxoby (2005) develop and test models in which a pro-commerce element among a divided elite (old agriculture versus new commerce) seeks a means of locking-in support for growth-promoting policies. Focusing on ancient Greece, Fleck and Hanssen (2006, 2013a) consider how the potential for economic development led to the expansion (and at times contraction) of political rights, and how those changes in political rights in turn supported economic development.

[8]See, e.g., Kaiser (2007), Ober (2008, 2010), Morris (2010), Fleck and Hanssen (2006, 2012), Pitsoulis (2011) Ober and Weingast (2013), Teergarden (2014), Caugati (2015).

II. THE MODEL

This section develops a model to illustrate how the quality of democratic decision-making can depend on the presence (or absence) of complementarities between political institutions and features of the economy. More specifically, we consider how the nature of policy decisions, combined with the pattern of asset ownership, determine whether majority rule decision-making will support wealth-enhancing policies. The model thus builds on Aristotle's insight with respect to maintaining a pattern of land ownership that generates well-aligned incentives among voters. When applying the model to the real world, the key implications relate to the tradeoffs faced when designing institutions and setting policy: Keeping voters' incentives aligned with respect to policy decisions may have costs (e.g., it may be necessary to forgo potential gains from market-based exchange); similarly, to obtain mutual gains from exchange, it may be necessary to incur costs related to voters' incentives (e.g., as a consequence of trade and specialization, some wealth-creating policies may lack sufficient voter support).

The main conclusion, in short, is that designing successful institutions can depend on recognizing when the best feasible democratic decisions are "second-best" – because a system based on majority rule decisions cannot always support the "first-best" (i.e., ideal) set of policies. To keep our exposition of this point straightforward, we will develop our model as follows. We begin by setting out a simple framework that allows for heterogeneity among the citizenry (two types of citizens), one type of asset (initially allocated uniformly among the citizenry), a yes/no policy decision with respect to allowing the asset to be bought and sold, and a yes/no policy decision regarding a public good. We then ask the very easy question of what the first-best policy would be – this provides a starting point to which we can compare a variety of outcomes, most importantly the

best feasible outcomes obtainable under democratic decision-making..

A. The Basic Framework

As an analytical tool, we consider individuals who begin behind a veil of ignorance with respect to which of two groups they will belong. Thus, the behind-the-veil citizenry can be viewed as homogeneous, knowing how many (but not which) individuals will become members of each group. This approach serves two purposes. First, it provides an intuitive way to characterize the "optimal" policy: What would the citizenry favor behind the veil? Second, it allows a simple but useful way to examine how the performance of democratic institutions will depend on the timing of key events – notably, the times when policy decisions are made in relation to the time when voters learn who stands to gain and who stands to lose as the consequence of a given policy decision. With this in mind, consider our formal assumptions:

Assumption 1: Assets. Each of n citizens is initially endowed with $\underline{A}$ units of a productive asset, with n a large number, $\underline{A}>0$, and $n\underline{A}$ the total (and fixed) amount of the asset in existence.

Assumption 2: Heterogeneity. The citizen population (after the veil is lifted) has $n\omega_H$ individuals of type H and $n\omega_L$ individuals of type L; $\omega_H + \omega_L = 1$; $\omega_H > 0$; $\omega_L > 0$; $\omega_H \neq \omega_L$.[9]

Assumption 3: Policy Options and Returns to Assets. The government makes two policy decisions. One is whether to allow (the alternative being to disallow) citizens to trade the asset A; if trade is allowed, it occurs at the market clearing price (as in a standard supply and demand model). The other policy decision is whether to provide a public good: At a cost of κ per citizen (i.e., at a total cost κn, shared equally among citizens), the government can set g=1 (the alternative being to set g=0, at zero cost), thus raising by γ (as indicated below) the per unit returns to the asset. The marginal benefit curves for asset ownership (which differ by type of citizen) are:

$MB_i = \alpha_H - \beta A_i + \gamma g$ if citizen i is type H
$MB_j = \alpha_L - \beta A_j + \gamma g$ if citizen j is type L

[9]By assuming $\omega_H \neq \omega_L$, we guarantee clear decisions under majority rule voting (by eliminating the cases in which the electorate would be evenly split).

where: $\alpha_H > \alpha_L > 0$; $\beta > 0$; $\gamma > 0$; $[(\alpha_H - \alpha_L)\omega_H]/\beta < \underline{A}$.[10]

Assumption 4: Information, Trade, and Voting. All citizens have full information about the model and, when trading or voting, each acts individually (and with rational expectations) to maximize his or her own expected benefits. Government decisions are made by majority rule: Each citizen (i) votes to disallow trade if and only if disallowing trade would increase that citizen's net benefits and (ii) votes to set g=1 if and only if setting g=1 would increase that citizen's net benefits.

Assumption 5: Order of Events. There are six time periods:

Period 1	Voters choose trade policy and public good policy
Period 2	Veil lifted
Period 3	Voters may be able to change trade policy and/or public good policy
Period 4	Trade occurs (if allowed)
Period 5	Voters may be able to change public good policy
Period 6	Outcomes realized (payoffs based on benefits defined in Assumption A3)

where the ability/inability to make trade policy changes in Period 3, along with the ability/inability to make public good policy changes in Periods 3 and 5, are exogenously determined.

B. First-Best Policy and the Cost of Deviating from the First-Best

How would one set policy (i.e., choose g=1 or g=0; allow or disallow trade) in the manner that maximizes the expected payoffs for behind-the-veil citizens? The solution is very easy:

First-Best Policy. Under the following rules, the expected payoffs to citizens will be maximized: allow trade; if $\gamma\underline{A} > \kappa$, set g=1; if $\gamma\underline{A} \leq \kappa$, set g=0.

To put the first-best policy in perspective, it is useful to calculate the costs incurred as a result of deviations from the first-best:

Loss from setting the non-optimal g. If g=0 when $\gamma\underline{A} > \kappa$, or if g=1 when $\gamma\underline{A} < \kappa$, the loss (i.e., the forgone gain) will be $|\gamma\underline{A}-\kappa|n$.

Loss from disallowing trade. If trade is disallowed, the loss can be calculated as the forgone gains

[10]By assuming $[(\alpha_H - \alpha_L)\omega_H]/\beta < \underline{A}$, we guarantee interior solutions for market-generated asset allocations (i.e., each type L individual will sell some, but not all, of his or her initial endowment).

from exchange (denoted T): $T = (.5n/\beta)\ (\omega_H\ \omega_L)\ (\alpha_H - \alpha_L)^2$.[11]

The factors influencing the costs of non-optimal decisions are straightforward to interpret. First, and quite obviously, when there is more at stake in the decision regarding g (i.e., when the gap between the benefits and costs, $|\gamma\underline{A}-\kappa|$, is greater), the cost of making the wrong decision with respect to g is larger. The gains from allowing A to be traded (or, equivalently, the cost of disallowing trade) will be greater when the population has more heterogeneity in the sense of A being much more valuable (at a given quantity) to type H citizens than to type L citizens (i.e., when α_H is much greater than α_L). Moreover, when there are substantial numbers of both types (when ω_H and ω_L are each close to half), rather than the two types occurring in very unequal numbers, there are more potential gains arising from the heterogeneity in marginal benefits (i.e., there are more gains to be obtained as a result of the difference between α_H and α_L).[12]

C. The Feasibility of First-Best Outcomes Under Majority Rule

Whether voters will choose policies that generate first-best outcomes depends on the values of exogenous parameters in combination with the timing of events. Recall that Assumption 5 specifies that the time at which trade will occur (if allowed) comes after the veil has been lifted (i.e., after potential traders know whether they would want to buy or sell), and that outcomes are realized

---

[11]Note that the calculations can be done separately for each decision (because the gains from exchange are independent of g, and the returns to setting g optimally are independent of trade). To see how the calculation of gains from trade works, note that if trade is allowed, transactions will occur at a market clearing price (denote this as P*), each type H citizen will purchase some A (denote the purchased quantity as A_D), and each type L citizen will sell some A (denote the sold quantity as A_S). Basic principles of supply and demand yield the following: $P^* = \alpha_H\ \omega_H + \alpha_L\ \omega_L - \beta A_i + \gamma g$; $A_D = [(\alpha_H - \alpha_L)\omega_L]/\beta$; $A_S = [(\alpha_H - \alpha_L)\omega_H]/\beta$. And this yields $T = (.5n/\beta)\ (\omega_H\ \omega_L)\ (\alpha_H - \alpha_L)^2$.

[12]Although less important for our analysis, the other parameter, β, also has a straightforward interpetation. For smaller values of β (i.e., when the demand curves for A are flatter), there will be greater quantities supplied and demanded; hence, there will be greater gains from trade.

in the final period. Thus, the critical timing issues pertain to how late (i.e., in which periods) the majority can still change trade and/or public goods policies. Also recall that voters have full information about the social benefits of policies even before the veil is lifted; thus, there is no option value (in terms of social benefits) to having the ability to change policy later. The inability to reach the first best arises when a time-inconsistency problem renders a majoritarian system incapable of maintaining the policies to which all individuals initially would desire a credible commitment. To show how the timing of policy-setting influences outcomes, we will consider each of the possibilities, starting with the simplest cases.

If there is no possibility to change the policy decisions made in Period 1, then voters will unanimously favor the first-best policies, which would then be implemented. Thus, we can state:

> Implication 1: When voters can commit to policy while still behind the veil, the first-best outcome will be obtained.

As we will argue, the idea of policy being set behind a veil is relevant to the real world because, in some cases (e.g., writing laws to be enforced in the long run, with the specific circumstances under which those laws will be enforced not known in advance), it is useful to think of policies being locked in before people know whether they will be on one side or another of a potential conflict (such as a trial).[13]

Now consider modifying the scenario just discussed so that voters in Period 3 (but not Period 5) have (i) the ability to change the trade policy, (ii) the ability to change the public good policy, or (iii) the ability to change both policies. In each of these three cases, the first-best will be obtained.

---

[13]See Fleck and Hanssen (2013b) on the importance of distinguishing cases in which constitution-style laws are written behind a veil from cases in which constitution-style laws are written after the veil has been lifted.

To understand why, recall that the pre-trading holdings of A are (by assumption) uniform across the population, and that means that trading A at the market-clearing price will result in a uniform distribution of the net benefits of setting g to its the optimal value. To summarize:

> Implication 2: When voters have equal endowments of the tradable asset (and full information), being able to commit to public good policy in the period before trade occurs will lead to the first-best outcome.

This, in turn, indicates that homogeneity in asset holdings can – as Aristotle suggested – be valuable for aligning preferences over public good provision.

Two further scenarios remain for us to analyze. In both of these, the voters in Period 5 – which is after trade (if allowed) occurs – have the ability to change the public good policy, and that may prevent a credible commitment to the optimal public goods policy. There are two scenarios to analyze here because we need to consider how outcomes vary based on whether or not voters in Period 3 (i.e., after the veil has been lifted) have the ability to change trade policy. Either way, however, a barrier to ideal policy can arise because voters who have heterogeneous assets holdings will have heterogeneous returns to their (assumed equal) contributions to the provision of the public good. To see why, note that at the time the choice between g=1 and g=0 is finalized, setting g=1 will be more attractive (or, equivalently, less unattractive) to type H individuals than it will be to type L individuals.

Under some (but not all) conditions, this will preclude first-best outcomes. In particular, if the first-best has g=1 (i.e., if $\gamma \underline{A} > \kappa$), then the first-best (g=1; trade allowed) will be obtained if one or both of the following hold: type L are a minority (so that type H are decisive); type L favor g=1

even after trade occurs.[14] Similarly, if the first-best has g=0 (i.e., if $\gamma\underline{A} < \kappa$), then the first-best (g=0; trade allowed) will be obtained if one or both of the following hold: type H are a minority (so that type L are decisive); type H favor g=0 even after trade occurs.[15]

This leads us directly to the cases in which the first-best is infeasible. More specifically, if the public good policy can be changed in Period 5, then for the cases not listed in the previous paragraph, the first-best choice of g will be blocked. The reason is that, subsequent to trade, the majority will favor the non-optimal choice of g.[16] To see why, note that if the first-best has g=1, but a post-trade type L majority would vote for g=0, the first best will be precluded. Similarly, if the first-best has g=0, but a post-trade type H majority would vote for g=1, the first best will be precluded. To summarize:

> Implication 3: When voters have the ability to set public good policy after trade has occurred, trade-generated heterogeneity in asset holdings may preclude voter support for first-best policies and, thus, preclude first-best outcomes.

D. Optimal Policy (i.e., "Second-Best") when the First-Best is Infeasible

When the first-best outcome is infeasible, rational voters will choose the best feasible option. In particular, voters will evaluate and choose one of the following two options: allow trade knowing that doing so will lead the majority subsequently to set the non-optimal g; disallow trade knowing that doing so will lead the majority subsequently to set the optimal g. The key tradeoff is, therefore,

---

[14]More formally: Type L will be a minority if and only if $\omega_L < \omega_H$. Type L will favor g=1 after trade occurs if and only if $\gamma[\underline{A} - [(\alpha_H - \alpha_L)\omega_H]/\beta] > \kappa$.

[15]More formally: Type H will be a minority if and only if $\omega_L > \omega_H$. Type H will favor g=0 after trade occurs if and only if $\gamma[\underline{A} + [(\alpha_H - \alpha_L)\omega_L]/\beta] \leq \kappa$.

[16]Recall that if trade is allowed, the optimal amount of trade will occur regardless of the choice of g. Thus, when the decision regarding g occurs after trade takes place, type L individuals will each have $\underline{A} - [(\alpha_H - \alpha_L)\omega_H]/\beta$ units of A, and type H individuals will each have $\underline{A} + [(\alpha_H - \alpha_L)\omega_L]/\beta]$ units of A.

obtaining gains from trade versus having the optimal level of the public good.

If the order of events for a second-best scenario is such that voters are behind the veil when deciding whether to allow trade (i.e., trade policy cannot be changed in Period 3), the voters will decide unanimously what to do: allowing trade if and only if the gains from trade at least equal the (forgone) gains from setting g optimally.[17] Either way this decision goes, the outcome will be the best one feasible. In other words, conditional on the incentive compatibility constraints, the outcome maximizes net social benefits.

If the order of events for a second-best scenario is such that the veil has already been lifted when voters decide whether to allow trade (i.e., trade policy can be changed in Period 3), the vote on whether to allow trade may be unanimous, but it may not be. For the case with sufficiently large benefits from setting g optimally (i.e., $|\gamma \underline{A} - \kappa|$ is sufficiently large), both types of citizens will agree that it is worth forgoing the gains from trade in order to obtain the optimal g. Similarly, if the benefits from setting g optimally are sufficiently low (i.e., $|\gamma \underline{A} - \kappa|$ is sufficiently small), then both types of citizens will agree that it is worth forgoing the optimal g in order to obtain the gains from trade. Thus, even when citizens are out from behind the veil, they may be in complete agreement to prohibit trade or to allow trade. Yet there can be disagreement at that stage, because the gains from trade are more valuable to each member of the minority type than to each member of the majority type.[18] Thus, the majority may vote (against the minority) to disallow trade even if the total

---

[17]More formally, trade will be allowed if and only if: $(.5n/\beta)\ (\omega_H\ \omega_L)\ (\alpha_H - \alpha_L)^2 \geq |\gamma \underline{A} - \kappa|n$.

[18]The per capita gains from trade for a type H individual i are $T_{Hi} = .5\beta\ [(\alpha_H - \alpha_L)\omega_L]^2$, and the per capita gains from trade for a type L individual j are $T_{Lj} = .5\beta\ [(\alpha_H - \alpha_L)\omega_H]^2$. The intuition here is that the smaller group necessarily trades more A per capita and, thus, given our assumption that β (the slope of the marginal benefit curve) is invariant between the groups, has greater per capita gains from trade.

(i.e., aggregated over all citizens) forgone gains from trade would exceed the costs of having the non-optimal g. In sum, the majority may vote (with or against the minority) to incur costs (i.e., forgo gains from trade) in order to prevent itself from voting undesirably (i.e., for the non-optimal g) in the future. To summarize the key tradeoffs:

> Implication 4: If the optimal (i.e., trade-generated) allocation of asset holdings would preclude voter support for the optimal choice of g, voters face a tradeoff: obtaining gains from trade versus having the optimal level of the public good. In the presence of sufficiently large potential gains from trade (equivalently, choosing g optimally is sufficiently unimportant), voters will unanimously favor allowing trade, even though it will lead to the non-optimal g. Similarly, if choosing g optimally is sufficiently important (equivalently, the potential gains from trade are sufficiently small), voters will unanimously favor disallowing trade, because homogeneity in asset holdings will lead to the optimal g.

E. Aristotle Versus Lindahl in the Presence of Distortionary Redistribution

In what follows, we extend our analysis to allow for cases in which aligning incentives requires costly redistribution. By doing so, we can move in the direction of realism for the task of identifying policies that generate "second-best" outcomes.

Aristotle in our theoretical framework: Engineer homogeneity

Although the theoretical analysis presented above assumes a homogeneous citizenry as a starting point (i.e., initially equal asset holdings, voters temporarily behind a veil, equal cost-sharing for public good provision), the logic of the model applies much more broadly. Consider our fundamental point that it may be beneficial to maintain less than the first-best degree of heterogeneity in asset holdings. Because our analysis above started with a uniform allocation of assets (Assumption 1), we examined the potential benefits of using prohibitions on trade to enforce more-than-market levels of homogeneity. For the same reasons, if the starting point is an initially heterogeneous allocation of assets, it may be beneficial to legislate a reallocation of assets – even

if the initial heterogeneity resulted from a well-functioning market (i.e., from a market that allocates assets to their most valuable uses). Indeed, our theoretical framework shows the following:

> Implication 5: When assets are initially allocated in the first-best manner, the equalization of asset holdings, even if distortionary and uncompensated, may benefit everyone, including those from whom assets are taken.

The reason is that an individual's loss of assets may be more than offset by the improvement in public good policy made possible by the effects of redistribution on voters' incentives.[19]

Of course, the real world is more complex, and policies need not be so extreme as an uncompensated equalization of asset holdings. The critical point is to recognize tradeoffs. Consider, for example, a policy that allocates parcels of land (which are more valuable when well defended) to the soldiers who contribute to national defense. This may be desirable even if it creates costs (e.g., deadweight losses associated with redistribution). But it will not be desirable if those costs are so large that they offset completely the benefits derived from voters having better aligned incentives.

Lindahl in our theoretical framework: Taxation in proportion to benefits

To align voters' incentives for public good provision, an alternative to engineering

---

[19]To see this more formally, consider the case where individuals of type H are a minority (i.e., $\omega_H<\omega_L$) and hold assets in quantities corresponding to the market equilibrium. An uncompensated equalization of asset holdings will cause a net loss of T (i.e., the gains from trade when the starting point is a uniform asset distribution). In other words, redistribution creates a deadweight loss equal to T. (Recall that T can be expressed as a function of the exogenous parameters: $T = (.5n/\beta)\,(\omega_H\,\omega_L)\,(\alpha_H - \alpha_L)^2$.) Type H individuals will benefit from such an uncompensated reallocation if their gains from setting the optimal g exceed their losses from having their assets taken. More formally, the condition is:

$$|\gamma \underline{A} - \kappa| > .5\beta\,[(\alpha_H - \alpha_L)\omega_L]^2 + [(\alpha_H - \alpha_L)\omega_L]/\beta]\,P^*, \text{ where } P^* = \alpha_H\,\omega_H + \alpha_L\,\omega_L - \beta A_i + \gamma g$$

Note that if compenstion were paid based on the market price (P*), the condition guaranteeing that type H individuals would benefit from the compensated equalization of asset holdings (relative to the market allocation of assets) would be: $|\gamma \underline{A} - \kappa| > .5\beta\,[(\alpha_H - \alpha_L)\omega_L]^2$. Similarly, type L individuals would benefit from the compensated equalization of asset holdings (where type L individuals are paying the compensation) if $|\gamma \underline{A} - \kappa| > .5\beta\,[(\alpha_H - \alpha_L)\omega_H]^2$.

homogeneity is to fund the public good in a manner so that individuals' cost shares match their benefit shares, as in the Lindahl prescription. As noted earlier, the informational requirements for Lindahl taxation render it better in theory than in practice. Nevertheless, our model can provide some practical guidance.

As a preliminary point, consider the case in which Period 1 voters can commit to a cost-sharing system specific to individuals' types, so that Period 5 voters choose g based on the cost-sharing system chosen behind the veil. For the two types of individuals in our model, the incentive-aligning system for funding g=1 would be $\tau_L = \kappa/[(\omega_H A_H/A_L) + \omega_L]$ and $\tau_H = \kappa/[(\omega_L A_L/A_H) + \omega_H]$, where A_L and A_H are market-determined and can be expressed in terms of the exogenous parameters: $A_H = \underline{A} + [(\alpha_H - \alpha_L)\omega_L]/\beta$ and $A_L = \underline{A} - [(\alpha_H - \alpha_L)\omega_H]/\beta$.[20] This would generate a first-best outcome, exactly in line with the Lindahl prescription, but it would require the ability to set tax rates based on individuals' types, as well as a credible commitment to do so.

An alternative would be to tax individuals in proportion to their holdings of A. Doing so may or may not allow an outcome equivalent to imposing the type-based tax rates (τ_L and τ_H) discussed in the previous paragraph. The potential equivalence depends on how a tax on A affects the quantity of A. Recall that, in our model, deadweight losses result from time-conconsistency problems, not distortionary taxation, because taxation takes the form of equal per capita contributions (κ per citizen). Thus, if the supply of A were perfectly inelastic (i.e., the total quantity of A remains fixed at $n\underline{A}$), a tax in proportion to asset holdings (i.e., $\tau_i = \kappa(A_i/\underline{A})$, where A_i indicates individual i's

[20]Note that that τ_H and τ_L satisfy the revenue requirement ($n\omega_H\tau_H + n\omega_L\tau_L = n\kappa$) in a manner that aligns tax rates with asset ownership ($\tau_H/\tau_L = A_H/A_L$). Also note that with full information about which individuals are of which type, taxes can be set as functions of the exogenous parameters given above, rather than as functions of tax-influenced asset holdings.

assets) would be exactly equivalent to τ_L and τ_H and, thus, generate a first-best outcome exactly in line with the Lindahl prescription.[21]

In a more realistic scenario, where A has a positively sloped supply curve, funding the public good with cost shares in proportion to asset holdings will cause a deadweight loss and, therefore, cannot be first-best, but it may allow a second-best solution. To illustrate this, we will introduce a new assumption:

Assumption 6: Distortionary Taxation of Assets: In the market for assets, let σ denote the slope of the supply curve (with $\sigma>0$ and the zero-tax equilibrium quantity supplied being $n\underline{A}$), and let τ_A denote a per-unit tax (the smaller one if two exist) on A sufficient to generate the revenue required to fund g=1.[22]

For a given set of parameter values, a ceteris paribus increase in σ (i.e., making the supply of A less elastic) would reduce the deadweight loss created by the tax, because it will increase the equilibrium quantity of A and reduce the tax rate necessary to fund g=1. As σ approaches infinity (i.e., perfectly inelastic supply), the deadweight loss approaches zero, with τ_A approaching κ, and the outcome approaching the first-best. This yields:

> Implication 6: When the first-best is precluded yet voters can commit to a (distortionary) tax on A, there will be a range of σ where supply is sufficiently inelastic that setting tax rates in proportion to asset holdings will yield a second-best outcome – that is, an outcome preferable to enforced homogeneity.

In the real world, of course, the ability to employ this type of mechanism will depend on the

---

[21]Note that this satisfies the revenue requirement ($\Sigma_i \tau_i = n\kappa$) in a manner that aligns tax rates with asset ownership ($\tau_i/\tau_j = A_i/A_j$).

[22]Following the basic logic of supply and demand: Because the equilibrium quantity declines in response to higher tax rates, there may be two tax rates that raise total tax revenue equal to $n\kappa$ (i.e., the amount required to fund g=1). Naturally, the preferable rate would be the lower of the two. Note that assuming the zero-tax equilibrium quantity supplied equals $n\underline{A}$ corresponds to the case when setting g=1 increases by γ the non-taxed marginal benefits of the non-taxable use of the asset. This is logically consistent with our earlier treatment (in Assumption 1) of the equilibrium quantity of the asset being independent of the level of g.

feasibility of establishing asset-based taxes in a manner that voters will not undo.

Aristotle versus Lindahl (again): Newly acquired assets

Now that we have a framework for considering deadweight losses from taxation, consider the relevance of Aristotle's prescription for assets again. The key issue here is to examine the relevance of the citizenry having a stock of newly acquired assets (such as conquered land and/or captured workers) to be divided among the citzenry. How would this differ from the case of having assets already distributed among the citizenry in the first-best manner? Even if dividing the newly acquired assets can be done without real costs (i.e., no deadweight losses from conflict or taxation), maintaining an equal division of those assets would still generate a deadweight loss corresponding to T; funding the public good would not create a deadweight, because a tax of κ per citizen (i.e., a head tax) could be employed. Thus, the "Aristotle versus Lindahl" comparison here, in the context of our model, is T (for Aristotle's prescription) to the deadweight loss of a distortionary Lindahl-style tax levied on A.[23] Yet if the stock of A is initially allocated in among the citizenry in the first-best manner, the equalization of asset holding via taxation would create a deadweight loss, thus reducing the value of the Aristotle prescription (what we call "engineering homogeneity") relative to the altnernatives: the Lindahl-like tax or having a non-optimal choice of g. The punchline is that a large-scale acquisition (e.g., discovery, conquest, or theft) of divisible assets increases the relative attractiveness of Aristotle's recommendation.

F. Practical Implications

The fundamental lesson from our theoretical model is that designing a successful democracy involves tradeoffs related to voters' incentives. To see how the model can apply to the real world,

[23]As noted earlier, the Lindahl-style tax is such that $\tau_i/\tau_j = A_i/A_j$ for all i and j, with $\Sigma_i \tau_i = n\kappa$.

it is essential to recall that our model relies on a time-inconsistency problem to generate those tradeoffs (Implication 1). This makes sense as a way to develop a model with real-world applicability: In the real world, as in the model, if it were possible, say, to write first-best policies into a credible constitution – or otherwise commit credibly – at a time when voters favored a commitment to those first-best policies, that would be a way around the tradeoffs our model highlights.[24] Put another way, our model's key implications pertain to cases in which voters lack the ability to employ a low-cost way to establish and enforce first-best policies.[25]

Another critical point to recognize is that we intend Implications 5 and 6 to serve as useful abstractions, not as descriptions of distinct types of public policy. For example, when modern democracies mandate schooling for children and provide tax revenue to fund schools, these policies will, if successful, help to equalize endowments of human capital in a manner relevant to our model (especially Implication 5) – even if that is not the main objective of those setting policy. Similarly,

---

[24]There is a large literature, for example, on the role of courts and other "counter-majoritarian checks" as mechanisms to constrain voters away from undesirable majority-rule outcomes (e.g., Glaeser, Johnson, and Shleifer 2001; Aghion, Alesina, and Trebbi 2004; Hanssen 2004; and Maskin and Tirole 2004). On the role that modern courts can (and cannot) play in addressing tyranny of the majority problems similar to voters' inability to commit to the optimal g in our model, see Fleck and Hanssen (2013b). The ancient Greeks faced problems similar to those of modern democracies, but designed very different legal institutions (Fleck and Hanssen 2012); for additional for economic analysis applied to Athenian institutions, see McCannon (2010a, 2010b) and Carugati, Hadfield, and Weingast (forthcoming). Another way to constrain voting outcomes is delegating power to a subset of the citizenry (or disenfranchising a segment of the population). Democracies have always had restrictions on the franchise, and one reason to expand or contract the set of voters is to influence future policy decisions; see, for example, Acemoglu and Robinson (2000, 2001), Fleck (2000), Lizzeri and Persico (2004), Llavador and Oxoby (2005), and Fleck and Hanssen (2006, 2013a).

[25]Our model can easily be extended to illustrate this point more formally. For example, the expected value of a counter-majoritarian check would depend on (i) how much making desirable policy decisions matters (indicated by T and by $|\gamma\underline{A}-\kappa|n$ in our model), (ii) how much the counter-majoritarian check reduces the probability of implementing an undesirable majority-supported policy, and (iii) how much the counter-majoritarian check increases the probability of failing to implement a desirable majority-supported policy.

the idea of taxing productive assets, especially those that are inelastically supplied and easily observable, has a long history (both as real-world sources of revenue and as examples in introductory economics textbooks). Such taxes may or may not be earmarked for public goods (or other types of government goods and services) complementary to the taxed assets – to the extent that they are, we expect our model (especially Implication 6) to be useful. Also note that the task of enforcing commitments to education, tax, and other policies has often been delegated to courts, especially when potentially temporary support among voters may create time-inconsistency problems for a purely majoritarian system (e.g., Fleck and Hanssen 2013b). In sum, the ideas in Implications 5 and 6, along with efforts constrain majority rule decisions, may function more as complements than as substitutes.

III. THE TIMELINE OF EVENTS IN ANCIENT GREECE

Before applying the model, it will be useful to provide a very brief overview of the events that transformed ancient Greece. Figure 1 provides a timeline that begins with Greece's Archaic Period, when the organization of the *polis* solidified and the Greeks re-integrated themselves into the trading networks of the Mediterranean, and finishes with the end of the Classical period, when Greek institutions were permanently altered following the invasions of Philip of Macedon and his son, Alexander the Great.[26] Events that brought fundamental changes in the wealth creation process

[26]Samons (2007, 4) writes, "The origins of the polis are controversial, but this form of settlement was firmly established by the eighth century BC. A typical polis comprised a town or city center (*astu*) surrounded by land (*chora*) farmed and owned by the polis's citizens." Scholars typically divide the history of Ancient Greece into four periods: the Mycenaean period (1600-1150 B.C.E.), the Dark Ages (1150-800 B.C.E.), the Archaic period (800-490 B.C.E.), and the Classical period (490-323 B.C.E.). The Mycenaean period was characterized by highly centralized, highly bureaucratic palace economies, similar to those seen in Crete at Knossos (and in Egypt and other near Eastern civilizations). The reason for the collapse of

are listed below the timeline; changes in institutions are listed above. In Sparta, institutional changes followed both its conquest and its loss of Messenia, a neighboring *polis* that possessed the largest and most fertile grain fields in the Peloponnese. Athens underwent two institutional redesigns, the first when its tribute-paying empire was consolidated, the second when the empire was dissolved following Athens' loss in the Peloponnesian War.

IV. INSTITUTIONAL DESIGN IN SPARTA AND ATHENS

In both Sparta and Athens, the citizens faced tradeoffs similar those illustrated in our model, and we will here discuss how this affected the design of the institutions for which the *poleis* are now famous. We will begin with the Spartans, who "engineered homogeneity" by assigning citizens very equal shares of the benefits and costs related to the state's main source of wealth – in the language of the model, Sparta established a system that provided each citizen with $\underline{A}$ units of the asset. We then discuss Athens, which faced a more complicated design problem: How to align benefits and costs, given that Spartan-style homogeneity would have impeded the enormous gains that Athens, a commercial powerhouse, obtained from allowing citizens to specialize, with respect to investments in human capital and other productive assets. The Athenian solution was analogous to the Lindahl-style mechanism illustrated in our model – broad-based taxes and broadly redistributive transfer

---

Mycenae remains a mystery; for background on Mycenae and its collapse, see Murray (1993). In the "Dark Ages" that followed (a time of "poverty, isolation, and illiteracy" according to Manville 1990, 35), the centralized palace bureaucracy that had controlled much of life vanished without a trace. Throughout Greece, precipitous declines in population occurred, estimated at 60 to 90 percent, depending on the region. Most Mycenaean sites were abandoned, and little collective memory of the earlier period appears to have survived, so that the Dark Ages represent a sharp break with the preceding era (see, e.g., Freeman 1999; Pomeroy et al. 1999, 41). At the end of the 9th century B.C.E. (the beginning of the Archaic period), the population began to grow again, reaching rates of two to three percent per year by the early 8th century; see Hanson (1999, 36) and the citations therein. Settled cultivation replaced the largely pastoral Dark Age economies, and *poleis* began to form.

payments allowed costs and benefits to be shared widely without equalizing asset holdings. That said, the Athenians also used institutional design to promote homogeneity at the group level, in order to align voting incentives without blunting individual initiative.

A. Sparta

Given that our model has voters deciding whether to restrict their own ability to engage in trade, an essential starting point is to recognize Sparta's contrast with some "revolutionary" regimes of the modern era. In regimes where leaders have attempted (or at least claimed to have attempted) to minimize differences among citizens, the consequences have often been disastrous (e.g., the Cultural Revolution in China, the Khmer Rouge in Cambodia). Sparta's experience is distinguished by the fact that the overt, state-sponsored suppression of differences was designed and carried out by the very citizens upon whom it was imposed (not a "vanguard" of them), and served to undergird the development and durability of one of the most powerful *poleis* of ancient Greece.

The source of wealth: Fertile land and captive labor

Much of classical Sparta's power was financed by the grain production of a neighboring territory, Messenia, which Sparta conquered circa 700 BCE.[27] The Messenians rebelled about thirty years later in what is known as the Second Messenian War, probably timing their revolt to correspond to a Spartan military defeat by Argos.[28] At great cost, Sparta crushed the Messenians and

[27]Semple (1921, 55) writes that the Messenian grain fields "enjoyed a rare reputation for productivity from very early days." The reasons for the initial attack on Messenia are uncertain, but it occurred at a time when many *poleis* were sending citizens to found colonies; some scholars have referred to the Messenian conquest as internal colonization (Forrest 1968, 38).

[28]See, e.g., Murray (1993, 157). Scholars believe that in the aftermath of the initial war, the Messenians may have had a status comparable to the *perioikoi* ("those who live around"), who inhabited other parts of Sparta. The *perioikoi* possessed no political rights and were required to provide assistance to the Spartan military, but faced no other obligations, and were left to govern their own communities in an autonomous fashion (except with respect to foreign policy). See Cartledge (1987, 16).

reduced them to serfdom.[29] The victory provided the Spartans with two very valuable assets – the most fertile land in the region and a captive labor force to work it – that would generate wealth only if the Spartans incurred the very large cost of maintaining control of, and extracting rents from, them. In the context of our model, one can view g=1 as the collectively optimal (for Spartans) use of an extremely strong Spartan military – a public good that yielded substantial net gains ($\gamma\underline{A}-\kappa$)n for Spartans, but that required large costs κn in order to obtain even larger benefits $\gamma\underline{A}$n.

Sharing the benefits and costs

Sparta's response to the victory was a wholesale re-design of its institutions. Scholars dating back to Thucydides have attributed Sparta's unique military-oriented institutions to the conquest of Messenia, and the institutions make particular sense when viewed through the lens of our model.[30] Spartans began by dividing the gains from the conquest with remarkable equality, splitting Messenia into similar sized estates and providing one to each Spartan citizen. In the context of our model, when dividing up Messenian land (and captive labor), each Spartan received an endowment

[29]The Messenian serfs were known as helots; Forrest (1968, 31) suggests that the word "helot" derives from the root *hel-*, implying seizure or capture. The helots had no political rights and no freedom of movement, but were bound to particular parcels of land, from which they were required to provide a portion (probably half) of their output to the Spartan masters. Helots belonged to the Spartan state – individual helots could not be bought and sold, although the parcels on which they worked may have changed from one Spartan owner to another from time to time (see Hodkinson 2000, 124). In addition to the Messenian helots, there were also helots located in Laconia, the site of Sparta itself. However, the vast majority of helots were in Messenia; for this reason, the helots collectively were often referred to as "Messenians" (Cartledge 1987, 15).

[30]Thucydides wrote that "most institutions among Spartans have always been established with regard to security against the helots" (*Hist.*4.80.3). Cartledge (2001, 89) sums it up as follows: "[I]t was the helots who, by freeing the Spartan citizens en bloc from all productive labor (other than warfare), enabled their masters to develop their uniquely military society, a workshop of war. But at the same time it was also the helots who so outnumbered the Spartan citizen population . . . who, as the enemy within, 'lying in wait for their masters' in the phrase of Aristotle (*Pol*. 1269a37-9), necessitated as well as enabled the Spartans military mode of life, and their unique transformation of a *polis* [city-state] into a military-police state."

analogous to A. And, of course, this is just what Aristotle had recommended.[31]

This unusually equal allocation of assets corresponded to a very equal division of the corresponding costs.[32] The Second Messenian War had made clear that the land could be held only by military force.[33] The war was followed by a series of institutional changes, the Lycourgan reforms (called also the Great *Rhetra*, or statement).[34] Finley (1975, 175-6) notes that these reforms contain features found in the constitutions of many other *poleis* (that of Athens, for example), but combined in Sparta to produce a unique system: "a state ruled by an elite of Spartiates whose prime concern was military preparedness."[35]

---

[31]Sparta was presumably the inspiration for the Aristotle quotation that opens our paper. As time passed, inequality may have been increased (see, e.g., Hodkinson 2000).

[32]Before the conquest of Messenia, there was little in the Spartan system that set it off from the systems used anywhere else in the Greek world. Murray (1993, 155) writes of the early Spartans, "They seem originally to have differed little from other early Greek communities. Their political constitution was normal in basic structure." The Dorian origins of the Spartans have been used by some to explain Sparta's unusual militarism; however, few scholars today accept this explanation (for one thing, other "Dorian" states, such as Argos, never developed institutions like those of Sparta).

[33]Xenophon (*Hell*.3.3.4-11), an Athenian who frequently visited Sparta, said of the Messenian serfs, "They would gladly eat the Spartans raw." The Messenians rose in rebellion numerous times over the centuries that followed the Spartan conquest, often timing their revolts to coincide with other problems faced by Sparta. For example, the Messenians revolted when Sparta suffered a major earthquake in 465 B.C. (Freeman 1999, 198). The threat of revolt was serious: Although the Messenians lacked weapons, they outnumbered the Spartans by as many as ten to one. Modern estimates place the total number of Messenian serfs in the range of 200,000, while Sparta at its peak had no more than 10,000 male citizens (Cartledge 1987, 174). One of the more conservative guesses was by Herodotus, who estimated that the Messenians outnumbered the Spartans on an order of seven to one (quoted in Jameson 1992, 136).

[34]The dating is somewhat imprecise; see, e.g., Forrest (1968, 55-8), Dillon and Garland (1996, 147), and Murray (1993, 165-73).

[35]Sparta made an annual declaration of war on its Messenian helots, regardless of the actual state of affairs at that moment, both to symbolize the underlying nature of the relationship and to allow a quick response to provocations. Because of the declaration, a Spartan suffered no "blood-guilt" (and faced no punishment) if he killed a helot, regardless of the circumstances under which the killing took place (Cartledge 2003, 73).

Spartans saw the maintenance of homogeneity among the citizenry as essential to the success of their society. Indicating the Spartan ideal, they referred to themselves as *homoioi* ("the equal ones" or "the similars"), and they took that ideal seriously.[36] The process of reducing heterogeneity began at birth, when all male infants were examined by a public commission to determine whether they should be allowed to live.[37] At the age of 7, each Spartan boy was taken from his family and enrolled in the *agoge*, a system of public education that required him to live, eat, and sleep with other boys until he reached the age of 20.[38] From age 20 through age 60, all Spartan males trained continually, living communally until marriage (at roughly age 30) and eating communally (in military messes known as *sussitae*) until age 60. Whereas citizens in most Greek *poleis* engaged in a variety of occupations (farming, trading, crafts, manufacturing), taking up arms temporarily as need dictated, Sparta restricted citizenship exclusively to those able to engage in *full-time* soldiering.[39] Sparta was the only *polis* in Greece to maintain a full-time army.

---

[36]See, e.g., Freeman (1999, 97) and Hanson (1999, 385).

[37]Plutarch (*Lyc*.16.1-2) writes, "The father did not decide whether to raise a baby; rather he took it and carried it to some place called Lesche where the elders of the tribe sat and examined the infant, and if it were well-built and sturdy, they ordered the father to rear it."

[38]The *agoge* (meaning "upbringing") separated those aged 7 to 17 (the *paides*, or "boys") from those aged 18 to 20 (the *paidiskoi*, or "youths"). For both groups, however, the physical aspect of the training was rigorous, designed to develop tough, brave men with the ability and mental preparedness to attack and kill Messenians. The youngest boys were required to go barefoot in all seasons, and from ages of 12 through 17, systematic surveillance and discipline by older boys was the norm; Cartledge (2001, 86) writes, "this second stage resembled nothing so much as a paramilitary assault course." An elite few of the older boys were assigned to a group known as the *Kryptoi*, and, armed only with knives, sent to Messenia and required to survive on their own (primarily by robbing and killing).

[39]Hanson (1999, 301) writes, "Outside of Sparta, hoplites [Greek infantrymen] spent little time training for war." By contrast, in Sparta, as Hanson (1999, 385) writes, "Males did not farm. They trained constantly for battle." Plutarch recounts a Spartan general calling out to his combined Spartan-allied army various trades and requesting each to rise until nearly the whole crowd was standing, "but of the Lacedaemonains [Spartans] not one. For they were forbidden to work at a craft and to learn servile occupations" (quoted in Anderson 1970, 243).

Moreover, citizenship required meeting criteria that also kept the citizenry's incentives well aligned.[40] Notably, a Spartan who fell below the required wealth level, or participated in a non-military activity such as commerce, was stripped of his citizenship.[41] This feature of the Spartan system matters for our argument because it shows that Spartans had concerns beyond just fielding a large army of good soldiers – they took the homogeneity seriously enough to exclude potential citizen-soldiers from their ranks and to forgo gains from exchange that could have generated gains that, at least in principle, could have been used to strengthen the military.[42]

To sum up, the Spartans engineered a system that looks very like the "promoting homogeneity" strategy of the model. First, with citizens having similar shares of land, similar shares of captive labor attached to the land, similar human capital (as soldiers), and highly restricted opportunities to engage in commerce, the Spartans engineered a system that had something extraordinarily close to each citizen owning $\underline{A}$ units of productive capital. Second, with the main public good being defense, the nature of the Spartan military organization meant that the total cost

---

[40]All citizens belonged to the assembly, in which they all had equal say. Sparta also had an advisory council, known as the *Gerousia*, whose membership was drawn exclusively from certain important families, and a dual kingship – Aristotle called the Spartan kingship "hereditary generalship for life" (*Pol*. 3.1285b). (Athens, by contrast, elected its generals to annual terms.) See, e.g., Forrest (1968).

[41]See, e.g., Freeman (1999, 97) and Hanson (1999, 385).

[42]Sparta was also unique in the way it treated women. Spartan women could inherit property, and, as a result, they held a subtantial amount of Sparta's wealth (e.g., Pomeroy 2002). Fleck and Hanssen (2009) provide an economic analysis of women's rights under the Spartan system, and argue that women's rights played an essential role in providing incentives to manage estates. Note that this, combined with assortative marriage (i.e., wealthy Spartan women marrying wealthy Spartan men) probably made it more difficult to maintain of relatively evenly distributed wealth – but that, in turn, is another reason for Spartans' to have such great concern with homogeneity. It is also important to note that (i) Spartan girls attended public schools, (ii) Spartan women were expected to engage in activities (such as managing estates) that helped Sparta maintain control and extract rents from the helots, and (iii) Spartan women were discouraged from engaging in many activities (such as commerce) that might have shifted the efforts in other directions. As a result, it was not just Spartan men, but Spartan households, that held similar endowments of human capital.

of providing that public good was shared remarkably equally, analogous to each citizen in our model bearing a cost κ in order to provide g=1. Thus, when it came to decisions regarding the principal public good, Spartans had well aligned incentives to support policies that would maximize the net gains for Spartans. Thus, the Spartans could – and did – rely on a majority rule system that allocated power in an exceptionally egalitarian manner.

What happened when source of wealth changed?

If the way we have thus far applied our model to Sparta is correct, we should expect to see that the eventual loss of Sparta's main source of wealth brought about a redesign of Sparta's political institutions. When Sparta lost a crucial battle at Leuctra in 371 BCE, its control of Messenia ended. This changed the benefits and costs of maintaining homogeneity, thus giving the Spartans an incentive to adopt institutions more like those in other *poleis*. Is there evidence of institutional redesign along the lines of what our model would predict? Spartan history in the decades between its loss of Messenia and the final Macedonian conquest of Greece is poorly documented; however, what can be determined suggests that Sparta altered its institutions.[43] First, there is some evidence of post-Leuctra reform in the system of landholding (specifically, an abandonment of rigid inheritance rules).[44] Second, it has been suggested that Sparta modified its rules regarding serving as soldiers (Anderson 1970, 229-51). Third, the strict prohibition against non-military activities ceased to be enforced; Cartledge and Spawforth (2002, 15) write, "we now hear for the first time of Spartans turning their own hands perforce to the plow."

---

[43]Forest (1968, 1142) writes, "No connected account of Spartan history is possible for this period." Cartledge and Spawforth (2002, 9) write that "The lack of . . . a competent narrative account of Spartan and Greek history between 359 and 338 is lamentable."

[44]Cartledge (1979, 167-9; 1987, 167).

Thus, after losing Messenia, Sparta became more like the other *poleis* of the Peloponnese, allowing substantial heterogeneity in among asset holdings (including human capital investment, as well as land). In the context of the model, the institutions that had kept the key assets uniformly distributed (analogous to $\underline{A}$ for each citizen), and thus voters' incentives so closely aligned, were no longer worth the opportunity cost of lost gains from trade (including, of course, specialization of labor market).[45] Thus, the dismantling of Sparta's highly unique institutions corroborates our model's explanation of the rise of those institutions.

B. Athens

Athens also faced tradeoffs similar to those illustrated by our model. Unlike Sparta, Athens depended heavily on commerce, especially seagoing trade, and this required specialization. Thus, in the context of the model, while Sparta chose something similar to "disallowing trade" (forcing homogeneous holdings of land, captive labor, and the citizenry's own human capital) in order to obtain aligned votes on public good provision, such an option would have been too costly for Athens – the potential gains from specialization were too great to forgo.

That Athens should have emphasized trade to a greater extent than Sparta is not surprising. It lacked the fertile agricultural land that Sparta controlled; it imported grain and exported oil from olives grown on hillsides (see Fleck and Hanssen 2006 for discussion and citations). Furthermore,

[45]The Spartans also developed a new source of wealth: Acting as mercenaries for rulers about the Mediterranean. Forest writes, "To mend the poverty Spartans from time to time would use the mercenary market to their advantage and even Spartan kings appear in foreign service. Agesilaos [a Spartan king] died in 360 returning from two years as a mercenary captain in Egypt with 230 talents for the Spartan treasury." Although mercenaries require military skills, the old Spartan institutions no longer would have been optimal, because the gains from specialization would presumably have been very large. In other words, while some citizens had a comparative advantage in providing mercenary services (others in commerce, management, etc.), the old institutions would not have precluded an allocation of human capital (and other resources) in a manner that made use of comparative advantage.

it and had an enviable position on the coast, well-placed to access the most thriving parts of the Mediterranean (the Black Sea region, Anatolia, the Middle East, and Italy). Nonetheless, it was far from inevitable that Athens would become ancient Greece's predominant commercial power. Indeed, at the start of the Classical period, Athens was engaged in fierce competition with the nearby – and relatively small – island of Aegina for *local* commercial preeminence.

For Athens, commercial dominance followed the establishment of institutions that, like Sparta's, were unique in the Greek world. To maintain its source of wealth, Athens (like Sparta) had to ensure that its voters would remain willing to devote a sufficient share of their resources to the production of public goods, most importantly the Athenian navy.

Tribute as the source of wealth

From 478 to 404 BCE, Athens headed an empire, known as the Delian League, founded in response to the threat posed by the Persian empire.[46] Athens earned substantial wealth from its empire, in the form of tribute payments from member states intended to finance the League's navy and, thereby, the safer seas that allowed trade to flourish. Because Athens provided virtually all the

[46]The Greek world at the time extended from as far west as Spain to as far east as the shores of the Black Sea. Many of the *poleis* to the east of Athens – in particular, those located on Aegean islands and the Anatolian coast – fell under the rule of the Persians in the mid-sixth century BCE. In 499 BCE, a number of them revolted, aided by Athens and several other mainland Greek *poleis*. The revolt failed and the two famous Persian invasions of the Greek mainland were launched in retaliation. The Greeks twice repelled the Persians, who never again attempted to invade the Greek mainland. However, the Persians remained a threat to the eastern *poleis*, who formed a defensive alliance in response, placing themselves under the command of Athens and its powerful navy. The alliance was named the Delian League, after the sacred island of Delos in the Aegean. More than 300 *poleis* joined the Delian League (see Hansen and Nielsen 2004 for a list). Each member state was given the choice between supplying the alliance with ships and men and paying tribute; the vast majority chose to pay tribute. Whether the League evolved into an Athens' dominated empire, or whether it was that way essentially from the beginning, has been debated – Thucydides (1.99) writes, "the Athenian navy grew strong at the cities' expense, and when they revolted they always found themselves inadequately armed and inexperienced in war" (quoted in Morris 2005, 40); Parker (2009) and Raaflaub (2009) argue that the League was always an empire.

ships and sailors used by the League, it kept all the tribute. Xenophon (*Anabasis*, 7.1.27) estimates that tribute flows brought Athens 1000 talents annually (between 10 and 20 percent of the total wealth held in Athens).[47] When the Persian threat receded in the mid-fifth century, Athens was able to build up a substantial surplus; the League treasury contained 6000 talents when it was relocated from Delos to Athens in 454 BCE.[48]

As head of the empire, Athens grew into the region's preeminent commercial power. Morris (2005, 24) writes, "Athens became the Aegean's economic central place, and imposed its own weights, measures and coinage on the Greek world."[49] Most trade went through Athens' huge port at Piraeus, and all commercial disputes were heard in Athenian courts.[50] As a result, commerce

---

[47]Morris (2005, 2). Meiggs (1972, 256-7) provides three comparisons intended to put the tribute received by Athens in context. First, the right to collect a two percent import duty at Athens' principal harbor (Piraeus) sold in the year 399 BC for 30 talents and generated profit of 6, suggesting 1800 talents worth of goods entered the harbor that year. Second, the first special tax on wealth to aid the war effort (*eisphora*), imposed in 428 BC, generated 200 talents, suggesting the existence of 10,000 talents of private capital. [We will add description of esiphora.] Third, a 377 BCE census valued total privately held wealth at 6000 talents.

[48]A peacetime navy was maintained at an estimated annual cost of about 480 talents (Hale 2009, 127). By the 460s at the latest, member *poleis* were not permitted to exit the alliance, and Athens punished attempts to do so harshly, destroying navies, confiscating land, and requiring reparation payments. Rhodes (2007, 36) writes, "From Naxos in the League's early years to various states in the last phase of the Peloponnesian war, including states on the Asiatic mainland . . . we can construct a substantial list of revolts." At the outbreak of the Peloponnesian War, Sparta had substantial support among Greek states that disapproved of Athens' "tyranny" towards its "allies."

[49]While Athens had drawn wealth from commerce for many years, it had previously been but one trading state among many; commercial powers like Aegina and Corinth were more than Athens' match in the sixth century. (We will add citation on 5th century imports.)

[50]The Old Oligarch writes, "Some think that the Athenian demos makes a mistake in this too, that they compel the allies to sail for their cases to Athens. The justification for this lies in the advantages this brings to the demos in Athens. For, first of all, the court fees provide the jurors fees for the whole year . . . In addition there are these gains for the demos of Athens when the allies' cases are heard at Athens: first of all the one percent [tax on trade] at Piraeus which the state gets is increased; then anyone who has lodgings does better and anyone who has a carriage-pair or a slave to earn money for him; then the criers do better when the allies stay here." Quoted in Meiggs (1972, 265).

flourished in the region. Meiggs (1973, 267) states that

> Aegean trade needed a strong fleet to suppress piracy and the sea lanes were probably more secure during the period of Athenian empire than at any other time in the ancient world except perhaps during the first two centuries of the Roman empire.

Sharing the benefits and costs

Scholars dating back to Thucydides have attributed the design of several uniquely Athenian institutions to the nature of its empire.[51] These institutions include: 1) the distribution of tribute payments to a large cross-section of the population, 2) the linking of the payments to tasks necessary to the running of the empire, and 3) the most radical democracy in ancient Greece, with unchecked power placed in the hands of a popular assembly in which all citizens, regardless of wealth or property holdings, served.

From the perspective of our model, the task faced by Athenians was more complicated than that faced by Spartans. Because the key wealth-producing asset for the Spartans was captive land and labor, the task of dividing the assets among the citizenry was relatively straightforward. The Athenian empire's productive assets could not be divvied up in a similar manner. Thus, the Athenians needed to find a way to allocate benefits so that a broad segment of the citizenry (i.e., not just those with direct commercial interests) would have incentives to support policies that maximized the returns from maintaining the empire. One critical feature of the Athenian solution to this

[51]Prior to empire, Athens was a "moderate" democracy, similar to many other *poleis* in ancient Greece in that certain public offices could be held only by the well-to-do, and in the vesting responsibilities in certain less representative bodies (as Sparta did with the *Gerousia*). Galpin (1984, 101, 107) writes that "the 'radical' democracy of Athens during the fifth century B.C. required imperialism for both ideological fulfillment and the establishment of certain characteristic institutions. . . . Perhaps more relevant to the relationship between democratic values and imperialism was the Athenians' perception that the empire was necessary to bring the democracy to fulfillment. The progressive completion of democratic institutions was simultaneous with the development of an extensive system for the distribution of public funds." Morris (2005) argues that the Empire accelerated Athenian state formation: "Athens expanded its democratic system in part by drawing on the *phoros* [tribute] paid by the subject cities."

incentive problem was to distribute shares of the tribute payments to poorer citizens – that is, to citizens who otherwise would have been unlikely to favor incurring the large costs, especially the manpower required for a navy, necessary to maintain the empire. As a result of the payments, writes Morris (2005, 40), "Most of the tribute ended up in the pockets of the poorer Athenian citizens." In addition, tribute payments were used to fund public works, such as the building of the Parthenon.[52] And these public works employed many laborers (thus increasing the returns to the less-skilled group of citizen-laborers), in addition to generating benefits directly enjoyed by the public.

For understandng how our model applies, an essentual point is that the Athenian tribute-sharing sytem linked payments to activities that helped maintain the empire. As a result, when making decisions with respect to allocating additional resources for the purpose of bringing in more tribute, members of the Athenian assembly had incentives to consider opportunity costs, not just the additional tribute. Perhaps most importantly, the essential activity of the rowing Athenian ships, was the first to be compensated.[53] Athenian triremes were fast and highly maneuverable, making them

---

[52]Meiggs (1973, 258) writes: "It was common knowledge that the Parthenon, the Parthenos, and the Propylaea had been paid for in large part from the tribute reserve brought from Delos in 454." Rhodes (2007, 29-30) writes "between ca. 449, when regular campaigning against Persia came to an end, and 431, when the Peloponnesian war began, the income from the tribute must greatly have exceeded the sums spent for League purposes. . . . In the 440s and 430s Athens was spending large sums on buildings on the Acropolis and elsewhere, which were paid for indirectly and probably to a considerable extent directly out of surplus tribute from the League." Pericles answered critics of such expenditures as follows: "They [the allies] contributed neither horse nor ship nor hoplite [soldier] but money only, and the money is the business not of those who give it but of those to whom it is given, so long as they supply the services for which the money was given" (quoted in Meiggs 1972, 133).

[53]Athens made a fateful decision to build ships with the proceeds from an unanticipatedly massive silver strike at its mines at Laurium in 483 BCE. Herodotus says that it had been originally proposed to distribute to the silver citizens at 10 drachmas per head, but Themistocles convinced Assembly to vote to use to build triremes instead. See Meiggs (1972, 37). The result was that by 480 BC (in time for the famous naval battle at Salamis), Athens had 200 new triremes, as compared to probably less than 100 ships (including more primitive penteconters) two decades earliers. "No Greek state could match this fleet and it proved to be the backbone of the combined navy in 480 at Artemisium and Salamis" (Morris, 2005, 37).

lethal in battle – *if* manned by trained rowers.[54] Athens' rowers were drawn largely from the lower ranks of Athenian society, those without sufficient wealth to purchase the hoplite panoply necessary to serve in the infantry (which is where Sparta's citizens were found).[55] Approximately 17,000 Athenian citizens served as rowers – nearly half the citizen population.[56] Rowing triremes was not only skilled work, it was dangerous. Thus, if rowers had instead been conscripted without pay, that would have created the incentives for non-rowers to favor an excessive use of rowers and, at the same time, created incentives for rowers to oppose ventures that would be unprofitable for rowers even though profitable for Athens.

In addition to defense, the empire required administration: Tribute payments had to be received and tracked, tribute-based distributions (for pay or public building programs) had to be monitored), and trade-based litigation had to be resolved (as noted, Athens mandated that all commercial disputes among its allies be heard in Athens). Such tasks were both made necessary by and funded by the empire – Meiggs (1973, 258) writes that "Without some such annual income from outside [tribute] it would not have been possible to provide pay for state services, from the jurors of

---

The largest states in the Delian League – Chios, Lesbos, and Samos – contributed ships and men initially instead of tribute; however, the fleets were disbanded after failed revolts (by Samos in 439, by Lesbos in 427, and by Chios in 412), from which point onwards each contributed tribute in addition to paying reparations.

[54]Thucydides writes of rowing a trireme, "Seamanship, just like anything else, is an art. It is not something that can be picked up and studied in one's spare time. Indeed, it allows one no spare time for anything else" (quoted in Hale 2009, 77). A trireme had three levels of oars, and was crewed by about 180 men.

[55]Rhodes (2007, 30) writes, "Rowing the Athenian navy's ships was primarily the responsibility of the poorer citizens, the *thetes*, those too poor to be able to equip themselves to fight as hoplites." Previously, this subset of the citizenry had accompanied the hoplite infantry as "the naked and the light" (add citation).

[56]At its peak, Athens operated 400 triremes, many employing paid rowers from other states. (citation)

the popular law-courts to the archons, and at the same time maintain a large enough fleet to ensure command of the seas."[57] Athens was the only *polis* to provide such payments.[58]

Thus, the benefits as well as the costs of maintaining the Athenian empire were distributed widely among citizens. Of course, we could not reasonably argue that Athens achieved anything close to the alignment of costs and benefits described in our Implication 6. What we do argue, however, is that the degree to which all Athenian citizens shared the benefits and costs of the empire allowed Athens to rely on a highly unique form of government – a radical democracy (Aristotle's term) that concentrated power in a largely unchecked popular assembly, to which each citizen belonged and in which each had an equal vote.[59]

The move to radical democracy with an unchecked assembly

To understand the "radical democracy" of Athens, it is essential to recognize that the process of designing that democracy included Athenians voting to disempower a pre-existing counter-majoritarian check on assembly actions (the Areopagus) and to abolish all property requirements for

[57]See Kaiser (2007) on the taxation for triremes. Key point is that the wealthiest citizens contributed a lot of resources to the navy. (we need to verify timing of tax policies)

[58]The work the payments funded (see what follows in text) was unpaid in every other Greek *polis*, with one possible exception, and that for the fourth century: "De Ste. Croix, op. cit., 602, n. 24, admits that he can prove the existence of public pay in other states only during the fourth century, and then only for Rhodes" (Galpin 1984, 101).

[59]The Athenian citizenry consisted of all free adult males who were born to a legally married Athenian couple (i.e., an Athenian citizen married to the daughter of another Athenian citizen), and it was very unusual for a non-Athenian (specifically, not born to a citizen and the daughter of a citizen) to become a citizen. Moreover, Athens had many slaves who, along with Athenian women, had no voting rights. Thus, while the distribution of the Athenian franchise was radically broad for its day, it was far narrower than what we expect to find in modern democracies.

holding public office or for voting in the assembly.[60] In contrast to Sparta – and to every other Greek *polis* – Athens placed no limits on what the assembly could do.[61] Accordingly, Hansen (1999, 174) writes that by the mid-fifth century:

> the assembly had more and more frequently used its increased power to legislate, and the traditional sense of the priority of the laws had given way to a sense that the people in their assembly were the highest power in the state.

In the context of our model, removing a counter-majoritarian check makes sense if the costs of (potentially preventable) undesirable majority-supported policies are below the costs of the counter-majoritarian check (e.g., socially undesirable policies favored by, and implemented under the influence of, the Areopagus).

What about the empire might have brought about such circumstances? We conjecture that Athens focused policymaking in a broadly representative assembly – without a countermajoritian check – to maintain the broad citizenry's direct control over the use of tribute flows. The key factor is that members of the Areopagus, if empowered to constrain the assembly, would have had incentives to use their power to divert the benefits of tribute flows in a manner that served their own

[60]The Areopagus was a council on which all men who had ever served as archon – "chief magistrate" – sat for life. Its mandate had been, in Starr's (1990, 8-9) words, "Guardianship of the laws." In the 460s, this "island of aristocratic power" (Hansen 1999, 37) was reduced to a single narrow function, hearing cases involving the homicide of a Athenian citizen. The name means "Hill of Ares," on which the Areopagus met. The precise size of the Areopagus is not known, but an average membership of about 150 is thought to be typical for the Classical period. See, e.g., Hansen (1999, 290) and MacDowell (1978, 27).

[61]In Sparta, no bill could be brought to the assembly until it had first been discussed and approved by the *Gerousia,* a council on which only some members (those over the age of 60) of certain families were allowed to sit. The *Gerousia* not only set the agenda for the assembly, but was empowered to overrule "crooked" (i.e., not in the general interest) assembly decisions by simply declaring the assembly adjourned. See Forrest (1968, 46, 49-50) for a discussion the *Gerousia* and of the clause which granted this right. The *Gerousia* also served as a criminal trial court for charges of murder, treason, and other serious crimes.

self-interest.[62] Yet the median voter in the assembly was a recipient of tribute payments and thus had the incentive to watch tribute funds carefully. Indeed, when the tribute was received from each allied state, it was paraded in front of the assembly, and detailed records of "Athena's share" (from which total payments by *polis* could be calculated) were kept on a carved stone block for public view.

In short, all Athenians were residual claimants to the profits of the empire. Thus, the incentives for voting on public goods – most importantly the military effort necessary to maintain the empire's tribute flows – were kept relatively well aligned.

Institutional redesign in response to a shift in the source of wealth

At the end of the fifth century, when Athens' defeat in the Peloponnesian War brought the end of its empire, the flow of tribute ended, and this gave the Athenians an incentive to redesign their institutions. Very importantly, the potential gains from seagoing trade did not disappear – nor did necessity of a powerful navy to keep the seas safe enough to allow that trade. But the loss of its previously massive inflow of tribute moved Athens into a position much more similar to that faced by modern market-oriented democracies: To realize the potential gains from seagoing trade and other commerce, Athenians needed an institutional setting that encouraged wealth creation – notably wealth creation based on voluntary exchange and mutually beneficial contracts – and Athenians recognized this need.[63] The recent literature on good institutions in the modern world suggests that

[62]Athenians were acutely aware of the effects (positive and negative) of concentrating power. (Athens had been ruled by a tyrant in the sixth century, and there were several attempts in the fifth century to swing Athens from democracy and to oligarchy.) As Morris writes of fifth century Athens, "classical democracy was a kind of compromise between the need to have state institutions and the desire to prevent anyone from capturing control of them."

[63]The is evident in the following speech by a litigant involved in a trade-related dispute: "Do not ignore the fact that by resolving one dispute you are passing a law for the entire port of Athens. Many of the men who have chosen to engage in overseas trade are watching you to see how you will decide this case. If you think that written contracts and agreements between partners should be binding and you will not take

Athenians would have done well if they established a commitment to relatively stable policy, including secure property rights and the rule of law, and a tax system sufficient to fund expenditures on public goods (including the navy in the case of Athens) and other government activities.

In fact, Athenians did just that. Most importantly, Athenians redesigned their institutions through a series of reforms that – and this is the critical point for our analysis – constrained the popular assembly in the manner consistent with what the modern literature on the rule of law suggests – and with an effort to avoid time-inconsistency problems similar to what our model illustrates. The reforms consisted of four things. 1) writing down laws so as to render them distinct from decrees issued by the assembly, and recognizing that the former trump the latter, 2) establishing a process for writing new laws and amending old laws, 3) establishing a process for challenging both newly written laws and assembly decrees for conformity with existing laws, 4) re-empowering the Areopagus to serve as "guardians of the law." (We describe these reforms and their effects in more detail in this paper's appendix). In short, the loss of empire appears to have inspired the Athenians to set up a system more akin to a modern "rule of law" (indeed, some classicists have described the reforms using precisely this phrase)?[64]

Our explanation of these changes is based on the tradeoffs illustrated in our model: When the Athenians lost their ability to collect tribute, reforms that established counter-majoritarian checks on the legislature became desirable, because the benefits of those checks became larger than the costs. As our model illustrates, an unchecked assembly may lack the ability to make credible

---

the side of those who break them, those involved in lending will more readily make their assets available. As a result, the port will thrive, and you will benefit." Quoted in Harris (2006, 143).

[64]See Ostwald (1987) and Harris (2006) on the "rule of law" in Athens.

commitments to (ex ante) desirable future policies. It was to this problem that Aristotle referred when he wrote of radical democracy (in the style Athens' 5th century assembly-centered government): "The laws are forever being changed by decree [i.e., by the assembly] and are seldom respected."[65]

Majority rule with a meaningful counter-majoritarian check

Can we be confident that the reforms did in fact constrain the previously unchecked assembly? Yes, as a comparison of one famous fifth century case and two well-known fourth century cases demonstrates. Perhaps the most famous example of the fifth century Athenian Assembly's refusing to obey its own rules involves the "Trial of the Generals." In 406 BCE, eight generals commanding newly commissioned ships with inexperienced crews won an important and unexpected victory over the Spartans near the island of Lesbos. However, the Assembly became

[65]Quoted in Hansen (1999, 68). A famous example involves the erstwhile Athenian ally Mytilene, and is described by Thucydides. In 428, Mytilene attempted to leave the Athenian-dominated Delian League and ally itself with Athens' enemy, Sparta. Athens dispatched its navy and subdued the Mytilenes (who did not receive the help from the Spartans they had counted on). Furious with the breaking of the treaty, the Athenian Assembly voted the ultimate punishment: All Mytilenean men were to be executed, and all women and children to be sold into slavery. A warship was dispatched to give the Athenian commanders in Mytilene the order. However, on the following day some citizens began to regret the severity of the decision, and a second meeting of the Assembly was called. After vigorous debate, the Assembly voted to dispatch a second warship to countermand the first order. Rowing day and night (rowers supposedly fed at the oars), the ship arrived in Mytilene shortly after the initial warship, in time to preserve the native population. A less harsh punishment was applied – confiscation of all Mytilene ships and Mytilene wealth held on the mainland. (A decade later, Melos, who sought to avoid taking sides in renewed Athens-Sparta hostilities, was not so lucky – after defeating the Milesians in battle, the Athenians executed all its men, sold it women and children into slavery, and settled 500 Athenians on the then-empty island.)

Note that even without a third-party enforcer or counter-majoritarian check, a majoritarian institution can, at least in principle, overcome time inconistency problems if the conditions match a repeated game. And, in fact, there were instances when the Athenian assembly apparently wished to force itself to commit and succeeded in doing so. A device existed known as the "entrenchment clause," which designated assembly decisions un-amendable, even setting death as the penalty for any individual who proposed modifications. As far as is known, the Athenian Assembly respected these clauses; however, entrenchment clauses were rarely used, employed mostly for treaties with foreign powers. See Schwartzberg (2004) for a detailed analysis of entrenchment clauses. See also Rhodes and Lewis (1997).

angry when messengers reported that the generals had failed to retrieve survivors (and corpses) from the water, as was typically expected (bad weather was the reason given). The generals were ordered home (two of them fled; the other six returned to Athens), and the Assembly debated the proper venue for a trial. After a recess at which the deceased were commemorated, the Assembly reconvened in an angry mood. It was proposed that rather than try the generals (who had been imprisoned), the Assembly simply vote on whether to execute them as a group. When a relative of one of the generals protested that the vote would be unconstitutional (a right to individual trial was ordained by law), he, himself, was threatened with execution. The philosopher Socrates was serving as chairman of the Assembly on that day (citizens were chosen for this duty by lot); when he refused to bring the motion for execution to a vote, saying he would not engage in an act that was contrary to the law, he was shouted down with the phrase, "It is shocking not to let the people do whatever they wish" (Starr 1990, 47). In the end, a vote was taken and the execution of the imprisoned six was approved and carried out.[66]

Contrast that story with the following fourth century accounts. In 348, a member of the Athenian council named Apollodoros proposed that certain funds be assigned for a military use despite the fact the funds had been allocated for a different purpose. The justification that Apollodoros offered was (echoing the earlier trial) that "the *demos* [people] should be free to do whatever it wished with its own money." The assembly passed the motion, but a rival prosecuted Apollodoros under a *graphe paranomon* – the charge of introducing in the assembly a decree that is in violation of the law. The charge was upheld by the court and Apollodoros was sentenced to pay

[66]Hansen (1999, 41) writes, "The Trial of the Generals was cited by contemporaries as evidence that assembly democracy was a bad form of government."

a fine of one talent.[67]

Another fourth century example involves the famous statesman Demosthenes. In 323 BCE, a court found Demosthenes guilty of accepting a bribe and sentenced to pay a fine of 50 talents; Demosthenes, unable to pay, fled into exile. Facing troubles abroad, the assembly sought to recall him to seek his advice, but found itself unable to act against a court decision. Instead, the assembly voted to allocate fifty talents to Demosthenes, so that he might pay the fine.

In sum, after the Athenians lost their empire, they redesigned their institutions in a manner that constrained the scope of policies subject to simple majority rule decisions. This made sense, because with the end of tribute flows, there came (i) an increased importance of a commerce-promoting institutional environment and, hence, counter-majoritarian checks to ensure stable policy and the rule of law, and (ii) a decreased concern that an institution not broadly respresentative of the population (such as the Areopagus) would use its power to divert wealth.

Our model also suggests a conjecture as to why the Athenian assembly expressed so much anger during the Trial of the Generals. News of death in battle, even a battle that ended in victory, would naturally upset the family and friends of those killed, and that was probably an important influence on the assembly's behavior. Also consider, though, the critical role of naval power – and the pivotal role of rowers – in the Athenian system. Managing the Athenian empire well under a majority rule system required a commitment to keep the benefits and costs of the empire shared in a manner that aligned voters' incentives. Citizen rowers accepted the risk of death in exchange for what they were paid to row and for the effect of naval power on tribute flows. Yet their acceptance of that risk was conditional on naval officers following the agreed upon rules – rules that influenced

[67]For details on both cases, see, e.g., Sealey (2007, 253-4).

how the costs of maintaining the empire were allocated among the citizens. Thus, breaking those rules was not just a threat to rowers, but a threat to the success of the Athenian system.

V. LESSONS FOR MODERN DEMOCRACY

Although our historical focus in this paper is ancient Greece, we argue that our model and analysis of Athens and Sparta demonstrate tradeoffs that remain fundamental to democracies. Thus, as a further illustration, we will consider the pioneer of modern democracy: the United States. Given the pioneering nature of American democracy, along with the country's institutional durability and economic success, we should expect to see that the citizenry – and to be specific, the subset of the population who had voting rights – started with sufficiently aligned incentives over the generation of wealth to allow the American experiment in institutional design to succeed. And, of course, if the tradeoffs in our model matter, we should see first-best outcomes precluded by the difficulty of aligning voters' incentives over the policies affecting the value of key assets.

One key fact is that, at the time of its founding, the potential for westward expanion provided the American government with ability to allocate rights to newly acquired assets – land originally held by Native Americans, then bought, conquered, stolen, or otherwise taken. Being able to make land ownership opportunities available to large numbers of people, without taking land away from enfranchised landowners, made it relatively easy to obtain alignment among voters with respect to supporting a legal system that promised credibly to enforce property rights to land – the same could not be said, for example, about France after the French Revolution. Put another way, even though the landholdings were not distributed among the citizenry as equally as they had been in Sparta, the typical citizen had little to gain and much to lose from a decision to weaken the protection of

property rights.

Another key fact is that the United States relied heavily on a captive labor force – enslaved African Americans. In sharp contrast to the distribution of helots among the Spartans, the distribution of slaves among the American citizenry was very unequal, and an Aristotle-style solution (e.g., as in Implication 5) requiring an equal distribution of slaves among the free, enfranchised population would not have worked.[68] Moreover, some types of land (e.g., land suitable for plantations) were assets complementary to slave labor (e.g., plantation land), while other land (notably on the frontier) was not – and many free citizens opposed slavery because they saw slave labor as a substitute for free labor. Quite obviously, the United States never did manage to align voters' incentives over policies to enforce slaveholders' claims to slave labor. Instead, the lack of alignment led to civil war – the antithesis of alignment over public good policy, given that each side set its military out to destroy the other.

Our model also provides insight into the modern literature on "polarization" in American politics, along with the prominence of the "One Percent" in political and academic debates.[69] Although partisan and/or liberal-conservative divisions have always been the norm in the United States (e.g., Poole and Rosenthal 1997), in the last several decades the degree of polarization in congressional roll call voting patterns has increased dramatically (e.g., Bonica, McCarty, Poole, and Rosenthal 2013). The literature identifies many contributing factors, including gerrymandering, the rising level of inequality driven by increased returns to human capital, and campaign finance. Our

[68]The cost of reallocation (whether compensated or not) would have been enormous, and the reallocation of slave labor would have reduced output (i.e., T would have been huge).

[69]See, e.g., the symposium on "The Top 1 Percent" in *The Journal of Economic Perspectives*, Summer 2013.

model suggests a simple question: When looking a major policy issues, has there been a change over time in the degree to which the benefits and costs are shared in a manner that aligns voters' (and politicians') incentives?

Because we seek to explain a change over time, the natural starting point is to identify an earlier point that contrasts with the present. Perhaps most importantly, the New Deal, despite creating serious controversy, put into effect a set of policies that had a "something for everyone" design; moreover, the New Deal Coalition famously brought otherwise disparate interest groups together in support of a broad-based, logrolled set of policies (e.g., Brady 1988; Fleck 2008). One legacy of this is Social Security, a program that redistributes a vast amount of income, but does so by requiring most of the public to pay taxes into the system (when working) and distributing benefits (when retired) using a formula that, in part, links individual-level benefits to the individual-level history of taxes paid into the system. By avoiding a system in which those bearing the costs were distinct from those receiving the benefits, the Social Security system – by design – built broad support. As Franklin Roosevelt famously explained (in 1941): "We put those payroll contributions there so as to give the contributors a legal, moral, and political right to collect their pensions and their unemployment benefits. With those taxes in there, no damn politician can ever scrap my social security program."[70] This is, in essence, an Athenian-style mechanism for keeping voters' incentives relatively well aligned or, in modern parlance, avoiding polarization.[71]

---

[70]See http://www.archives.gov/exhibits/treasures_of_congress/text/page19_text.html.

[71]This is not to say that the Social Security Act was sufficient to maintain the broad coalition in support of the public pension system it established. Indeed, given that the early beneficiaries received unsustainably generous benefits, adjustments to taxes and/or benefits had to be made, and in principle these adjustments could have been made in a way that weakened the coalition supporting the program. However, the form of the first major adjustments (most notably those implemented during the Nixon Administration)

How does this differ from today? If voters expect that expansions of federal programs will be paid for by scaling up the current federal personal income tax rates (and if voters assume those rates approximinate the actual incidence of the tax burden), then voters will expect the costs of those expansions to fall entirely on a distinct (tax-paying) subset of the population. In this light, it is unsurprising to observe that a wide variety of programs – from the Affordable Care Act to various types of "stimulus" spending – generate very similar "polarizing" divisions in Congress. This type of political division is what the Athenians and Spartans sought to avoid. An Athenian style solution might be to fund stimulus programs with a consumption tax.

VI. CONCLUSION

Sparta and Athens famously differed on many dimensions, but in this paper we argue that their contrasting institutions can be interpreted as responses to the same basic problem: How can the designers of a majority rule system provide incentives to voters – or how can voters constrain themselves – so that policy decisions will reflect the broad interest of the citizenry? Addressing that question inevitably involves making tradeoffs. After the conquest of Messenia, the Spartans' top priority was to keep the citizenry's incentives aligned with respect to the military. Thus, the Spartans set up a unique system (one that enforced an exceptional degree of homogeneity among the citizenry) that kept those incentives highly aligned. The opportunity cost of this system was that Spartans had to forgo gains from specialization. In Athens, however, forgoing the gains from specialization would have been too costly, and the Athenians therefore designed complex institutions (relying on a variety

---

was to increase taxes along with benefits, thus further entrenching the durability-enchancing mechanism described by Roosevelt.

of taxes, payments, and other rules) that kept incentives relatively well aligned among a heterogeneous population. To govern its empire, Athenians designed a system of governance dominated by a broadly representative – and largely unchecked – assembly. When Athens lost its empire, the benefits of relying on (necessarily imperfect) counter-majoritarian institutions increased, and the Athenians responded by redesigning their institutions. Thus, Sparta and Athens both engineered successful (from their citizens' perspectives) institutions given the constraints they faced. Understanding how they did this provides new insight into the design of successful democracies.

Many of the themes we explore in the context of ancient Greece are similar to those emphasized in Acemoglu and Robinson's (2012) book, though we reach somewhat different conclusions. Similar to the role of well aligned versus poorly aligned incentives in our analysis, Acemoglu and Robinson find that the establishment of "inclusive" institutions supports incentives that lead in the long run to wealth and prosperity, while in the presence of "extractive" institutions, incentives to create wealth will be weaker. Yet Acemoglu and Robinson (2012) express skepticism about the prospect for redesigning institutions: "You cannot engineer prosperity" is the title of a chapter subheading. The reason is that time and chance played a large role in the emergence of "good" institutions, and extant beneficiaries of "bad" institutions tend to resist reform. An important conclusion of our paper is that prosperity *can* be engineered – and indeed was engineered in ancient Greece – but the success of efforts to engineer prosperity depends upon exogenous conditions, as well as on the incentives of the would-be engineers and the timing of the engineering.

BIBLIOGRAPHY

Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared. 2008. "Income and Democracy." *American Economic Review* 98:808-842.

Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared. 2009. "Reevaluating the Modernization Hypothesis." *Journal of Monetary Economics* 56:1043-1058.

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A. Robinson. 2014. "Democracy Does Cause Growth." NBER Working Paper No. 20004.

Acemoglu, Daron, and James A. Robinson. 2000. "Why Did the West Extend the Franchise? Democracy, Inequality and Growth in Historical Perspective." *Quarterly Journal of Economics* 115:1167-1199.

Acemoglu, Daron, and James A. Robinson. 2001. "A Theory of Political Transitions." *American Economic Review* 91:938-63.

Aghion, Alesina, Trebbi. 2004.

Anderson, J. K. 1970. *Military theory and practice in the age of Xenophon*. Berkeley: University of California Press.

Barro, Robert J. 1997. *Determinants of Economic Growth*. Cambridge: The MIT Press.

Boix, Carles. 2003. *Democracy and Redistribution*. Cambridge: Cambridge University Press.

Bonica, Adam, Nolan McCarty, Keith T. Poole, and Howard Rosenthal. 2013. "Why Hasn't Democracy Slowed Rising Inequality?" *Journal of Economic Perspectives* 27:103-124.

Brady, David W. 1988. *Critical Elections and Congressional Policy Making*. Stanford: Stanford University Press.

Buchanan, James M., and Gordon Tullock. 1962. *The Calculus of Consent: Logical Foundations of Constitutional Democracy*.

Cartledge, Paul A. 1979. *Sparta and Lakonia: A Regional History 1300-362 BC* London: Henley & Boston

Cartledge, Paul A. 1987. *Agesilaos and the Crisis of Sparta* London: Henley & Boston

Cartledge, Paul and Antony Spawforth. 2002. *Hellenistic and Roman Sparta*. New York: Routledge.

Carugati, Federica. 2015.

Carugati, Hadfield, & Weingast. Forthcoming.

Clarke, Edward H. (1971). "Multipart Pricing of Public Goods". *Public Choice* 11 (1): 17-3.

Defense Base Closure and Realignment Commission. 2005. "Report."

Falaschetti, Dino. 2008. "Can Lobbying Prevent Anticompetitive Outcomes? Evidence on Consumer Monopsony in Telecommunications." *Journal of Competition Law and Economics* 4:1065-1096.

Fernandez, Raquel and Dani Rodrik. 1991. "Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty." *American Economic Review* 81:1146-1155.

Finley, Moses I. 1973. *Democracy: Ancient and Modern* New Brunswick: Rutgers Universwity Press.

Fischel, William A. 2001. *The Homevoter Hypothesis: How Home Values Influence Local Government Taxation, School Finance, and Land-Use Policies*. Cambridge: Harvard University Press.

Fleck, Robert K. 2000. "When Should Market-Supporting Institutions Be Established?" *Journal of Law, Economics, & Organization* 16:129-154.

Fleck, Robert K. 2008. "Voter Influence and Big Policy Change: The Positive Political Economy of the New Deal." *Journal of Political Economy* 116:1-37.

Fleck, Robert K., and F. Andrew Hanssen. 2006. "The Origins of Democracy: A Model with Application to Ancient Greece." *Journal of Law and Economics* 49:115-146.

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.

Fleck, Robert K., and F. Andrew Hanssen. 2012. "On the Benefits and Costs of Legal Expertise: Adjudication in Ancient Athens." *Review of Law and Economics* 8:367-399.

Fleck, Robert K., and F. Andrew Hanssen. 2013a. "How Tyranny Paved the Way to Democracy: The Democratic Transition in Ancient Greece." *Journal of Law and Economics* 56:389-416.

Fleck, Robert K., and F. Andrew Hanssen. 2013b. "Judicial Review as a Constraint on Tyranny of the Majority." *Journal of Law, Economics, & Organization* 29:303-331.

Fleck, Robert K., and F. Andrew Hanssen. 2013c. "When Voice Fails: Potential Exit as a Constraint on Government Quality." *International Review of Law and Economics* 35:26-41.

Forrest, W.G. 1969. *A History of Sparta 950-192 B.C.*, Norton: New York

Freeman, Charles. 1999. *The Greek Achievement.* New York: Allen Lane Penguin Press.

Galpin, Timothy J. 1984. "The Democratic Roots of Athenian Imperialism in the Fifth Century B.C.", *The Classical Journal* 79: 100-109

Glaeser, Johnson, Shleifer. 2001.

Groves, Theodore. (1973). "Incentives in Teams". *Econometrica* 41 (4): 617-631

Hale, John R. 2009. *Lords of the Sea: The Epic Story of the Athenian Navy and the Birth of Democracy*, NY

Hansen, Mogens H. 1999. *The Athenian Democracy in the Age of Demosthenes.* Norman, OK: University of Oklahoma Press

Hansen, Mogens H., and Thomas H. Nielsen. 2004. *An Inventory of Archaic and Classical Poleis*. Oxford: Oxford University Press.

Hanson, Victor D. 1999. *The Other Greeks: The Family Farm and the Agrarian Roots of Western Civilization*. Berkeley: University of California Press

Hanssen, F. Andrew. 2004. "Is there a Politically Optimal Level of Judicial Independence?" *American Economic Review* 94:712-729.

Harris, Edward M. 2006. *Democracy and the Rule of Law in Classical Athens*. Cambridge: Cambridge University Press.

Hodkinson, Stephen. 2000. *Property and Wealth in Ancient Sparta*. London: Gerald Duckworth and Co

Jack, William, and Roger Lagunoff. 2006. "Dynamic Enfranchisement." *Journal of Public Economics* 90:551-572.

Jew, Daniel. 1999. "Food, silver, trade and liturgies: modeling the Athenian economy." Unpublished M.Phil. thesis, Cambridge University

Johnson, Ronald N., and Gary D. Libecap. 2003. "Transaction Costs and Coalition Stability Under Majority Rule." *Economic Inquiry* 41:193-207.

Justman, Moshe, and Mark Gradstein. 1999. "The Industrial Revolution, Political Transition, and the Subsequent Decline in Inequality in 19th-Century Britain." *Explorations in Economic History* 36:109-127.

Kaiser, Brooks. 2007. "The Athenian Trierarchy: Mechanism Design for the Private Provision of Public Goods." *Journal of Economic History* 67:445-480.

Lipset, Seymour Martin. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy," 53 *American Journal of Political Science* 69-105.

Lizzeri, Alessandro, and Nicola Persico. 2004. "Why Did the Elites Extend the Suffrage? Democracy and the Scope of Government, With an Application to Britain's 'Age of Reform'", *Quarterly Journal of Economics* 119: 707-65.

Llavador, Humberto, and Robert J. Oxoby. 2005. "Partisan Competition, Growth, and the Franchise." *Quarterly Journal of Economics* 120:1155-1189.

MacDowell, Douglas M. 1978. *The Law in Classical Athens*. Ithaca: Cornell University Press.

Manville, Philip Brook. 1990. *The Origins of Citizenship in Ancient Athens*, Princeton University Press: Princeton, NJ

Maskin & Tirole. 2004.

McCannon, Bryan C. 2010a. "Homicide Trials in Classical Athens." *International Review of Law and Economics* 30:46-51.

McCannon, Bryan C. 2010b. "The Median Juror and the Trial of Socrates." *European Journal of Political Economy* 26:533-540.

Meiggs, Russell. 1972 (reprinted 2002). *The Athenian Empire* Oxford: Clarendon Press

Morris, Ian. 2005. "The Athenian Empire (478-404 BC)"

Morris, Ian. 2010. "Why the West Rules – For Now." New York: Farrar, Straus, and Giroux.

Murray, Oswyn. 1993. *Early Greece*. Cambridge: Harvard University Press

North, Douglass C., and Barry R. Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth Century England," 49 *Journal of Economic History* 803-832.

Ober, Josiah. 2008. *Democracy and Knowledge*. Princeton: Princeton University Press.

Ober, Josiah. 2010. "Wealthy Hellas." *Transactions of the American Philological Association* 140:241-286.

Ober, Josiah, and Barry R. Weingast. 2013. "Is Development Uniquely Modern? Athens on the Doorstep." Working paper, Stanford University.

Olson, Mancur. 1982. *The Rise and Decline of Nations*. New Haven: Yale University Press.

Parker, Robert. 2009. "Introduction", in Ma, John and Nikolaos Papazarkadas (eds.) *Interpreting the Athenian Empire London*: Duckworth and Co. (1-12)

Pitsoulis, Athanassios. 2011. "The egalitarian battlefield: Reflections on the origins of majority rule in archaic Greece." European Journal of Political Economy 27:87-103.

Pomeroy SB, Burstein SM, Donlan W, Tolbert Roberts J. 1999. *Ancient Greece: A Political, Social, and Cultural History*. New York: Oxford University Press.

Pomeroy, Sarah B. 2002. *Spartan Women*. New York: Oxford University Press.

Poole, Keith T., and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford University Press.

Raaflaub, Kurt A. 2009. "Learning from the Enemy: Athenian and Persian 'Instruments of Empire'", in Ma, John and Nikolaos Papazarkadas (eds.) *Interpreting the Athenian Empire London*: Duckworth and Co. (89-124)

Rhodes, P.J. 2007. "Democracy and Empire", in Samons II, Lorren J. (ed.) *The Cambridge Companion to the Age of Pericles* Cambridge: Cambridge University Press (Chapter 1)

Rhodes, P.J., and David Lewis. 1997. *The Decrees of the Greek States*. Oxford: Clarendon Press.

Riker, William H. 1982. *Liberalism Against Populism*. San Francisco: W. H. Freeman.

Samons, II, Loren J. 2007. *The Cambridge Companion to the Age of Pericles*, Cambridge: Cambridge University Press

Schwartzberg, Melissa. 2004. "Athenian Democracy and Legal Change", *American Political Science Review* 98: 311-325

Sealey, Raphael. 1987. *The Athenian Republic: Democracy or the Rule of Law?* London: Pennsylvania State University Press.

Semple, Ellen Churchill. 1921. "Geographic Factors In the Ancient Mediterranean Grain Trade," 11 *Annals of the Association of American Geographers* 47-74.

Shaw 1944. Everybody's Political What's What

Starr, Chester C. 1990. *The Birth of Athenian Democracy: The Assembly in the Fifth Century B.C.*, New York: Oxford Press

Strauss, Barry S., and Josiah Ober. 1990. *The Anatomy of Error: Ancient Military Disasters and Their Lessons for Modern Strategists*, St. Martin's Press: New York.

Teegarden, David A. 2014. *Death to Tyrants! Ancient Greek Democracy and the Struggle against Tyranny*. Princeton: Princeton University Press.

Treisman, Daniel. 2012. “Income, Democracy, and the Cunning of Reason.” Working paper, University of California, Los Angeles.

Vickrey, William (1961). "Counterspeculation, Auctions, and Competitive Sealed Tenders". *The Journal of Finance* 16 (1): 8-37.

APPENDIX
The Athenian reforms

In 403 BC, Athens began a process of revision, ratification, and codification its laws. As noted above, the various rules by which Athens was governed emanated from (or were attributed to) the earlier law codes of Draco and (principally) Solon, and decrees by the Assembly.[72] Indeed, the proper distinction to be made (if any) between a longstanding law and a recent Assembly decree was murky at best, and adjudicators chose rules and used them as they saw fit. Most laws were written down somewhere (generally on stone and thus difficult to move), but not in one place nor in a particular order (as evinced by the fact that the codification launched in 403 took four years to complete). How conflicting rules should be dealt with was unclear.

The first task was thus to develop a formal hierarchy of laws that would be understood and respected by all. Through the laborious collection and inscription of the various rules, Athens established a single legal code. These written laws were termed *nomoi*, or "fundamental laws." They were intended to be permanent (although an amendment process was developed – see what follows) and to supercede and govern any subsequent acts by the Assembly. Assembly acts were termed *psephisma*, or "decrees." Decrees generally dealt with specific issues – funding a given infrastructure project, awarding citizenship to a particular resident alien, rewarding a particular general. But the "law" (*nomos*) was to govern what types of decrees could be issued, and in what manner the decrees could be applied. This distinction between laws and decrees had been articulated earlier in Athenian history (for example, in Pericles' famous funeral oration), but until the end of the fifth century, had rested on no formal base.[73] The new laws not only specified expectations regarding individual behavior, but the obligations and rules under which political bodies were expected to operate. The Athenian legal code thus became a de facto written

---

[72]Solon served as *archon* in the early sixth century BC, and was classical Athens' most noteworthy "law giver" (he was one of the famous "Seven Sages"). Parties seeking to endow preferred rules with authority would attribute them to Solon, making it difficult to determine which laws were actually his (relatively little of his original law code has survived). Even less of Draco's earlier code has survived, other than the harsh penalties for murder which exemplified its "Draconian" nature.

[73]For all practical purposes, the terms *nomoi* and *psephisma* were used indiscriminately before the reforms of 403 BC. Kleisthenes' laws had been called *nomoi*, presumably with the intention of distinguishing them from the earlier laws of Solon and Draco, which were called *thesmos* (something "laid down"). *Psephisma* refers to decision taken by means of *psephoi*, or pebble, which apparently were used for voting in the Assembly early in the fifth century – see Hansen (1999, 161). However, the later clear distinction between *nomoi* (laws) and *psephisma* (decrees) is evident in this passage from the orator Andokides (400/399 BC), "No decree [*psephisma*] passed by the Assembly or the people may have higher validity than a law [*nomos*]. No law [*nomos*] may be passed that applies only to a single person. The same law [*nomos*] shall apply to all Athenians, unless otherwise decided [in a meeting of the Assembly] with a quorum of 6000, by secret ballot." Quoted in Hansen (1999, 170).

constitution.[74]

Of course, for a constitution to govern, it must be enforced. In Athens, this meant three things. First, the Assembly had to be unable to alter the fundamental law easily or capriciously. Second, a process was required to reconcile acts of the Assembly with fundamental law. Finally, a body was required to ensure that those who applied the fundamental law did so properly. The Athenians thus did the following. First, they established a new body charged specifically with scrutinizing the laws and approving changes in the law – only under well-defined procedures could the law be altered. Second, they established a new public action whereby any citizen could challenge "unsuitable laws" in the people's law courts. They also adapted an existing procedure and focused it on prosecuting the proposers of "unconstitutional" decrees. Third, they empowered an existing institution to oversee the application of the law.

The new institution created by the Athenians to consider changes to the fundamental law was called the *nomothetai*; literally, the "layers down of the law."[75] The members of the *nonomothetai* were drawn by lots from the list of citizens available to serve on *dikasteria*, rendering the *nomothetai* effectively a specialized people's court. As a result, it enjoyed the same democratic legitimacy as the *dikasteria*, which ensured that the *nonomthetai* had sufficient popular support to play the necessary role of gatekeeper for proposed changes to the fundamental law.

Once per year, at the first meeting, the entire law code was read aloud to the Assembly (Hansen 1999, 139). The membership of the Assembly would then vote on whether to revise any of the laws. Most often, no changes were proposed; however, any member of the Assembly (i.e., any citizen, because all citizens belonged) had the right to propose and argue for a particular revision. If this happened the proposed new law, and any existing law that would be altered or invalidated as a result, were posted in the marketplace and filed with the clerk of the Assembly, who read them at the two succeeding meetings of the Assembly (under normal circumstances, the Assembly met (? – add citation) times per year). At that point, if the proposer wished to continue, a *nomothetai* of ten men was established.[76] At the same time, five members of the

---

[74]It did not define in a *de novo* sense the bodies of government, and thus was not a true constitution as we think of the term. The collection and ratification (or rejection) of the new laws were entrusted to two separate bodies; membership of the first chosen by the Council of 500, and of the second by the *demes*. The Assembly played no direct role in the process, other than setting it in motion.

[75]During the initial process of revision of the law code, two *nomothetai* were created; one charged with collecting the laws, and the second (as above) with ratifying. The former was established by the Council of 500 and dissolved once the collection process had been completed. See Hansen (1999, 163). The latter was reconstituted every time a change in the law code was proposed – see what follows. See Bonner (1969, 98-9), Hansen (1999, 165-78).

[76]Add citation. Had to be 30. Usually ? members (add citation).

Assembly were designated to defend the old law. It was the duty of the *nomothetai* to listen to the proposer of the new law, to the five public advocates defending the existing law, and to any other citizen who cared to speak. The *nomothetai* then voted among themselves, and the issue was decided by simple majority (a tie preserved the old law intact). The Assembly had no power to overrule the *nomothetai*'s vote. If the *nomothetai* rejected a proposed change, the revision was set aside; if it accepted the change, the law code was altered.

However, the process did not end there. Even after the formal revision to the legal code, individual citizens retained the right to challenge revised laws, either on the basis of the character of the new law, or on the basis of having failed to observe proper procedure in proposing and enacting the new law. If this occurred, the case was heard in court. However, the challenge took place under a newly constituted legal procedure: *graphe nomon me epitedeion theinai*, or "public action for having proposed and carried out an unsuitable law." The citizen who proposed the original revision to the law code in the Assembly was required to defend his revision, while the citizen who brought the challenge played role of prosecutor, explaining in what way the fundamental law was thus violated.[77] If the *dikasts* agreed that the new law was indeed "unsuitable," the original proposer paid a fine; at he extreme, he could lose the right to bring any future revision to the laws.[78]

With the establishment of the *graphe nomon me epitedeion theinai*, the previously existing public action directed at laws and Assembly decrees (when the two were less clearly distinguished) – *graphe paranomon* – evolved into a means of challenging the conformance of Assembly decrees with the fundamental law – i.e., the "constitutionality" of the Assembly's actions.[79] The proposer of the challenged decree was required to defend himself in court (the challenger acting as prosecutor); if the decree was deemed to violate the fundamental law, the proposer faced a fine, loss of legal rights, or (in extreme cases) exile.[80]

Thus, the reforms of 403 produced in Athens a new concept of the law; specifically, a *formal* distinction between fundamental law (*nomoi*) and acts of the Assembly (*psephisma*); a new review body; and a new legal procedure to adjudicate challenges to the fundamental law. The Athenians distinguished clearly between fundamental laws and decrees from that time

---

[77]Professional advocates were not used in ancient Athens in legal cases of any kind – the party claiming injury brought a case and prosecuted it himself, while the accused party acted as his own defense attorney. A majority vote of *dikasts* decided the issue. (add citation)

[78]After a year had passed, the proposer of the law was free from personal liability, but the law could still be challenged at any time. (add citation)

[79]It is not clear whether *graphe paranomon* was introduced under Cleisthenes or Ephialates – see the discussion in Starr (1990, 26-8).

[80]See Starr (1990, 27); add citation to Hansen

forward: If a decision involved a fundamental law, the public declaration began with phrase, "It was decided by the *nomothetai*," while if the decision involved a decree, the public declaration began with the phrase, "It was decided by the Assembly", or "It was decided by the Council and the people" (Hansen 1999, 167).

The final step involved the establishment an institution to oversee application of the law. And the final of the major institutional changes involved re-empowering the Areopagus to serve as a formal and long-lived overseer of the law. This re-empowerment proceeded in two major steps. The first occurred in 403, at the same time the law code was being formalized. In addition to its powers to hear homicide cases, the Areopagus was empowered to once again supervise the administration of laws by the *archons*. Such supervision was an important part of the formalization of the laws. Under the pre-reform regime, *archons* relied on customary (unwritten) law as much as on formal (written) law – as noted, the distinction between the two was far from sharp. Thus, when the law code was formalized, all controlling law was written down, and subsequently, the application of informal (unwritten) law was prohibited. This was important in reducing not only the discretion of the Assembly, but the discretion of magistrates – Hansen writes, "It is important to notice that the prohibition of unwritten law was directed against *magistrates* [emphasis in original]." The Areopagus, being staffed by live-serving ex-chef magistrates, brought a level of ability (and an incentive to develop expertise) lacking in a typical court, whose membership changed with each case.[81]

The second reform occurred about 50 years later. In the mid-fourth century, a new procedure known as *apophasis* was established to prosecute offenses against the state (such as treason and attempts to overthrow democracy). Responsibility for bringing the charge rested not with private citizens or individual magistrates (as was the case for most other offenses), but with the Assembly or Areopagus directly. Regardless of which body made the initial charge, the Areopagus was responsible for conducting a preliminary investigation (including summoning of witnesses, taking testimony, and so forth). The word "*apophasis*" refers to the report the Areopagus made to the Assembly following its investigation. The report included a provisional recommendation of "guilty" or "not guilty". After the report had been read and discussed, if the Areopagus' recommendation was for acquittal the case was abandoned; if it was for condemnation, the Assembly would confirm (or not) by a show of hands. If it were confirmed, the case passed to the *dikasteria* for final judgement (Hansen 1999, 134).[82]

---

[81]There were no judges under the Athenian legal regime, so that each case was decided by judge-like juries several hundred strong, who were not allowed to cross examine or consult among themselves before voting. See Fleck and Hanssen (2012) for more detail and an explanation for why a system eschewing formal legal expertise would have been established.

[82]The Areopagus was also given joint responsibility with the Council of 500 (and several other boards) for supervising all religious sanctuaries in Athens and Attica (Hansen 1999, 291).

Finally, in the latter half of the fourth century, the Areopagus became a venue for judging all citizens charged with a broad range of political offenses (not just treason). In effect, the Areopagus was transformed into a specialized court of last resort, available to prosecutors who preferred the *dikasteria* with their constantly changing *dikasts.*[83]

[83]This last power is believed to have been pushed through by Demosthenes during the original panic over invasion from Macedon, but was still in force more than a decade later. See Hansen (1999, 292)

Figure 1: Timeline of Changes