Origins of Early Democracy

Ali Ahmed
New York University
kta230@nyu.edu

David Stasavage
New York University
ds166@nyu.edu

September 2017

Abstract

The idea that rulers must seek consent before making policy is a key component of democracy. Using data from the Standard Cross Cultural Sample, we suggest that this practice evolved independently in a substantial fraction of human societies where executives ruled jointly with councils. We argue that council governance was more likely to emerge when information asymmetries made it harder for rulers to extract revenue from populations, and we illustrate this mechanism with a theoretical model. In such situations, giving the population a role in governance became one means of overcoming these information problems. We test this hypothesis by examining the correlation between exogenous localized variation in potential agricultural suitability and the presence of council governance, obtaining robust results. As a further step, we suggest that executives facing substantial information asymmetries could also have an alternative route for resource extraction - develop a bureaucracy that could measure local variation in productivity and then directly assess taxes. We provide further empirical results to suggest that rule through a bureaucracy could substitute for shared rule with a council.

1 Introduction

The Greeks may have invented the word democracy, but we will argue in this paper that the fundamental democratic practice whereby rulers must seek consent from a council arose independently in a broad number of human societies.[1] Standard accounts by political scientists generally overlook this critical fact. Even as astute an observer as Robert Dahl (1998) suggested that while some form of democracy may have been commonplace when humans were hunter gatherers, as people settled into agricultural communities, monarchy and despotism became the rule. We will suggest that Dahl was only right about a portion of human societies. There were many cases in settled societies where rulers were constrained by councils. We do not claim that council governance also implied the sort of very broad political participation that exists in many countries today; in most cases it certainly did not. Even so, the core democratic practice of obtaining consent was widespread. We suggest further that council governance was more likely to emerge in areas where rulers were at an informational disadvantage with respect to those they sought to govern. Under these conditions sharing power with a council that could provide information allowed both rulers and those they ruled to achieve better outcomes.

A number of recent authors have explored early state development and both its short-term and long term consequences.[2] This work has generally limited itself to asking whether a centralized state developed or not.[3] We explore not only whether a state existed but also what type of state existed and in particular whether it was more autocratic or instead had collective governance via a council.[4]

In an important contribution, Mayshar, Moav, and Neeman (2017) have recently argued that in early societies where production was more transparent, a state was more likely to take on a coercive character. They show this with a principal agent model and a

[1] See Hansen (1986) on the debate over when the term *demokratia* was first used.

[2] See Michalopoulos and Papaioannou (2014, 2013), Mayshar, Moav, Neeman and Pascali (2017), Hariri (2012), Angeles and Elizalde (2017), Gennaioli and Rainer (2007). Osafo-Kwaako and Robinson (2013), Baker, Bulte, and Weisdorf (2010), Bockstette, Chanda, and Putterman (2002).

[3] A notable recent exception is Boix (2015)

[4] Among archaeologists Blanton and Fargher (2016, 2008) have recently advocated a similar approach

case study of Ancient Egypt, a society where the high predictability of agricultural production arguably favored the development of a more coercive state. Starting from this same basic insight regarding transparency of production, we consider how actual institutions of governance might evolve to deal with information asymmetries. In the presence of information asymmetries, rulers were more likely to share power with a council whose members could provide information about local conditions.

To illustrate how council governance could emerge, we build a very simple theoretical model that motivates our empirical analysis and in particular the use of the variance in localized agricultural suitability as our core explanatory variable. We consider a state defined following North (1981) as an entity that extracts revenue in exchange for protection. In the model a ruler attempts to extract revenue from a representative citizen whose income is stochastic, but the citizen can first send a costless message about his or her income. In this setting it is possible to have an informative equilibrium in which a citizen communicates truthful information about income prior to the ruler's choice of a tax rate. The informative equilibrium can exist as long as two conditions hold. First, there must be some commonality of interest between ruler and citizen over what revenues are used for. Second, the ruler's choice of the tax policy must be constrained by the threat that a citizen will revolt or simply exit if the tax rate is set too high. This informative equilibrium, which we interpret as having a council exist, is pareto superior to an equilibrium without communication. Also, the net utility benefit from the informative equilibrium is increasing in the variance of citizen income. For this reason we suggest that rulers are more likely to share power with councils in societies with high variance in agricultural suitability from location to location.

For our empirical analysis we make use of data from the Standard Cross Cultural Sample on the presence of council governance in a set of 186 societies covering all world regions. We discuss this evidence in detail in Section 2 while also drawing on material from other sources. It is clear that council governance has existed in a very broad set of human societies. We then make use of a measure of local agricultural suitability produced

by Galor and Ozak (2016, 2015) which measures the maximum number of calories that can be drawn from a 5' x 5' cell (roughly 10km x 10km). We will refer to this measure as *caloric potential*. Importantly, the Galor and Ozak measure is designed to capture exogenous features of the natural environment rather than factors that might depend on human intervention, such as if yields are improved by irrigation or adding inputs to the soil. For each of the 186 SCCS societies we then construct a measure of *caloric variability* which is the standard deviation of *caloric potential* across space. This is a useful empirical measure to the extent that a higher standard deviation in agricultural suitability from location to location implies greater information asymmetries for rulers, which seems very plausible.

Using the SCCS and Galor and Ozak data we provide evidence of a positive and statistically significant correlation between council governance and caloric variability. This correlation is robust to the inclusion of a number of different pre-treatment covariates. This correlation is not driven by societies in only one or two world regions; it appears to be quite general. Further support for the theoretical interpretation we give to our result comes from the fact that the council and caloric variability correlation is stronger in societies where there is some evidence of formal taxation and the correlation is absent in a sub-sample of societies that did not practice agriculture. As part of our empirical analysis, we also show that the council and caloric variability correlation is not simply capturing something about the presence of a state in general. In recent work, a number of scholars have used an SCCS measure of jurisdictional hierarchy to capture early presence of a centralized state.[5] We show that while the presence of a centralized state is correlated with high caloric potential, it is not correlated with high caloric variability. Likewise, council governance is correlated with high caloric variability, but it is not correlated with caloric potential.

While we focus on geographic/environmental factors in our empirical analysis, we make

[5]See Michalopoulos and Papaioannou (2014, 2013), Mayshar, Moav, Neeman and Pascali (2017), Angeles and Elizalde (2017), Gennaioli and Rainer (2007). Osafo-Kwaako and Robinson (2013), Baker, Bulte, and Weisdorf (2010)

no claim that geography is or was destiny for council governance. First of all, environmental constraints change over time, most notably in response to changes in technology. Second, geography is just one of the many potential sources of information asymmetries between rulers and ruled. Our more general point is that council governance is more likely to prevail when rulers are at an informational disadvantage with respect to those they seek to rule. This informational disadvantage could come from a wide variety of sources, geographic or other.

As a further step in our analysis, we also provide evidence to suggest that rulers could find other means of dealing with information asymmetries other than relying on a council. In particular, they could build a bureaucracy. Using SCCS data on bureaucrats, we show that in the absence of full-time bureaucrats there is a strong correlation between caloric variability and council presence. When a full-time bureaucracy is present this correlation is essentially zero. This evidence is only suggestive because the presence or absence of bureaucrats is a post-treatment condition. However, it does suggest that in the presence of information asymmetries rulers could pursue either a democratic or a non-democratic route to governance.

The remainder of this paper is organized as follows. We first present our core empirical measures of council governance and variation in council presence across regions. We then provide a simple theoretical model to illustrate why caloric variability might make it more likely that rulers would share power with a council. This is followed by a presentation of our measure of caloric variability. We then present our core empirical estimates, followed by a consideration of whether we are simply capturing an effect whereby caloric variability promotes state formation in general. Following this we then explore two separate sets of heterogeneous effects, showing that councils at the local and central levels could act as substitutes for each other and that in the presence of a trained bureaucracy there was no correlation between caloric variability and council presence. We then briefly explore whether there is evidence of a correlation between democracy today and prior council governance, and we offer a brief conclusion.

2 Council Governance Was Widespread

That council governance has been widespread in human societies is not a new idea. In 1877 Lewis Henry Morgan, an early American anthropologist suggested that "The council was the great feature of ancient society, Asiatic, European, and American.."[6] Even if Morgan sometimes stretched this idea too far, most notably by falsely concluding that Aztec rulers governed jointly with a council, in this section we will show that council governance has been indeed widespread historically.[7] To do this we will make use of data from the Standard Cross Cultural Sample (SCCS), first proposed by Murdock and White (1969). This covers a set of 186 societies, each thought to be representative and the best described society in a specific geographic area. The intent of this design was to allow for statistical analysis with greater confidence that individual observations are independent from each other since only one society is chosen from a geographic area. The SCCS societies were drawn from a broader set of 1265 societies in Murdock's (1967) *Ethnographic Atlas* for which a more limited range of variables have been coded. Subsequent coding of SCCS societies has resulted in a very large set of measures, some of which pertain directly to political organization. Tuden and Marshall (1972) coded a number of political characteristics of those SCCS societies that had some sort of central political organization about the community level. This was the case in 85 societies. Murdock and Wilson (1972) coded features of political organization at the community level in all SCCS societies that had some form of political organization that extended beyond individual families. This was the case with 173 societies. In what follows we will make use of evidence on councils at both the community level and the central state level.

As part of their effort, Tuden and Marshall (1972) coded whether supreme decision making authority at the central state level was "concentrated in a single authoritative leader" or whether "authority is shared more or less equally by a single (or plural) executive and a deliberative body," or finally whether "authority is vested, in a council, assembly,

[6]Morgan (1877) p.84

[7]See Gibson (1947) on Morgan's erroneous conclusions about the Aztecs.

or other deliberative body with no single executive." Based on this coding, we generated a variable called *central council* that takes a value of 1 if either of the latter two conditions prevails and zero otherwise. Central council governance was present in 29 of the 85 societies that had some form of political organization above the community level. According to the data collected by Tuden and Marshall (1972) membership in a central council almost invariably belonged to individuals with some elevated status, either heads of major class or ethnic components, or sometimes the equivalent of aristocrats. In a minority of cases council members were elected.

In considering political organization at the community level Murdock and Wilson (1972) also coded whether an individual leader shared power with a council. They distinguished between cases where the community has a single leader or headman (or plural leader or headman), as opposed to those instances where a ruler shares power with a "formal council or assembly" or where the community lacks a single leader and is governed collectively by a council. Using this information, we have a coded a variable *local council* that takes a value of 1 if either of the latter two conditions prevailed and zero otherwise. Local councils played a prominent governance role in 87 of the 159 communities where some local authority existed and there was sufficient information for political characteristics to be coded. Though Murdock and Wilson (1972) did not report information on who served in local councils, based on the individual sources that they used, council membership patterns probably paralleled those observed by Tuden and Marshall (1972) at the central level with heads of lineage groups often being represented.

The conclusion from the SCCS data that council governance was widespread in human societies fits with what we know from other sources. In Mesoamerica before the Spanish conquest Aztec rulers governed without councils through a bureaucracy, but their neighbors in Tlaxcala had what many have called a republican form of government with executive power vested in a council of several hundred individuals.[8] In precolonial Africa the Hausa kingdoms had an autocratic form of rule, but rulers in the Asante empire to

[8]See Carrasco (1971) and Fargher, Espinoza, and Blanton (2011)

the west and the Kuba kingdom to the south shared power with councils that placed substantial checks on their rule.[9] Likewise, in ancient Mesopotamia it is recognized that autocratic states, such as Babylon under Hammurabi, coexisted with other forms of rule where councils at the city level played a much more important role in governance.[10] Similar examples can be provided for variations in governance patterns in North America, Oceania, and Ancient India.[11]

The SCCS data suggest that when council governance existed it took three main forms. First, in the case of societies without central authority, a council could exist at the community level. Second, in societies with central authority, a council could exist at that level. In each of these two cases it is straightforward to see how sharing rule with a council might help rulers to reduce asymmetries of information when it came to tax collection if council members had better information about production. However, there was also a third pattern of council governance that could have the same effect. In this case there might be a central executive who relied on councils at the local level for information. This was a pattern evident in societies as diverse as early Anglo-Saxon England and Ancient Mesopotamia. In England, prior to the establishment of active central assemblies under Æthelstan (r.924-939) rulers relied principally on assemblies at the level of the Hundred and the Shire.[12] In Mesopotamia the rulers of Mari relied similarly on the collective responsibility of local councils for raising taxes (Fleming, 2004).

Given that any of the three above patterns of council governance could be used by rulers to deal with information asymmetries in tax collection, we also created a variable *any council* that takes a value of one if a council shared power at either the central or the local level and zero otherwise. This will be the principal variable that we will use for our main analysis. In section 7 we will provide further support for this choice with

[9] See Wilks (1975) for Asante and Vansina (1978) for the Kuba state.

[10] Daniel Fleming's 2004 book on the kingdom of Mari provides thorough evidence of this latter phenomenon based on cuneiform texts.

[11] For North America see Trigger (1976) for the Huron as one example among many. See Sharma (1968) for India and Panoff (1984) for Western Samoa.

[12] Loyn (1984) Maddicott (2010)

Table 1: Descriptive Statistics by Region

Any Council	No Council (%)	Council Present (%)	Total
Subsaharan Africa	29.03	70.97	100.00
	(9)	(22)	(31)
Middle Old World	25.00	75.00	100.00
	(7)	(21)	(28)
Southeast Asia/Insular Pacific	30.43	69.57	100.00
	(7)	(16)	(23)
Sahul (Australia et al.)	50.00	50.00	100.00
	(6)	(6)	(12)
North Eurasia/Circumpolar	57.14	42.86	100.00
	(8)	(6)	(14)
Northwest Coast of NA	50.00	50.00	100.00
	(3)	(3)	(6)
North and west of NA	50.00	50.00	100.00
	(6)	(6)	(12)
Eastern Americas	56.00	44.00	100.00
	(14)	(11)	(25)
Mesoamerica/Andes	60.00	40.00	100.00
	(6)	(4)	(10)
Far south America	33.33	66.67	100.00
	(1)	(2)	(3)
Any Council Total	40.85	59.15	100.00
	(67)	(97)	(164)

Note: Fraction of *any council* overall and by region. Frequencies shown in parentheses.

statistical evidence by showing that local and central councils could substitute for each other in reducing information asymmetries for rulers.

Table 1 shows the overall distribution of our *any council* variable for the overall SCCS sample as well as for specific regions. We can see that council governance was not specific to one or two regions. It was a widespread phenomenon. The associated map showing SCCS societies with and without council governance helps to reinforce this point.

There are several further questions we should ask before accepting the idea that governance by council evolved independently in a broad number of human societies. The first of these involves the date of observation for the individual societies in the Standard Cross Cultural Sample. The goal of those who constructed the sample was to find the earliest available ethnographic accounts for each society. In some cases this was quite early, 1520 for the Aztecs and 1634 for the Huron. In many other cases, though, this "pinpoint" date was quite late, even well into the twentieth century. The median pinpoint date in the sample is 1915.[13] A supplemental reason for coding some societies at a later date was to obtain greater variation in cultural practices. This opens the possibility, for example, that SCCS societies coded at a later date might be more likely to have a council because of learning from others (and in particular Europeans) or because a council arrangement was imposed upon them.

A few simple tests suggest that this potential outside influence is not biasing our conclusions. While the fraction of societies with council governance (*any council = 1*) is indeed higher when the pinpoint date is after the median date of 1915 (65 percent), it is still quite high in those societies with pinpoint dates before the median (53 percent). It is also the case that SCCS societies that had already adopted words into their language from other (and in particular European) cultures were not more likely to have councils.[14] We can also consider the possibility of measurement error of the following form - European ethnographers with incomplete knowledge of local conditions may have been more likely

[13]This is also the case with the commonly used Ethnographic Atlas produced by Murdock (1967) where the median pinpoint date is 1920.

[14]This is based on a regression of *any council* on v1832 in the SCCS.

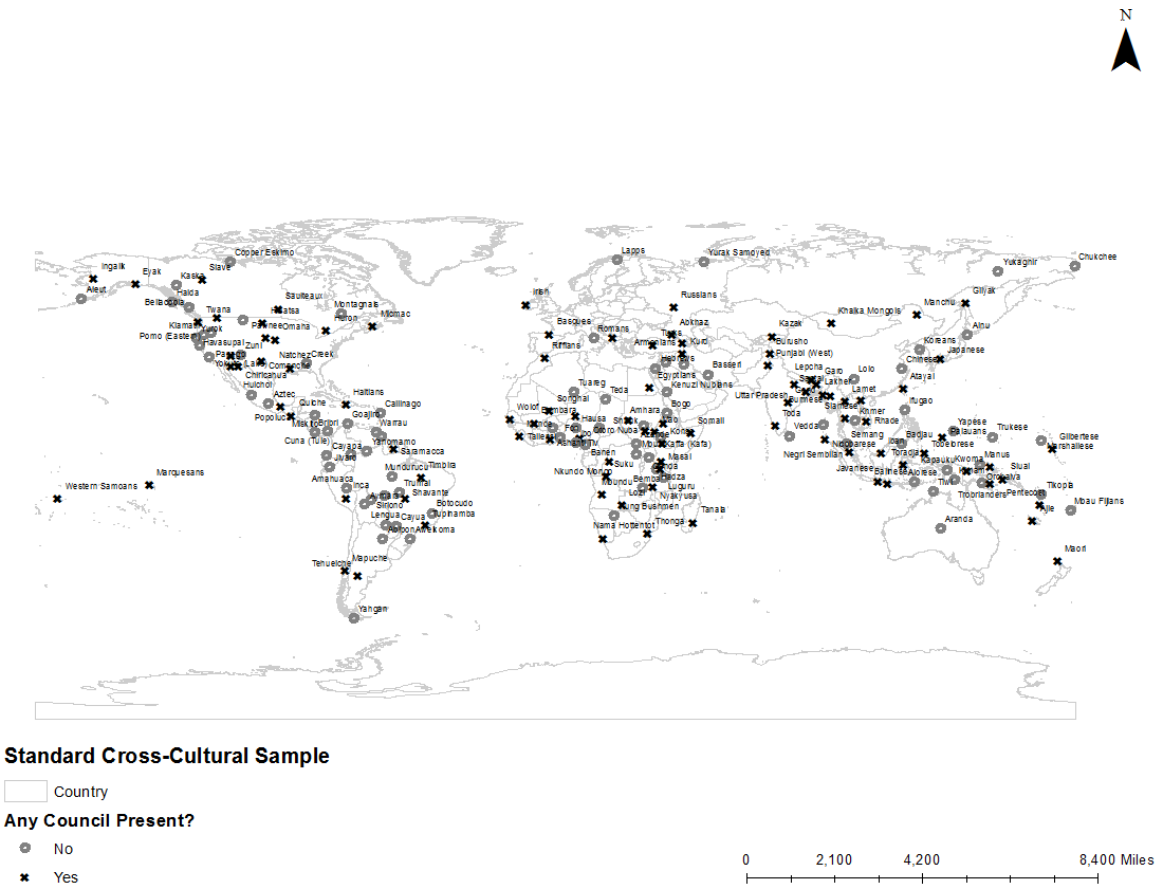

Figure 1: This map depicts the location of the 186 societies in the Standard Cross Cultural Sample distinguishing between those where council governance was present and where it was absent. Council governance is taken to be present if executives at either the central or local (i.e. community) level rule jointly with a council. See text for description of coding.

to see that councils existed when they did not. This would be based on the idea that they would infer back from their home experience. In fact, European ethnographers who did not know the language of the society in question were not more likely to report that a council was present.[15]

A further question one might want to ask about council governance is how it related in SCCS societies to the way in which executives were selected. Two recent papers by Giuliano annd Nunn (2013) and Bentzen, Hariri, and Robinson (2016) have used informa-

[15]This is based on a bivariate regression of our *any council* variable on variable v724 ("Knowledge of native language **(e.g., by ethnographers")) from the SCCS.

tion from the related *Ethnographic Atlas* (Murdock, 1967) to examine whether selection of community leaders by election in SCCS societies is correlated with measures of democracy in the same locations today. They find that this is indeed the case. The SCCS dataset also includes codes for how leaders (either central or local) are selected. Selection of central leadership was coded by Tuden and Marshall (1972). Selection of local leadership was reported by Murdock and Wilson (1972) based on earlier coding for the larger set of *Ethnographic Atlas* societies. One thing that is immediately clear from this data is that if council governance was widespread, formal election of leaders was much less so. At the central level leaders were formally elected in 26 percent of cases but in only 5 percent of cases prior to the median pinpoint date of 1915. At the local level leaders were formally elected in only 12 percent of cases and in only 5 percent of cases prior to the median pinpoint date.[16] One reason for the dearth of formal elections in the sample is that it is clear from ethnographic accounts that those who were ruled often had some influence on choosing or retaining a ruler, but this took place in in an informal manner. It is inevitably difficult to code whether this type of influence exists but Murdock and Wilson (1972) suggested that it existed at the local level in a further 20 percent of cases. It is also the case that leaders could be significantly constrained in their actions by the presence of a council even if they had not been formally elected. This fits with the fact that for centuries in the European context rulers in some states were constrained by councils even when they were not subject to elections.

3 How Income Variance Could Drive Council Presence

In this section we provide a simple theoretical model in which asymmetries of information over income prompt both rulers and those they rule to participate in a council form of governance. Existing literature on councils in the European context has mostly assumed

[16]These results for local leaders fit the pattern seen in the broader sample of the *Ethnographic Atlas* where community leaders were formally elected in only 8 percent of cases and in only 5 percent of cases if the sample is restricted to cases prior to the median pinpoint date of 1920.

that a council has veto power that prevents *ex post* opportunism by rulers.[17] How this veto power is sustained remains open to question. The historian James Tracy (1994) has argued that European assemblies rarely enjoyed formal veto powers over taxation, as they instead served as fora for discussion.

It turns out that we do not need to assume veto power to understand how a council can emerge and make a difference for governance. We simply need to show that rulers and those they rule have incentives to exchange information. We do this in a context of a state as defined by North (1981), an entity that extracts revenue in exchange for protection. The fact that both ruler and ruled in this context can benefit from revenue creates an (imperfect) commonality of interest that makes council governance possible. In what follows we will show this in a simple game setting where a ruler attempts to extract revenue from a representative citizen in a context where citizen income is stochastic. We will show that both ruler and citizen can often achieve higher utility in a game where the citizen can communicate about his income prior to tax policy being chosen, compared to a game with no communication. The net utility benefit from communication is increasing in the variance of the citizen's income. If we think of the communication game as governance by council, then we have a theoretical result motivating our empirical analysis - council governance is more likely to be observed the higher the exogenous variance in citizen income.

3.1 Revenue Raising Without Communication

Consider the following one period environment where a ruler seeks to raise revenue g from a citizen. The reference to "revenues" here should be thought of in broad terms, ranging from a society where the ruler collects taxes in monetary form, to one where taxes in monetary form are absent, but the ruler still has some role in redistributing resources such as foodstuffs. The citizen receives exogenous income y that has a deterministic

[17] This would be true of most work since the foundational contribution by North and Weingast (1989), though see Levi (1988) for a greater emphasis on bargaining than on veto rights.

component $y^d = 1$ and a stochastic component y^s drawn from a Bernoulli distribution with $Pr(y^s = 1) = p$. The ruler and citizen know y^d and p. Only the citizen knows y^s.

The ruler has utility $U^r = g^{\frac{1}{2}}$. The citizen has a subsistence requirement of 1, with a utility loss f if this is not satisfied, and then derives utility $U^c = g^{\frac{1}{2}} + (y - g)^{\frac{1}{2}}$ from income above subsistence.[18] The ruler's preferred tax rate τ is 1. The citizen's preferred τ is $\frac{1}{4}$ when $y^s = 1$ and 0 otherwise.

The order of play is as follows: (1) y^s is realized, (2) the ruler chooses τ, (3) the citizen chooses whether to revolt. A revolt succeeds with exogenous probability q. If a revolt is successful the citizen chooses τ, and the ruler suffers an additional cost c. After a failed revolt τ is set at 1.[19]

If $y^s = 0$, the citizen will revolt for all $\tau > 0$ due to the subsistence requirement. If $y^s = 1$, the citizen will revolt if the following inequality is satisfied (for cases where $\tau < \frac{1}{2}$).[20]

$$2q(\frac{1}{2})^{\frac{1}{2}} - (1 - q)f > \tau^{\frac{1}{2}} + (1 - \tau)^{\frac{1}{2}} \tag{1}$$

Given the revolt constraint, the ruler's expected utility is maximized by one of three policies. First, $\tau = 1$ will be chosen if q is sufficiently low that maximum extraction is optimal. Second, $\tau = 0$ will be chosen if q is sufficiently high that zero extraction is optimal. Finally, for intermediate values of q the ruler will choose a tax rate τ^* that just satisfies the revolt constraint. Excluding cases where $\tau = 1$ is optimal, the ruler will choose τ^* over $\tau = 0$ when the following inequality is satisfied.

$$p(2\tau^*)^{\frac{1}{2}} + (1 - p)[(1 - q)2^{\frac{1}{2}} + q((\frac{1}{2})^{\frac{1}{2}} - c)] > 0 \tag{2}$$

[18] This formulation is in the spirit of a Stone-Geary utility function where an individual has minimum needs.

[19] As an alternative we could incorporate an exit constraint into the model. In many instances, the ability of populations to move elsewhere had a significant impact on state formation and the type of state that formed. This idea was explored most notably by Carneiro (1970) for formation of a state in general, but he did not consider what type of state institutions formed.

[20] For $\tau > \frac{1}{2}$ the high income citizen will always revolt.

In the game without communication, unless $\tau = 1$ is the equilibrium, there are instances where *ex post* the ruler and citizen would each prefer a different τ. Take the case where $\tau = 0$ but $y^s = 1$. Here both ruler and citizen would prefer a higher tax rate. Take the case where τ^* is chosen. Here both ruler and citizen could prefer $\tau = 0$ when $y^s = 0$. This raises the question whether communication about citizen income could be sustained as part of an equilibrium.

3.2 Revenue Raising With Communication

Now consider a variant of the game in which the citizen can send a costless message $m \in \{0, 1\}$ about y^s. The timing is as follows: (1) y^s is revealed to the citizen (2) The citizen sends a message (3) the ruler chooses a tax rate (4) the citizen decides whether to revolt.

Take a putative equilibrium with a truthful message strategy $m = y^s$ where the ruler chooses τ conditional on m. The key question is whether a citizen with $y^s = 1$ will send a false message.[21] A citizen with $y^s = 1$ would receive utility of 1 from sending a message $m = 0$ because that would prompt the ruler to choose $\tau = 0$. If the high income citizen instead sends a truthful message $m = 1$, then the ruler will choose a tax rate τ^* that just satisfies the revolt constraint, as in (1) above.[22] We know the expected utility for a citizen when $\tau = \tau^*$ from (1). Given this, sending a truthful message is an equilibrium strategy for a citizen with $y^s = 1$ as long as the following inequality holds.

$$2q(\frac{1}{2})^{\frac{1}{2}} - (1 - q)f > 1 \tag{3}$$

.

The above inequality holds for sufficiently low values of f and sufficiently high values

[21] A citizen with $y^s = 0$ would not deviate, because they could not do better with a higher tax rate. Likewise if the message is $m = 0$ the ruler will choose $\tau = 0$ as long as qc is sufficiently high, and if the message is $m = 1$ the ruler will choose τ^* as long as q is sufficiently low.

[22] An equilibrium with informative communication cannot exist in the case where the ruler would opt for maximal extraction $\tau = 1$ if he knew $y^s = 1$. In that case the left hand side of the inequality in (3) would be $-f$.

of q. This is intuitive. As the citizen's expected utility from revolt increases, a ruler will choose a lower equilibrium tax rate, creating less incentive for a high income citizen to deviate and send a message $m = 0$.

3.3 Comparative Statics

The net expected utility benefit that the ruler and citizen each receive from the game with communication compared to the game without communication will depend on the difference in equilibrium tax rates when $y^s = 1$ and $y^s = 0$, which is a function of the exogenous parameters q, c, and f. The net expected utility benefit will also be increasing in the variance of y^s, which is equal to $p(1-p)$. As p approaches $\frac{1}{2}$, if the equilibrium tax rate in the non-communication game is $\tau = 0$ then there is a greater likelihood that both ruler and citizen would have preferred a higher tax rate, and if the equilibrium tax rate is τ^*, then there is a greater likelihood that both ruler and citizen would have preferred $\tau = 0$.

If the net utility benefit from an informative equilibrium in the game with communication is increasing in the variance of citizen income, then this equilibrium may be more likely to exist, precisely because it pareto dominates an equilibrium without communication. An alternative way to generate the same prediction is to imagine that transmission of the message m from citizen to ruler involved a direct cost incurred at the outset of the game, such as that for travel.[23] If at the outset of the communication game either the ruler or citizen, or both ruler and citizen in combination, had to commit to bearing a fixed cost of sending a message, and they were able to commit to this before y^s is revealed, then the communication variant of the game would be more likely to occur in areas with higher income variance (as $p \to \frac{1}{2}$). Under such conditions it would be more likely to be

[23]The literature on European representative assemblies is abundant with examples where under premodern conditions, representatives found it costly and difficult to attend assemblies. See Blockmans (1998) for this argument and Stasavage (2010) for empirical evidence. Evidence from other regions points to precisely this same phenomenon. Wilks (1975) emphasized how travel to central assemblies in the Asante empire became a burden for this reason. Trigger (2003) suggests that in premodern societies in general larger polities required more hierarchical decision making structures for this reason.

worthwhile to pay the fixed cost necessary to send a message. If we consider that the communication game involves joint rule with a council whereas the game without communication represents rule without a council, then we should be more likely to observe council governance the higher is the level of income variance. That is precisely the hypothesis we will test in the following sections.

4 Measuring Caloric Variability

We argued above that any exogenous factor that makes it more difficult for rulers to know whether citizens will bear a given rate of taxation makes it more likely that both ruler and citizens will prefer a council form of governance. One such factor involves transparency of agricultural production, something that has been emphasized by Mayshar, Moav, and Neeman (2017) who explore when and why more coercive labor arrangements prevail.[24] One important factor affecting transparency of production is the extent to which agricultural yields vary from location to location. If whoever is setting tax rates has incomplete knowledge of this variation, then they may face an incentive to consult with members of a council who do have such knowledge. In the case of a state that unites multiple communities, this could occur either by having representatives attend a central council or by having a ruler, or his or her agents, consult with local councils. In the case of a community that is not part of a larger state, the same dynamic could take place.

To measure geographic variation in agricultural potential we will make use of a measure produced by Galor and Ozak (2016, 2015). For each cell (with a cell defined as 5' x 5' or roughly 10km x 10km) on the Earth, Galor and Ozak produce a measure of the maximum number of calories that can be extracted per year if an optimal mix of crops is grown. The underlying data for Galor and Ozak's measure derives from UN Food and Agriculture Organisation data on agricultural potential and USDA data on caloric values for different crops. Galor and Ozak produced estimates of caloric potential based on both those crops

[24] In doing so they draw on the work of Allen (1997) who himself draws on the "circumscription" theory in Carneiro (1970).

available prior to the Columbian exchange (circa 1500) and those after the Columbian exchange. We use the post-1500 measure since the pinpoint date for almost all of the SCCS societies is well after 1500. In the rest of this paper we will refer to the Galor and Ozak maximum calorie measure as *caloric potential*. The Standard Cross Cultural Sample provides a geographic point location for each of the 186 societies in the dataset. Using this point we were able to obtain a measure of *caloric potential* for each of the SCCS societies.

Using the Galor and Ozak measure of *caloric potential* we then constructed a measure of *caloric variability* for each of the SCCS societies. This measure is simply the standard deviation of *caloric potential* in a sample of nine 5' x 5' cells including the cell where the SCCS society is located and the eight surrounding cells. This fits well with the core comparative static from our theoretical model where the probability of observing council governance depends on the variance in productive potential. For our empirical analysis we then took the average of this *caloric variability* measure within a 20 kilometer buffer. The idea behind this choice is that having only a single pinpoint location, we lack information on the full geographic area covered by each SCCS society. However, we would also like to avoid focusing on too small an area that may be atypical for the society in question. As we say this, we should note that our core empirical results are robust to alternative buffer dimensions ranging from zero kilometers to two hundred kilometers.[25]

One feature of the Galor Ozak *caloric potential* measure is that it is designed to be exogenous to human intervention. The underlying FAO data provide agricultural yields with four different levels of inputs. Galor and Ozak used the lowest level of inputs. The FAO data also provide yields assuming different water sources. Galor and Ozak assumed rain fed agriculture. This helps rule out the possibility that council governance might have an effect on the *caloric potential* and *caloric variability* measure.

[25] We should also note that Huning and Wahl (2016) have recently used a similar measure to explore state formation in medieval Germany, albeit with a different theoretical focus. Rather than attempting to explain the pattern of rule (with or without councils), they focus on the extent of political fragmentation. Also, their caloric variability measure differs from ours in that rather than focusing on the standard deviation of caloric variability, their measure is based on the square root of the sum of all squared deviations (not deviations from the mean).

Table 2: Descriptive Statistics – Caloric Variability

Any Council	No Council	Council Present	Total
Caloric Variability	260.3	332.6	303.6
	(450.8)	(521.5)	(494.1)
Central Council	No Council	Council Present	Total
Caloric Variability	414.1	147.7	318.7
	(580.2)	(152.5)	(489.2)
Local Council	No Council	Council Present	Total
Caloric Variability	245.3	346.3	301.1
	(413.9)	(547.5)	(493.3)

Note: Mean of *caloric variability* for a 20 km buffer around Standard Cross Cultural Sample (SCCS) societies. Summary statistics are shown by *any council*, *local council*, *central council*. Standard deviations in parentheses.

Table 2 provides summary measures of our *caloric variability* measure, providing evidence for our *any council* and *local council* variables of higher caloric variability in societies where council governance prevailed. We do not see the same pattern for the *central council* variable. All of these differences of course take no account of other factors that might simultaneously influence caloric variability and council presence.

5 Empirical Strategy and Main Results

The fact that we are using a cross-sectional dataset poses obvious challenges for causal identification. We cannot use change over time to help us identify whether caloric variability precedes council formation. With that said, the measure of caloric potential that we used to measure caloric variability is designed to be relatively immune to human intervention. Given this, barring the possibility that an entire society might choose to locate itself in a high caloric variability area because it was governed by a council, we should be immune to reverse causality. The likely greater risk is that because we cannot use change over time for identification, any statistical result regarding councils and caloric variability might be biased by the presence of fixed, unobserved factors at the societal level.

As a first step in considering this, in Table 3 we report balance tests to examine whether caloric variability is correlated with a range of different exogenous geographic covariates. As in all subsequent analyses we include *caloric variability* in log form with 1.0 added to prevent exclusion of societies where the *caloric variability* is zero. There are indications in Table 3 of *caloric variability* being correlated with several characteristics associated with the natural environment. This is not surprising given that features such as rainfall, temperature, or altitude certainly could be expected to influence caloric variability. Since we might be worried that one or more these factors could have an effect on council presence for other reasons, in one of our regression specifications we will include all covariates from Table 3 where there was not balance between high and low caloric variability societies.

In our core regression specification shown below where A is our *any council* variable

Table 3: Balance Table – ln(Caloric Variability)

	Higher Caloric Variability [1]	Lower Caloric Variability [2]	Difference [3]	Std. Err. [4]	Obs. [5]
Mean yearly annual rainfall	138.2410	139.0016	-0.7606	(3.8752)	182
Coefficient of variation in mean annual rainfall	31.9077	34.4269	-2.5191	(1.1922)	182
Lowest yearly rainfall in the n years sampled	95.4865	96.2101	-0.7235	(2.8800)	182
Highest yearly rainfall in the n years sampled	186.8530	186.5952	0.2578	(4.9875)	182
Difference between maxrain and minrain rainfall	91.3664	90.3851	0.9813	(2.2922)	182
Mean Annual Temperature	19.8087	20.0403	-0.2316	(0.4370)	176
Hottest North Mean Temperature	24.6073	24.6921	-0.0848	(0.3266)	176
Coldest Month Mean Temperature (xc)	11.3616	11.0331	0.3285	(0.6876)	176
Niche Temperature	2.2388	2.2873	-0.0485	(0.0981)	182
Land Slope	6.9894	7.1081	-0.1187	(0.0608)	182
Altitude in Meters	180.7961	98.1261	82.6700	(27.1291)	178

Note: Each cell reports a separate regression of the outcome of interest on the natural log of the caloric variability variable (i.e., $y = \alpha + \beta(CaloricVariability) + \varepsilon$). Column [1] represents the mean of the outcome of interest for societies with high caloric variability, $\alpha + \beta$. Column [2] represents the constant, α, interpreted as mean of the outcome for societies with low caloric variability (i.e. the 'comparison' group). Finally, column [3] shows the regression coefficient, β, interpreted as the 'difference' in the outcome of interest between high and low caloric variability across societies in the sample. Robust Standard errors are reported in parentheses in column [4], while column [5] gives the total observations.

for SCCS society i in region r and C is our *caloric variability* measure we will control for region fixed effects as well as latitude, longitude, and their product to help rule out the possibility that unobserved factors at the regional level might be biasing our results.

$$A_{ir} = \alpha + \beta ln(C)_{ir} + \gamma Lat_{ir} + \delta Lon_{ir} + \zeta LatX Lon_{ir} + \theta_r + \epsilon_{ir} \tag{4}$$

The results of linear probability estimates of equation (1) are shown in Table (4). In each of the four cases the coefficient on the measure of *caloric variability* is positive and statistically significant at least at the 1% level. If we interpret this result causally it would imply that a one standard deviation increase in the log caloric variability measure (4.3) would be associated with an increase of between 0.11 and 0.12 in the probability of having a council. This is a sizable effect. One further way to get a sense of the magnitude of this effect is to compare it to the probability of having a council in some comparison group of societies. If we had a binary treatment, identifying the control group would be straightforward. Since we have a continuous treatment variable, we defined a control mean as the mean of *any council* using the mean of *caloric variability* as the threshold to define the cut-off b/w high and low variability societies.

A next step in the analysis is to examine whether the positive correlation between councils and caloric variability is driven by one particular world region. Using a set of pre-defined regions from the SCCS (the same as those shown in Table 1) we ran an interactive model where the the coefficient on *caloric variability* was allowed to vary by region. There were significant differences in these interaction term coefficients between regions (a test of the null that they were jointly equal was rejected). This is not at all surprising. Our theoretical model, for one, predicts that caloric variability should only matter in instances where the revolt constraint is sufficiently tight, and this is something that may vary across regions. There is no indication in this data that the positive correlation between *any council* and *caloric variability* was specific to one region instead of being a broader

Table 4: Caloric Variability and Council Presence

	Any Council (*=1 Council, 0 No Council*)				
	[1]	[2]	[3]	[4]	[5]
Caloric Variability	0.055	0.060	0.057	0.063	0.066
	(0.019)	(0.020)	(0.019)	(0.021)	(0.023)
Region Fixed Effects	No	Yes	No	Yes	Yes
Latitude × Longitude Controls	No	No	Yes	Yes	Yes
Additional Controls	No	No	No	No	Yes
Adj. R-squared	0.0461	0.0704	0.0487	0.102	0.101
Observations	160	160	160	160	158
Control Mean	0.581	0.581	0.581	0.581	0.586

Note: Each cell reports a separate OLS regression of *any council* on *caloric variability*. Robust standard errors are shown in parentheses. The specification in column (5) includes as controls all variables from Table 3 for which there was not balance between the high and low caloric variability subsamples.

phenomenon.[26]

While the estimates in Table 4 provide strong evidence of a correlation between council governance and caloric variability, they also provide quite strong evidence that caloric variability was not all that mattered. In the first column in Table (4) that includes no controls, the r-squared is only 0.05. This low figure could be attributable to measurement error, but it could also most certainly be attributable to the fact that there are many other potential sources information asymmetries between rulers and those they govern. We see the Table 4 results as providing support for this general proposition rather than as evidence that geography is destiny.

Our interpretation of the correlation between councils and caloric variability is that high variability poses challenges for rulers attempting to extract revenue. We should think of revenue in broad terms here. Whether rulers are attempting to extract and/or redistribute taxation in monetary form, taxation paid in labor, or taxation paid in terms of agricultural production, the same problem involving information asymmetries could arise. This could exist even in non-agricultural societies. Among the Kwakiutl of the Pacific Northwest, Franz Boas (1913) reported that chiefs regularly obliged those who hunted or fished to turn over up to one half of their catch.

There are few variables in the Standard Cross Cultural Sample that refer directly to taxation. However, among those that do there is a record of some explicit form of taxation (and not the tribute referred to by Boas) in two thirds of societies. When repeating the Table 4 estimates while restricting the sample only to societies where there is evidence of explicit forms of taxation, we actually see that the coefficient on *caloric variability* is larger in magnitude and remains statistically significant, as one would expect if our theoretical interpretation was correct.[27] We should caution, however, that the existence of taxes is

[26]The coefficients and standard errors on the interaction terms in this model were Middle Old World 0.084 (0.056) Southeast Asia/Insular Pacific 0.176 (0.075) Sahul 0.040 (0.099) North Eurasia/Circumpolar 0.126 (0.496) Northwest Coast North America 0.0306 (0.073) North and West of North America -0.028 (0.111) Eastern Americas 0.255 (0.075) Mesoamerica/Andes 0.203 (0.056) Far South America 0.232 (0.034) Africa -0.029 (0.043).

[27]In this regression the coefficient on *caloric variability* was 0.096 with a standard error of 0.034. Societies with some form of taxation were selected based on variables v784 and v1736 in the SCCS dataset.

endogenous, and so we prefer to refer mainly to our Table 4 results that will not suffer from post-treatment bias.

Though forms of taxation, tribute, and redistribution can occur in both agricultural and non-agricultural societies, our *caloric variability* variable should ultimately be a better proxy for information asymmetries in agricultural societies. When restricting our sample to SCCS societies practicing agriculture we do indeed find that the coefficient on *caloric variability* remains statistically significant whereas this is not the case in a subsample of non-agricultural societies. However, the presence or absence of agriculture is clearly post-treatment, so we should be cautious in interpreting these results.

6 Caloric Potential and Political Integration

In this section we will expand the analysis to consider whether the correlation we have identified between councils and caloric variability is simply reflecting a more general correlation between state formation and aspects of the physical environment, and not the effect of information asymmetries that we have in mind. Ideally, we would like to be able to control for other aspects of state formation in our regression specifications and in particular for the extent of political integration. The problem is that a measure of political integration would be post-treatment. We will instead rely on exogenous determinants of political integration. Our contribution will be to show that while the level of political integration is correlated with total *caloric potential*, it is not correlated with *caloric variability*.[28] For council governance the exact opposite pattern holds. Our *any council* variable is correlated with *caloric variability* but uncorrelated with *caloric potential*.

To measure the degree of political integration we will make use of a measure of *jurisdictional hierarchy* from the SCCS dataset that has also been used in a number of other recent studies.[29] The measure of jurisdictional hierarchy we use, which is reported in

[28]This correlation between political integration and agricultural potential has previously been observed by Boix (2015). Mayshar, Moav, Neeman, and Pascali (2017) emphasize the importance for suitability of cereal agriculture in particular as does James Scott (2017).

[29]Michalopoulos and Papaioannou (2013, 2014) Mayshar, Moav, Neeman, and Pascali (2017)

Murdock and Provost (1971), is a five point that scale that takes a value of 1 if there is no political integration above the level of individual families, 2 if there is authority only at the community level, 3 if there is integration to one level above the community, 4 is there is integration to two levels, and 5 if there is integration to three or more levels above the community. Not surprisingly, there is a strong pairwise correlation (of 0.31) between our *any council* variable and *jurisdictional hierarchy*. However, we cannot simply add *jurisdictional hierarchy* to our existing specifications without inducing a risk of post-treatment bias.[30] To deal with this we will instead exploit the fact that *caloric potential* is an exogenous determinant of *jurisdictional hierarchy*.

In Table 5 we first report a set of reduced form relationships, alternately regressing either *any council* or *jurisdictional hierarchy* on both *caloric potential* and *caloric variability*. Including *caloric potential* will capture any direct effect it has on council development while also controlling for any indirect effects it might have on council development via the development of *jurisdictional hierarchy*. As we suggested above, the results here are quite consistent across specifications, and they suggest that while council presence is driven by *caloric variability*, political integration is driven by *caloric potential*.

In Table 6 as a next step we adopt an instrumental variables strategy where we use *caloric potential* to instrument for *jurisdictional hierarchy*. We caution that the exclusion restriction for this strategy may nor may not be satisfied. It is possible that *caloric potential* has a direct effect on council presence that we have not captured. It is also possible that *caloric potential* is correlated with other features that determine council presence. Those who doubt the validity of the exclusion restriction here should focus primarily on our reduced form results in Table 5. To the extent the exclusion restriction is satisfied, these estimates provide further support for the idea that the observed correlation between council governance and caloric variability is not biased by the failure to control for other aspects of state formation.

[30]For the record, when adding *jurisdictional hierarchy* to the first specification in table four we observe that the coefficient on *caloric variability* 0.065) with a standard error of .025. The coefficient on *jurisdictional hierarchy* is positive and statistically significant.

Table 5: Councils, Jurisdictional Hierarchy and Caloric Conditions (OLS Estimates)

	[1] Any Council	[2] Any Council	[3] Any Council	[4] Any Council	[5] Jurisdictional Hierarchy	[6] Jurisdictional Hierarchy	[7] Jurisdictional Hierarchy	[8] Jurisdictional Hierarchy
Caloric Variability	0.055	0.068	0.055	0.061	0.033	0.018	0.046	0.016
	(0.021)	(0.025)	(0.022)	(0.026)	(0.045)	(0.048)	(0.047)	(0.050)
Caloric Potential	-0.004	-0.045	-0.003	-0.020	0.226	0.321	0.262	0.356
	(0.049)	(0.060)	(0.055)	(0.064)	(0.110)	(0.130)	(0.117)	(0.134)
Region Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Latitude × Longitude Controls	No	No	Yes	Yes	No	No	Yes	Yes
Observations	157	157	157	157	179	179	179	179
Control Mean	0.583	0.583	0.583	0.583	3	3	3	3

Note: Each cell reports a separate regression with either *any council* or *jurisdictional hierarchy* as the dependent variable. Both *caloric variability* and *caloric potential* are entered in log terms. *jurisdictional hierarchy* is a five point scale, as described in the text. Robust Standard errors are reported in parentheses.

Table 6: Councils, Jurisdictional Hierarchy and Caloric Conditions (2SLS)

	[1] Any Council (IV)	[2] Any Council (IV)	[3] Any Council (IV)	[4] Any Council (IV)
Caloric Variability	0.055	0.056	0.072	0.062
	(0.026)	(0.028)	(0.028)	(0.026)
Jurisdictional Hierarchy	-0.015	-0.008	-0.133	-0.054
	(0.178)	(0.177)	(0.178)	(0.165)
First-stage IV				
Caloric Potential	0.278	0.304	0.339	0.374
	0.106	0.113	0.115	0.121
Region Fixed Effects	No	No	Yes	Yes
Latitude × Longitude Controls	No	Yes	No	Yes
F-stat first-stage IV	6.832	7.177	8.716	9.527
Adj. R-squared	0.0171	0.0231	.	0.0595
Observations	157	157	157	157
Control Mean	0.583	0.583	0.583	0.583

Note: Each cell reports a separate 2SLS regression where *any council* is the dependent variable. Robust Standard errors are reported in parentheses. *caloric potential* is the excluded instrument for *jurisdictional hierarchy*.

Taken together, the above results suggest that as has been suggested before, areas with a high level of agricultural suitability were more likely to develop states with hierarchies. But this seems to be a separate story from what determines whether governance occurs with or without a council. That appears to depend more on caloric variability, and this is consistent with a story where information asymmetries regarding local conditions create incentives for both rulers and ruled to participate in a council form of governance.

7 Central and Local Councils as Substitutes

So far we have shown a robust correlation between caloric variability and council governance where council governance takes place either at the level of the community or at a higher level of political integration. We have not explored the question why council governance sometimes occurs primarily at one level and not the other. Exploring this question in full is a task that lies beyond the scope of this paper. Previous work in a European context has shown that council governance is less likely to occur in larger scale polities, but unfortunately we lack appropriate scale measures for the individual societies in the Standard Cross Cultural Sample.[31] We will limit ourselves in this section to showing that local and central councils may have served as alternative means of dealing with information asymmetries in local production. We previously provided individual examples to suggest that this could be the case. In our empirical findings we will show that when a local council is absent, there is a positive correlation between central council and caloric variability. When a central council is absent, there is a positive correlation between local councils and caloric variability. When a local (central) council is present, there is not a positive correlation between central council (local council) and caloric variability.

In the regression specifications reported in Table 7 we investigate the correlation between separate types of council (local or central) and caloric variability while conditioning on whether the other type of council is also present. Since the presence of the other type of

[31] See Stasavage (2011, 2010) for empirical evidence and Blockmans (1998) for the original argument.

council is, of course, an endogenous development, we should add a caveat about potential post-treatment bias before considering these. In these interactive models we see evidence consistent with the interpretation that the effect of caloric variability on council presence at a given level of political integration is contingent on whether a council is already present at another level of political integration.

8 Bureaucrats as An Alternative to Councils

So far we have suggested that governance by council provides a means for rulers to overcome asymmetries of information that could impede tax collection. We have not considered alternative ways in which rulers could deal with these information asymmetries. The key alternative strategy would be for rulers to develop a bureaucracy that could accumulate localized knowledge about production. Instead of relying on a council representative to communicate about what level of taxation is feasible in a particular area, an appointed bureaucrat could be sent to the area to report the same information. In this section we will explore whether the correlation between council governance and caloric variability depends on whether a bureaucracy is present. The presence or absence of a bureaucracy is most definitely a post-treatment condition, but it may still be useful to consider this possibility. First, however, we will discuss what conditions would be necessary for a bureaucracy to emerge.

There would be clear tradeoffs between relying on bureaucrats versus a local community or its representatives. A bureaucrat appointed by the ruler might have a higher degree of loyalty, in the sense of not biasing they information they provided, but might also have local knowledge. One way for a bureaucracy to overcome the local knowledge problem would be if assessment of local conditions could be facilitated through the application of systematic knowledge. One thing that is clear is that some early societies had much better knowledge than others of how to classify soil according to its productivity. In China the text known as the *Yu Gong*, or the *Tribute of Yu*, holds that the mythic

Table 7: Central and Local Council Substitutability

	[1] Central Council	[2] Central Council	[3] Central Council	[4] Central Council	[5] Local Council	[6] Local Council	[7] Local Council	[8] Local Council
Panel A: *Local Council Present/Absent*								
Caloric Variability × Local Council (0/1)	-0.207	-0.210	-0.199	-0.212				
	(0.036)	(0.043)	(0.043)	(0.049)				
Local Council (0/1)	1.107	1.178	1.088	1.207				
	(0.161)	(0.209)	(0.198)	(0.230)				
Caloric Variability	0.057	0.051	0.054	0.047				
	(0.025)	(0.031)	(0.028)	(0.035)				
Panel B: *Central Council Present/Absent*								
Caloric Variability × Central Council (0/1)					-0.212	-0.200	-0.226	-0.224
					(0.048)	(0.055)	(0.049)	(0.058)
Central Council (0/1)					1.101	1.098	1.189	1.229
					(0.185)	(0.208)	(0.198)	(0.216)
Caloric Variability					0.095	0.111	0.102	0.124
					(0.023)	(0.029)	(0.027)	(0.033)
Region Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Latitude × Longitude Controls	No	No	Yes	Yes	No	No	Yes	Yes
Observations	72	72	72	72	72	72	72	72
Control Mean	0.400	0.400	0.400	0.400	0.640	0.640	0.640	0.640

Note: Each cell reports a separate regression where either *local council* or *central council* is the dependent variable. Robust Standard errors are reported in parentheses.

Emperor Yu, founder of the Xia Dynasty (c.2070 - c.1600 BC) toured different provinces and established different tax levels according to the quality of the soil.[32] While this story is undoubtedly apocryphal, firm evidence does exist that China's rulers began levying taxes according to the quality of the soil from a very early date.[33] Later, by the time of the Song dynasty (960-1279) and the Ming dynasty (1368-1644) Chinese tax assessments involved the use of "fish scale maps" that measured soil quality at a very localized level. No analogous tradition of soil classification existed in European states. By the sixteenth century Machiavelli mused about the fact that more fertile soil would allow for sustaining a larger population for a city, but that was about as far as the argument went.[34] Europeans would not actually begin producing accurate soil maps until the nineteenth century.[35] Under these conditions, even if one could hire a bureaucracy it is not clear that bureaucrats could effectively assess localized asymmetries in production.

Apart from systematic knowledge, the second requirement for a bureaucracy would, of course, be that one has the funds to pay for bureaucrats. This is where the chicken and egg aspect of the problem comes in because one of course needs funds in the first place to hire and train bureaucrats. Besley and Persson (2011) have conceived of state capacity as involving a large initial fixed cost investment.

We make no claim here to be able to answer how, when, and why bureaucracies emerged in SCCS countries. What we will instead limit ourselves to doing is to showing that in the presence of a bureaucracy, there is no longer any correlation between caloric variability and council governance. Table 8 reports a series of estimates where we estimate the correlation between *any council* and *caloric variability* conditional on the presence of full-time bureaucrats who are no related to the government head. To assess the presence of bureaucrats we use the variable coded by Martin Whyte (1978), which is

[32]The Tribute of Yu is part of the larger compilation known as the *Book of Documents*. See the translation by James Legge ([1879] 2016).

[33]See Gong, Zhang, Chen, and Zhang (2003)

[34]*Discourses on Livy* Chapter 1. Book 1.

[35]See Hartemink, Krasilnikov, and Bockheim (2013)

unfortunately available for only half of the SCCS societies.[36] We include this variable in the specifications while including an interaction term between caloric variability and the presence of bureaucrats. Once again, we acknowledge that the presence of bureaucrats is a post-treatment condition. It may still be useful to consider these results.

Across all four specifications we see a very clear result. In the absence of full-time bureaucrats there is a strong, positive correlation between council governance and caloric variability. In the presence of bureaucrats the negative coefficient on the interaction term nearly perfectly offsets the positive coefficient on the council variable, so the estimated effect of *caloric variability* is essentially zero. This is a striking result that provides some evidence that a bureaucracy and a council may be substitute ways of dealing with caloric variability. It is important to note that if though can have substitute functions, there are also many societies in the SCCS dataset that have both a council and bureaucrats

9 Was the Effect of Caloric Variability Persistent?

The theoretical mechanism that we have proposed does not imply any lock-in or long term persistence running from geography to early democracy to democracy today. As the information constraints to which leaders are subject evolve in response to exogenous (or endogenous) changes, we should expect patterns of council governance to evolve as well. Technological change making it easier to observe production is one reason this might happen. In the previous section we explored a reason why endogenous change could impact prospects for council governance if a leader builds a bureaucracy that can be used to better assess taxation, or in other words reduce information asymmetries. Another reason for non-persistence is that in many cases SCCS societies located within a country no longer exist or have been marginalized, and so it is implausible that their political traditions would have been transmitted to the current period. With all this said, authors of two recent papers have considered the specific issue of whether political practices observed in

[36]This is variable 701 in the SCCS data set. It was only coded for half of SCCS societies for reasons of feasibility. The societies coded were selected at random from the full SCCS sample.

Table 8: Councils and Caloric Variability With and Without Bureaucrats

	[1] Any Council	[2] Any Council	[3] Any Council	[4] Any Council
Caloric Variability × Bureaucrat (0/1)	-0.100	-0.109	-0.094	-0.106
	(0.052)	(0.057)	(0.054)	(0.059)
Caloric Variability	0.077	0.092	0.082	0.093
	(0.026)	(0.031)	(0.026)	(0.032)
=1 if full-time bureaucrat present, 0 otw	0.733	0.746	0.671	0.733
	(0.245)	(0.282)	(0.259)	(0.292)
Region Fixed Effects	No	Yes	No	Yes
Latitude × Longitude Controls	No	No	Yes	Yes
Adj. R-squared	0.150	0.127	0.143	0.0986
Observations	81	81	81	81
Control Mean	0.600	0.600	0.600	0.600

Note: Each cell reports a separate regression where *any council* is the dependent variable. Robust Standard errors are reported in parentheses.

societies in Murdock's *Ethnographic Atlas* are correlated with political practices today. In current countries where past societies had leaders that were elected, there tends to be a higher level of democracy today.[37]. The most appropriate current institutional measure for us to consider is the seven point "executive constraints" index from the Polity data set. Just as our *anycouncil* variable is intended to capture the extent to which an executive shares power with another group, the executive constraints index is designed to capture the extent to which individual executives cannot make decisions on their own.

Table 9 reports the results of three alternative specifications where we regress the value of the executive constraint index in a given country in 2016 on either our *any council* or our *jurisdictional hierarchy* variable from the Standard Cross Cultural Sample. One immediate problem with this strategy is that the SCCS societies and current country boundaries match very imperfectly. Some current countries have multiple SCCS societies within them while others have no SCCS societies. Also, a number of SCCS societies are located in small Pacific Island countries that the Polity dataset does not cover. The strategy we adopted was to average SCCS political institutions values for each country. We then regressed this on either our *any council* variable or our *jurisdictional hierarchy* variable, or the two of them simultaneously together with a set of region fixed effects and coordinate fixed effects.

The results of the Table 9 regressions suggest that previous governance by council is indeed associated with a higher level of executive constraints today. They provide no indication that the prior degree of political integration either helped or hindered democracy today.[38] In the first and third specifications the implied magnitude of the effect of having had council governance is large, equivalent to a one step on the seven point executive constraints scale, or more than one half of a standard deviation. Taken together, the Table 9 estimates provide some indication of persistence in the degree of consent-based decisionmaking.

[37] Giuliano and Nunn 2013, and Bentzen et al. 2017

[38] Jacob Hariri (2012) has previously shown evidence that a high degree of early state development tends to be correlated with lower levels of democracy today.

Table 9: Early Councils and Executive Constraints Today

	Polity Index		
	[1]	[2]	[3]
Any Council (0/1)	1.151		1.192
	(0.444)		(0.453)
Political Integration		0.086	-0.055
		(0.204)	(0.231)
Region Fixed Effects	Yes	Yes	Yes
Latitude × Longitude Controls	Yes	Yes	Yes
Adj. R-squared	0.269	0.172	0.257
Observations	73	82	73

Note: Each cell reports a separate regression where the Executive Constraints index from the Polity dataset is the dependent variable. Robust Standard errors are reported in parentheses.

10 Conclusion

In this paper we have provided evidence to suggest that consent-based governance via local or central councils arose independently in a broad set of human societies in numerous different regions. We have also provided evidence that consent-based governance was more frequent in societies where exogenous factors made local agricultural production more variable and thus less transparent. This finding is consistent with a theoretical framework in which those who rule and those who are ruled find it useful to hold a council to exchange information about taxation. We have also emphasized, however, that our empirical results do not support a claim that geography was destiny when it came to explaining the emergence of autocratic rule or the alternative of council governance. Though the effect of caloric variability that we have identified is sizable, it still explains only a small portion of the observed variance in our council measure between different societies. Our results should instead be taken as evidence of a more general phenomenon - when exogenous factors make it harder for rulers to know how much they can tax their citizens, we are more likely to observe consent-based governance. It is also possible, however, that rulers under such circumstances have an alternative route to dealing with information asymmetries. Building a bureaucracy presents an alternative route that avoids the need to govern by council.

References

[1] Acemoglu, Daron and James Robinson. 2012. *Why Nations Fail* New York: Crown Business.

[2] Allen, Robert C. 1997). "Agriculture and the Origins of the State in Ancient Egypt." *Explorations in Economic History* 34:135-154.

[3] Angeles, Luis and Aldo Elizalde. 2017. "Pre-colonial institutions and socioeconomic development: The case of Latin America" *Journal of Development Economics.* 124:22-40.

[4] Baker, Matthew, Erwin Bulte, and Jacob Weisdorf. 2010. "The origins of governments: from anarchy to hierarchy". *Journal of Institutional Economics.* 6:215?242

[5] Bentzen, Jeanet, Jacob Gerner Hariri, and James A. Robinson. "The Indigenous Roots of Representative Democracy."

[6] Besley, Timothy and Torsten Persson. 2011. *Pillars of Prosperity: The Political Economics of Developmental Clusters.* Princeton, NJ: Princeton University Press.

[7] Blanton, Richard and Lane Fargher. 2016. *How Humans Cooperate: Confronting the Challenges of Collective Action.* Boulder: University of Colorado Press.

[8] Blanton, Richard and Lane Fargher. 2008. *Collective Action in the Formation of Pre-Modern States.* New York, NY: Springer.

[9] Blockmans, Wim. 1998. "Representation (since the Thirteenth Century)." In *The New Cambridge Medieval History*, ed. Christopher Allmand. Cambridge: Cambridge University Press.

[10] Boas, Franz. 1913.
textitEthnology of the Kwakiutl *Thirty Fifth Annual Report of the Bureau of American Ethnology to the Secretary of the Smithsonian Institution.* Washington D.C.. Government Printing Office.

[11] Bockstette, Valerie, Chanda, and Louis Putterman. 2002. "States and Markets: The Advantage of an Early Start." *Journal of Economic Growth.* 7:347-369.

[12] Boix, Carles. 2015. *Political Order and Inequality: Their Foundations and Their Consequences for Human Welfare.* New York, NY: Cambridge University Press.

[13] Carneiro, Robert L. 1970. "A Theory of the Origin of the State,." *Science.*" 169:733-738.

[14] Carrasco, Pedro. 1971. "Social Organization of Ancient Mexico." In *Handbook of Middle American Indians* - Volume 10, eds. Gordon Ekholm and Ignacio Bernal. Austin: University of Texas Press. 349-75.

[15] Dahl, Robert. 1998. *On Democracy* New Haven. Yale University Press.

[16] Fargher, Lane F., Verenice Y. Heredia Espinoza, and Richard E. Blanton. 2011. "Alternative Pathways to Power in Late Postclassic Highland Mesoamerica." *Journal of Anthropological Archaeology.* 30: 306-26.

[17] Ferejohn, John and Frances Rosenbluth. 2016. *Forged Through Fire: War, Peace, and the Democratic Bargain.* New York: W.W. Norton.

[18] Fleming, Daniel. 2004. *Democracy's Ancient Ancestors: Mari and Early Collective Governance.* Cambridge: Cambridge University Press.

[19] Galor Oded and Omer Ozak. 2016. "The Agricultural Origins of Time Preference." *American Economic Review.* 106:3064-3103

[20] Galor, Oded and Omer Ozak. 2015. "Land Productivity and Economic Development: Caloric Suitability vs. Agricultural Suitability."

[21] Gennaioli, Nicola and Ilya Rainer. 2007. "The modern impact of precolonial centralisation in Africa." *Journal of Economic Growth.* 12:185-234

[22] Gibson, Charles. 1947. "Lewis Henry Morgan and the Aztec 'Monarchy'." *Southwestern Journal of Anthropology.* 3:78-84.

[23] Giuliano, Paola, and Nathan Nunn. 2013. "The Transmission of Democracy: From the Village to the Nation-State." *American Economic Review: Papers & Proceedings.*

[24] Gong, Zitong, Xuelei Zhang, Jie Chen, and Ganlin Zhang. 2003. "Origin and Development of Soil Science in Ancient China." *Geoderma* 115:3-13.

[25] Hansen, M. H. 1986. "The Origin of the Term Demokratia." *Liverpool Classical Monthly.* 11.3:35-36.

[26] Hariri, Jacob Gerner. 2012. "The Autocratic Legacy of Early Statehood." *American Political Science Review* 106: 471-94.

[27] Hartemink, Alfred, Pavel Krasilnikov, and J.G. Bockheim. 2013. "Soil Maps of the World." *Geoderma* 207:256-267.

[28] Huning, Thilo, and Fabian Wahl. 2016. "You Reap What You Know." Working Paper no.101. European Historical Economics Society.

[29] Legge, James. 2016 [1865] *Book of Documents, Shangshu: Bilingual Edition, Chinese and English: Chinese Classic of History* CreateSpace Independent Publishing Platform.

[30] Levi, Margaret. 1988. *Of Rule and Revenue.* Cambridge Cambridge University Press.

[31] Loyn, Henry. 1984. *The Governance of Anglo-Saxon Engand, 500-1087.* Stanford: Stanford University Press.

[32] Maddicott, J.R. 2010. *The Origins of the English Parliament, 924-1327.* New York: Oxford University Press.

[33] Mayshar, Joram, Omer Moav, and Zvika Neeman. 2017. "Geography, Transparency and Institutions." *American Political Science Review* 111:622-636.

[34] Mayshar Joram, Omer Moav, Zvika Neeman, and Luigi Pascali. 2017. "Cereals, Appropriability and Hierarchy."

[35] Michalopoulos, Stelios and Elias Papaioannou. 2013. "Pre-Colonial Ethnic Institutions and Contemporary African Development." *Econometrica.* 81:113-52.

[36] Michalopoulos, Stelios and Elias Papaioannou. 2014. "National Institutions and Subnational Development in Africa,." *Quarterly Journal of Economics.* 129, 151-213.

[37] Morgan, Lewis Henry. 1877. *Ancient Society* New York: Henry Holt and Company.

[38] Murdock, George Peter. 1967. *Ethnographic Atlas* Pittsburgh: University of Pittsburgh Press.

[39] Murdock, George Peter and Diana Morrow. 1970. "Subsistence Economy and Supportive Practices: Cross-Cultural Codes 1." *Ethnology* 9:302-330.

[40] Murdock, George Peter and Caterina Provost. 1973. "Measurement of Cultural Complexity." *Ethnology* 12:379-392.

[41] Murdock, George Peter and Douglas R. White. 1969. "Standard Cross-Cultural Sample." *Ethnology.* 8:329-369.

[42] Murdock, George Peter and Suzanne Wilson. 1972. "Settlement Patterns and Community Organization: Cross-Cultural Codes." *Ethnology* 3 11:254-295.

[43] North, Douglass. 1981. *Structure and Change in Economic History* New York: W.W. Norton.

[44] North, Douglass C., and Barry R. Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England." *The Journal of Economic History.* 49: 803-32.

[45] Osafo-Kwaako, Philip and James Robinson. 2013. "Political centralization in pre-colonial Africa." *Journal of Comparative Economics.* 41:6-21.

[46] Scott, James. 2017. *Against the Grain.* New Haven Yale University Press.

[47] Stasavage, David. 2010. "When Distance Mattered: Geographic Scale and the Development of European Representative Assemblies." *American Political Science Review* 104: 625-43.

[48] Stasavage, David. 2011. *State of Credit: Size, Power, and the Development of European Polities.* Princeton: Princeton University Press.

[49] Tracy, James. 1994. "Taxation and State Debt." In *Handbook of European History, 1400-1600*, ed. Heiko Oberman Thomas Brady, and James D. Tracy. Leiden: E.J. Brill.

[50] Trigger, Bruce 1976. *The Children of Aatasentsic: A History of the Huron People to 1660*. Montral: Mcgill-Queen's University Press.

[51] Trigger, Bruce. 2003. *Understanding Early Civilizations*, New York Cambridge University Press.

[52] Tuden, Arthur and Catherine Marshall. 1972. "Political Organization: Cross-Cultural Codes 4." *Ethnology*. 11:436-464.

[53] Vansina, Jan. "The Kuba State." in Henri J.M. Claessen (ed.). *The Early State*. New York: Mouton.

[54] Wilks, Ivor. 1975. *Asante in the Nineteenth Century*. New York: Cambridge University Press.

[55] Whyte, Martin. 1978. *The Status of Women in Preindustrial Societies.* Princeton: Princeton University Press.