Online Appendix

A. Theory

A.1. A Formal Model. This section presents a simple formal model of state consolidation. Its main goal is to highlight the channels through which environmental circumscription determines the ability to collect taxes, and how this affects political instability. The degree of circumscription changes due to productivity variations in the core and/or the hinterland. However, productivity shocks in these locations manifest themselves in different ways. On the one hand, shocks at the core have a *direct effect* on the tax collection, as they determine the total amount to be taxed. As a result, they have a contemporaneous effect on tax revenues. On the other hand, shocks at the hinterland have an *indirect effect*, as they manifest themselves by creating migration pressures which typically take a longer time to build up. This implies that their effect will become apparent in future periods.

The main features of the model are as follows. A state has developed in a territory (core) which is surrounded by a territory that lies outside the control of the state (hinterland).[35] The main difference between these areas is that the main activity at the core is easily taxable while that in the hinterland is very difficult to tax. Our goal is to analyze how political stability in this state responds to changes in the productivity of both the core (c) *and* the hinterland (h). As mentioned above, the model below can be interpreted in a more general way. Under this more general interpretation, c is an activity that is easily taxable (typically, growing cereals), while h is any other activity that is difficult to tax, regardless of where production takes place. For instance, h can be interpreted as growing tubers, as in Mayshar et al. (2022), and the main features of the model that follows would apply in a similar way. Thus, our approach invites new thinking and empirical work about alternative activities under h.

There are just two time periods, $t = \{1, 2\}$ and the world ends after period 2. Although artificial, this two-period approach allows us to highlight the main points of interest avoiding unnecessary complications.

A.1.1. Individuals. There is a population of size 1 that at time $t = \{1, 2\}$ lives in the core (c), L_t^c, or in the hinterland (h), $L_t^h = 1 - L_t^c$. The productive activities in those areas are different.[36] In the core, individuals have 1 unit of land and average productivity per worker is given by A_t^c. Productivity depends on a weather shock such that A_t^c will be high (low) whenever weather

[35]Thus, in our analysis we take the "core" and the "hinterland" territories as given, being exogenously determined by geographic and climatic conditions.

[36]Typically, the main activity in the core is cereal cultivation. As convincingly argued by Scott (2017) and Mayshar et al (2022), cereal crops have a number of characteristics (storability, transparency, divisibility) that reduce taxation costs making extraction feasible. By contrast, the activities developed in the hinterland are varied, such as hunting and gathering, cattle raising, etc.

at t is good (bad). Average productivity in the hinterland is given by A_t^h, which depends on a different weather shock in a similar fashion as A_t^c. Since typically weather shocks are autocorrelated, we assume that $A_{t+1}^j = \rho^j A_t^j + r_{t+1}^j$, with $\rho^j \in (0,1)$, $j = \{c, h\}$.[37] We also assume that the shocks r_t^c and r_t^h are mutually independent, i.i.d. random variables with positive support and means $\{\mu^c, \mu^h\}$, respectively.

Each individual i has an ability (α_i, β_i) to carry out the activity developed in the core and in the hinterland, respectively. The production achieved by individual i in location c is $\alpha_i A_t^c$, and in location h is $\beta_i A_t^h$. Let $\psi_i = \beta_i/\alpha_i$ be the relative ability of i to execute the activity developed in h, whose distribution over the entire population is given by $F(\cdot)$.

All the individuals in c are subject to a tax rate τ on their produced output (see below). To simplify matters, we assume that people in the hinterland are not taxed at all.

Timing is as follows: after realising A_t^c and and A_t^h at the beginning of $t = 1$, individuals produce an output according to their ability and their location, pay taxes accordingly, and derive utility from consuming whatever is left. The only decision individuals take at the end of time $t = 1$ is whether to stay where they are or to migrate at the end of this period. We assume no costs of migration which implies that individuals can decide whether to stay or to migrate considering each period in isolation. Individual i will decide at the end of $t = 1$ to stay at the core in $(t + 1)$ if the expected value of being at the core is higher than the expected value of being at the hinterland, that is,

$$E_t((1-\tau)A_{t+1}^c | A_t^c) > E_t(\psi_i A_{t+1}^h | A_t^h),$$

where $E_t(\cdot)$ denotes expected value conditional on the information available at time t. Under the assumptions above, there is a threshold $\overline{\psi}_t$ defined as

$$\overline{\psi}_t = (1-\tau)\frac{\rho^c A_t^c + \mu^c}{\rho^h A_t^h + \mu^h}, \tag{1}$$

such that individuals with $\psi_i < \overline{\psi}_t$ ($\psi_i > \overline{\psi}_t$) will be in the core (hinterland), respectively at time $(t + 1)$. Therefore, the population at time $t + 1$ in c is

$$L_{t+1}^c = F(\overline{\psi}_t). \tag{2}$$

It follows that the larger the productivity gap between the core and the hinterland at time $t = 1$, the higher the share of population in the core at time $t + 1$. From the characterization of the threshold $\overline{\psi}_t$, it is straightforward to see what the determinants of L_{t+1}^c are, which for convenience are summarized in the following Lemma.

Lemma 1. *The size of the population in c at time t + 1*

[37] Under these assumptions, A_j is a stationary, mean-reverting process.

(1) depends negatively on the tax rate τ;
(2) depends positively on the productivity gap between the core and the hinterland at $t = 1$; Specifically, weather shocks that increase A_t^c or that decrease A_t^h will increase population in c at time $t + 1$.

A.1.2. The state. We now describe the role of the state in our model and introduce our measure of political instability.

State revenue. The main source of state's revenue are the taxes collected from the population living in c. For simplicity, it is assumed that τ is constant and exogenous (or conditioned by a long-term commitment by the ruler).[38] The tax revenue (TR) at t is given by

$$TR_t = \tau E_t^c(\alpha) A_t^c L_t^c, \tag{3}$$

where $E_t^c(\alpha)$ denotes the expected value of α for the individuals living in c at time t.

From (3), it is easy to see the link between weather shocks and TR_t. A weather shock that increases the productivity of the core, A_t^c, *at time* $t = 1$, will have a *direct* effect on the tax revenue, as it will increase the amount of taxable output. On the other hand, a shock that increases the productivity of h *at time* $t = 1$ and/or a shock that decreases the productivity of c in the same period will have a (negative) *indirect* effect on the tax revenue in the following period, TR_{t+1} through its effect on the population at the core at $t = 1$, L_t^c. This is due to the fact that, ceteris paribus, an increase in A_t^h (or a decrease in A_t^c) will reduce $\overline{\psi}_t$ and, therefore, will have a negative impact on the population at the core at $t + 1$.

The state faces a fixed cost κ of maintaining the state apparatus and the infrastructure needed to enforce tax payments, so that the net revenue is given by $NR_t = TR_t - \kappa$.[39] This simple expression highlights some of the key factors for state formation and consolidation. For the state to be viable, the net revenue has to be positive, i.e., $NR_t \geq 0$. Mayshar et al. (2022) and Scott (2017) focus on the importance of a low κ for state formation, which in their theory is attained by growing cereals, as cereals can be taxed cheaply. Circumscription can be thought of as another factor that contributes to a large NR_t: on the one hand it tends to increase TR_t since it contributes to having a larger population and a (relative) high land productivity; and on the other it lowers κ, since costs associated with retaining the population and with enforcing tax payments (i.e., by building walls or by funding a bureaucracy to monitor workers, etc.) will go down with circumscription.

[38]Alternatively, one could consider that each period the state chooses the tax rate optimally to maximise taxation. Results would be very similar, and therefore we adopt the simpler interpretation that the tax rate is constant.

[39]More realistically, the cost of maintaining the state also has a variable component that is increasing with population. We omit this term and consider instead the fixed cost exclusively to highlight the fact that in the very sparsely populated world we're considering, a necessary condition for state formation and consolidation was to count with sufficient population density in order to be able to cover the fixed cost component of state management.

Political instability and state capacity. Political instability can take several forms. One extreme form of political instability might occur if the state ceases to be viable, i.e., if the costs of maintaining it exceed the revenues. If the tax collection is unable to cover the costs of maintaining the state (i.e., $TR_t < \kappa$), then the state will be forced to dissolve and central rule will break up.

Other forms of political instability can also affect the state without provoking its collapse. In the following we focus on political instability arising in periods where the state viability condition ($TR_t > \kappa$) is satisfied and therefore central rule can be maintained.

The executive head of state is a ruler who controls the revenue collected by the state bureaucracy. In cases where $TR_t > \kappa$, the ruler can spend state revenue on two basic categories of goods: defensive capacity a_t, and state consumption P_t. Defensive capacity a_t includes primarily the cost of an army.[40] State consumption P_t includes cult centers and elite tombs, like temples and pyramids, but might also be thought of as loot from military campaigns in foreign lands and luxury consumption of no value to anyone but to the ruling elite.[41] We make the key assumption that the incumbent ruler aims to maximize state consumption, P_t, which is entirely consumed each year.

Every period, the ruler and his elite face an attack of intensity d_t, which is observed at the beginning of t, together with A_t^i, $i = \{c, h\}$. This quantity is a realisation of a random variable D_t, which has strictly positive support. After observing d_t, the ruler decides how much to spend on his defensive capacity a_t. The probability that he will defeat the attack is given by a standard contest function $\eta(a_t) = \frac{a_t}{d_t + a_t}$. If the aggression is defeated, state consumption is $P_t = TR_t - \kappa - a_t$. However, with a probability $\left(1 - \eta(a_t)\right)$, the incumbent ruler is defeated and state's consumption P_t and utility is zero.[42] We assume that in addition to the loss of state revenue, defeat can also bring about political turmoil, such as the destitution of the ruler, who will be replaced by a new ruler with identical preferences in the next period and who will also control the taxes collected by the state bureaucracy. Hence, a_t is chosen by the ruler such that the expected utility is maximimised, where the latter is given by[43]

$$E(P_t) = \frac{a_t}{d_t + a_t}(TR_t - \kappa - a_t).$$

[40] See Dal Bo et al (2016) for a model that focuses on how the rulers of early states had to invest in defensive capacity to ward off predatory challengers.

[41] The Egyptian population seemed to have believed that pyramids potentially strengthened the ruler's divine powers, which in turn might benefit the people through optimal levels of Nile floods abundant harvests. Although we recognize that such beliefs probably played a role in the state consolidation process, we abstain from explicitly modelling them here.

[42] This formulation implicitly assumes that conflict does not destroy output in case of victory. It also makes the simplifying assumption that the attack only affects State consumption, P_t, but not individual consumption.

[43] For simplicity, this formulation additionally assumes that the outcome of conflict at t does not affect P_{t+1}, which implies that each ruler's decisions can be considered in each period in isolation.

Provided the state is viable ($TR_t - \kappa > 0$), the solution to the above-mentioned problem will be interior ($a_t^* > 0$) iff $d_t > 0$. In the following, we focus on the non-trivial cases where the solution is interior.[44] In these cases, the optimal size of the army, a_t^*, is

$$a_t^* = -d_t + \sqrt{d_t(d_t + TR_t - \kappa)}. \tag{4}$$

Equation (4) implies that military expenditure a_t increases with the size of the threat d_t and with the tax collection TR_t.

So far we have considered that d_t is simply a realization from a random variable D_t. However, it is reasonable to expect that D_t and TR_t are related. The sign of this association is likely to be related to the ultimate goal of the attack. There is ample support in the literature that economic downturns (that in our framework translate into a reduction in TR_t) increase the probability of political instability (Acemoglu and Robinson 2006, Burke and Leigh 2010, Chaney 2013, etc.). Periods of economic stress facilitate the organization of internal revolts, as factions of the elite can take advantage of the weakness of the state (and thus, the smaller size of mobilized defensive resources) and the discontent in the population (due to scarcity and famine) to organize an insurrection and replace the leader. This is the so-called "opportunity cost" effect. An alternative scenario would be one in which the aggressor is primarily interested in conquering TR_t, in which case D_t are TR_t will be positively related (the "rapacity" effect, see Dube and Vargas, 2013).

In the case of Egypt, we argue that $\lambda \leq 0$ could possibly be the most relevant situation. The pharaoh was believed to be responsible for the Nile floods. Then, in case of famines derived from extreme Nile floods he was often directly blamed for them (Bell, 1971). This could have led to internal revolts, whereby factions of the elite could have taken advantage of the discontent in the population and/or of the weakness of the army to organize an insurrection to depose the ruler.

A simple way to model this is to consider that D_t is a function of the same weather shocks as TR_t such that $D_t(w_t) = \lambda TR_t(w_t) + \chi_t$, where w_t is a vector of variables containing the relevant weather shocks and χ_t is a random variable independent from w_t, with strictly positive support such that D_t also has strictly positive support.

The following proposition summarizes the relationship between political instability and economic shocks under the above-mentioned assumptions.

Proposition 1. *Under the assumptions above, a weather shock that deteriorates TR_t will increase political instability* $(1 - \eta(a_t^*))$. *More specifically, ceteris paribus, a reduction in A_t^c, A_{t-1}^c or an increase in A_{t-1}^h, will increase political instability at t.*

[44]Provided $TR_t - \kappa \leq 0$, a corner solution ($a_t^* = 0$) will be obtained if d_t is zero. In this case political instability is trivially 0.

Proposition 1 is the main result of the paper and implies that, ceteris paribus, an increase in state revenue will lead to a decrease in the probability that the ruler is defeated. If D_t and TR_t are negatively related ($\lambda < 0$), the intuition of this result is straightforward. In this case, an increase in state's revenue is associated with a decrease in the size of the threat, and therefore the state has more resources to defend itself from a less severe attack. When D_t and TR_t are positively related, both the size of the threat and state revenue move in the same direction. As shown in Appendix A.2, an increase in TR_t will lead to a decrease in political instability if the elasticity of D_t with respect to the net revenue, NR_t, is smaller than 1, i.e., if the relative change in state revenue is larger than the relative change in the size of the threat. If this is the case, the state's additional resources will overcompensate the increase in the size of the threat, leading to a decrease in the probability of ruler change. Appendix A.2 shows that under the simple framework considered above this is always the case. This implies that TR and political instability move in opposite directions, regardless of the sign of λ, i.e., regardless of whether the attack is motivated by "rapacity" ($\lambda > 0$) or "'oportunity cost" ($\lambda < 0$).

Additionally, it's possible to establish some predictions about the evolution of P_t as a function of the productivity shocks. For simplicity, Proposition 2 focuses on the case where $\lambda = 0$. Appendix A.2 provides the proof of this proposition and discusses the case where $\lambda \neq 0$.

Proposition 2. *Under the assumptions above and $\lambda \leq 0$, a weather shock that increases TR_t will increase $E(P_t)$.*

See Appendix A.2 for the proof of this proposition. The intuition of this result is simple. An increase in TR_t reduces political instability (Proposition 1), and a fraction of the increase in TR_t will be devoted, ceteris paribus, to state consumption, as the size of the threat will be unaffected by the increase in TR_t.

The case where $\lambda > 0$ is more ambiguous as now both TR_t and d_t will move in the same direction. Therefore, depending on the relative sizes of the movements, the increase in TR_t might be invested in a_t rather than in P_t to counter the attack. In the case of Egypt, however, we argue that the case where $\lambda \leq 0$ could be the most relevant one. The pharaoh was believed to be responsible for the Nile floods. Then, in case of famines derived from extreme Nile floods he was often directly blamed for them (Bell, 1971). This could have led to *internal revolts*, whereby factions of the elite could have taken advantage of the discontent in the population and/or of the weakness of the army to organize an insurrection to depose the ruler.

In summary, our dynamic model of environmental circumscription predicts that negative weather shocks in the core will lead to more political instability and lower state capacity directly through lower tax revenues (causing a weaker defensive capacity) and indirectly through a lower level of population (and, therefore, total taxable production) in the next period. A negative weather shock in the hinterland will indirectly lead to less political instability and more state capacity due to a greater level of population in the core in the next period. These hypotheses are tested in the empirical section.

A.2. Proofs.. This section presents the proof to Propositions 1 and 2.

Proof of Proposition 1.

Proof. Let's consider the case where $D_t(w_t) = \lambda TR_t(w_t) + \chi_t$, where w_t is a vector of variables containing the relevant weather shocks and χ_t is a random variable independent from w_t, with strictly positive support such that D_t also has strictly positive support. Proposition 1 focuses on the case where $\lambda = 0$, however in this proof we also discuss the case where λ is different from zero.

Under the assumptions in the main text, D_t has strictly positive support which implies that every year there is a strictly positive threat, d_t. Consider a weather shock w_t that increases TR_t (either because it increases A^c_t or A^c_{t-1} or because it decreases A^h_{t-1}). Thus, $TR'_t(w_t) > 0$. We are interested in the sign of the first derivative of $(1 - \eta(a^*))$ as a function of w_t. This first derivative is given by

$$\frac{\partial(1-\eta(a^*))}{\partial w} = \frac{D_t[D'_t(TR_t - \kappa) - D_t TR'_t]}{2\Lambda^{3/2}} \tag{5}$$

where $\Lambda = D_t(D_t + TR_t - \kappa) > 0$.

If $D'_t = 0$ (i.e., if $\lambda = 0$), then (5) is negative as the first term in the numerator will vanish and the second term is negative.

If $D'_t < 0$ (i.e., if $\lambda < 0$), then (5) is negative, as both terms in the numerator are negative. This implies that shocks that deteriorate TR_t will increase political instability.

Finally, if $D'_t > 0$, (i.e., if $\lambda > 0$), then (5) is negative provided

$$\frac{D'_t/D_t}{TR'_t/(TR_t - \kappa)} < 1.$$

that is, if the elasticity of D_t with respect to NR_t is smaller than 1. Using the assumed relationship between D_t and TR_t it follows that,

$$\frac{D'_t/D_t}{TR'_t/(TR_t - \kappa)} = \frac{\lambda TR_t - \lambda\kappa}{\lambda TR_t + \chi_t} < 1,$$

since λ, κ and are positive numbers and χ_t has strictly positive support. ∎

Proof of Proposition 2

Proof. In this proposition we are interested in the sign of the first derivative of $E(P_t)$ as a function of a shock w_t that increases TR_t (either because it increases A^c_t or A^c_{t-1} or because it decreases A^h_{t-1}). We consider here the case where λ=0 and discuss at the end the general case. Under the assumptions of Proposition 2 this derivative is given by

$$\frac{\partial E(P_t)}{\partial w} = \eta(a_t)'(TR_t - \kappa - a_t) + \eta(a_t)(TR'_t - a'_t) \tag{6}$$

The first term in equation (6) is positive, since $\eta(a_t)'$ is positive (Proposition 1) and the budget needs to balance every year (therefore $(TR_t - \kappa - a_t)$ cannot be negative). The second term is positive provided $TR'_t \geq a'_t$. Notice that

$$TR'_t - a'_t = TR'_t(1 - \frac{d_t}{2(d_t^2 + d_t NR_t)^{1/2}}) \geq 0 \tag{7}$$

The last inequality follows because $TR'_t > 0$ and $d_t NR_t$ is also positive, which implies that $d_t/2(d_t^2 + d_t NR_t)^{1/2}$ is smaller than 1.

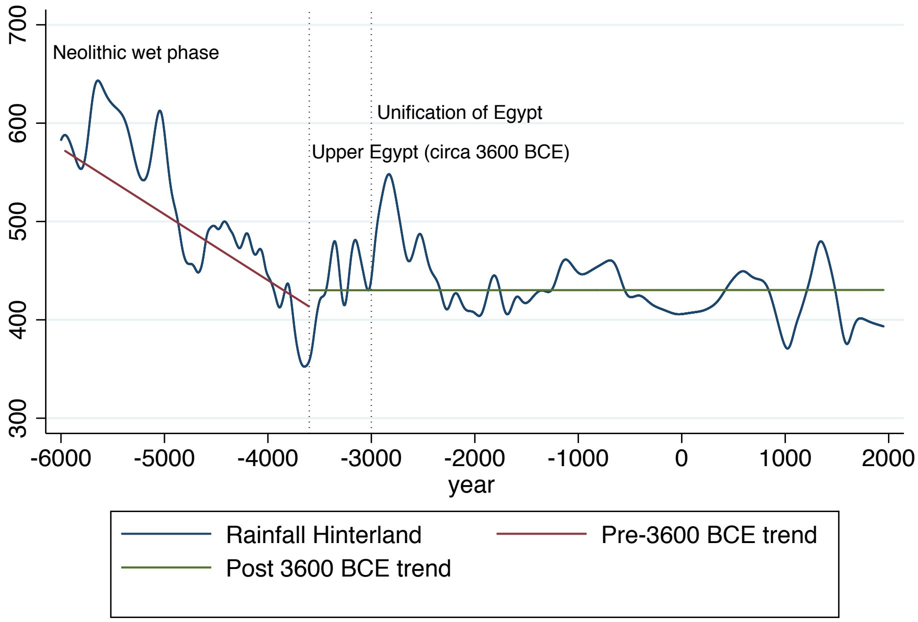
If λ is allowed to be different from zero, the same result as that shown above will hold in the case where $\lambda < 0$. In this case, as TR_t goes up the size of the threat goes down and therefore the pressures for increasing a_t will also go down. The case where $\lambda > 0$ is more ambiguous as now both TR_t and d_t will move in the same direction. Therefore, depending on the relative sizes of the movements, the increase in TR_t might be invested in a_t rather than in P_t to counter the attack.

In the case of Egypt, however, we argue that the case where $\lambda \leq 0$ could be the most relevant one. The pharaoh was believed to be responsible for the Nile floods. Then, in case of famines derived from extreme Nile floods he was often directly blamed for them (Bell, 1971). This could have led to *internal revolts*, whereby factions of the elite could have taken advantage of the discontent in the population and/or of the weakness of the army to organize an insurrection to depose the ruler. ∎

B. Historical Background: Additional documentation.

B.1. The Neolithic wet phase and the creation of the Egyptian state. Figure B.1 plots our proxy for rainfall in the areas surrounding the Nile Valley, obtained from the Soreq cave. As referred to in the main text, the first state in Egypt was unified around 3000 BCE. The emergence of the Egyptian state roughly coincides with the most severe local drought during the last 8000 years.

Figure B.1. NEOLITHIC WET PHASE AND EGYPTIAN UNIFICATION .



B.2. Literature overview. In this subsection we present a literature overview of theories of early state formation and development with a focus on Carneiro's (1970) model of environmental circumscription. We then review existing work on the impact of hydrological shocks on political stability on historical Egypt. As in Carneiro (1970), by *early state* we mean an autonomous political unit with centralised government that encompasses many communities within its territory and has the power to collect taxes, draft men and decree and enforce laws.

In the analysis below, we distinguish between two different phases in the "career" of early states: (i) The process leading up to the *initial state formation* and (ii) the state's *early development* of its capacity.[45] Mayshar et al (2017) assign a key role to the *transparency of farming* for the rise and development of the early states in Egypt and Mesopotamia. In Egypt, Nile floods in June-Oct were highly predictable and readily observable, with the result that taxation in the form of appropriated cereals was relatively efficient. These geographical and technological conditions led to a strongly centralized state that existed (with notable interruptions) for

[45]Ever since the days of Plato and Aristotle, scholars have studied the question of why and how the first states were created. Classical contributions include works such as Weber (1919), defining a state as an entity that upholds the monopoly of power and Wittfogel's (1957) *hydraulic hypothesis*, arguing that the natural need for central coordination in early irrigation agriculture led to highly centralized governments in Mesopotamia and Egypt. See Bentzen et al (2016) for an empirical test of this hypothesis. More recently, Spencer (2010) outlines a *territorial expansion model* claiming that there was a close correspondence in time between the first appearance of state organization and the earliest expansion of the state's control of regions beyond more than one day of travel from the center.

thousands of years. The more erratic and less transparent inundations of Euphrates and Tigris in Mesopotamia implied a greater number of independent political units and weaker states. In a related paper, Mayshar et al (2022) argue that historical political entities in territories relying on cereals typically saw the emergence of more stratified states earlier than in territories where tubers and roots were the staple crop of the population. The key reason for this regularity was the greater *appropriability* for taxation of cereals compared to tubers. A similar analysis is offered by Scott (2017).

Two recent papers focus on early state development in Mesopotamia. Allen et al (2023) exploit random shifts to the course of Euphrates and Tigris and show that city state formation is more likely to emerge when a dependable source of water diverges away from a site. The authors argue that this tendency should be understood as a response to the demand for coordinated irrigation facilities in already established communities. Benati and Guerriero (2021) argue that a state was formed in Mesopotamia when adverse production conditions pushed the existing elite to establish a state by giving strong political and property rights to non-elite groups that possessed complementary skills.

Carneiro (1970) offered an alternative hypothesis based on coercion and conflict. A central argument was that every early society where people had adopted some form of sedentary agriculture in a core region would sooner or later run into a conflictual situation when population growth made further geographical expansion impossible. In this encounter between growing farming communities, there would inevitably at some point arise a conflict over available farmland that could not be solved peacefully. Hence, war would typically ensue in which some group would win. In Carneiro's hypothesis, the crucial element is what happens to the losers of the war. There are essentially two plausible scenarios: The losing group(s) will either leave the core region and settle in a hinterland in order to escape the dominant rule of the winners, or they will stay and let themselves be exploited by the winning group.

A key factor that influenced this choice was, according to Carneiro (1970), the core region's degree of *environmental circumscription*. If the level of environmental circumscription around the agricultural core was high, perhaps due to surrounding deserts, mountains, or oceans, then the losing group would have no choice but to stay and let themselves be ruled by the winners. If environmental circumscription was low so that the losers could for instance hide in a forest in the hinterland, then the losers would typically rather escape than letting themselves be ruled by the winners. According to this argument, the consolidation into a larger political unit such as a state was easier if environmental circumscription of the core was high and that we should find the first states in places with a large productivity gap between the core and the surrounding territories. This empirical prediction finds support in Schönholzer's (2020) global analysis of 176,512 grid cells.

Carneiro (1970) discusses Egypt as an example of a circumscribed region where a state arose early, a view which is further elaborated by Allen (1997). Like us, Allen (1997) considers how the Egyptian state was able to successfully maintain power over such a long time period. He concludes that a key factor was the ability of the elite to extract taxes, which in turn depended importantly on the good storability of cereal harvests and the immobility of farmers in times of a dry hinterland.

The theory of the importance of environmental circumscription for original state formation can easily be extended into a dynamic model of early state development.[46] Effective circumscription might be thought of as reflecting the gradient in agricultural productivity between a core and a hinterland. If this gradient is very high due to an extremely productive core and a more or less unproductive hinterland, then the area is highly circumscribed and political stability and levels of taxation should be high. However, over time, climate change or weather shocks might cause changes in the core/hinterland productivity gradient which also change the strength of political consolidation. For instance, a shock that increased the relative productivity of the hinterland would make a successful escape from the core more likely and would therefore cause decreased political stability and lower state capacity in the core. In the conceptual framework below, we present an informal model of such an intertemporal interaction between exogenous environmental circumscription and political stability.

B.3. Exogenous shocks and political instability in historical Egypt. Two recent studies have examined the relationship between exogenous geophysical shocks and political instability in historical Egypt. Using the Roda Island (Cairo) Nilometer data from the Muslim period in Egypt, Chaney (2013) studies the impact of extreme Nile floods on the probability of replacement of the Head Judge, i.e. the main religious authority in the country during the Abbusid dynasty. The empirical study covers mainly the years 1169-1425 CE. Using text data from an annual chronicle of the period, Chaney (2013) finds that extreme Nile floods are associated with more references to high food prices and a more frequent use of words like "combat, riots and looting". However, the main result of the paper is that extreme floods implied a lower risk of replacement of the Head Judge and a greater construction of religious monuments.[47] Chaney's interpretation is that extreme floods strengthened the position of the Head Judge vis-à-vis the Sultan since revolts against the latter were typically

[46]Our dynamic model of state development shares certain features with the analysis in Olson (1993) and Sanchez de la Sierra (2020). In the spirit of Olson's account, gangs of "roving bandits" might terrorize the country and create disorder in times of famine. In times when resources are more abundant, opportunistic leaders and their followers will instead settle down and become "stationary bandits" who invest in tax extraction capacity in a state-like organization. Sanchez de la Sierra (2020) studies similar dynamics in DR Congo where an increase in the price of a highly transparent mineral (coltan) provided incentives for roaming warlords to settle down, exploit mineral resources, and install state-like institutions. See also Hidalgo et al (2010) for an account of how negative contemporary rainfall shocks led to land invasions by poor farmers.

[47]In contrast to our main result, the evidence relating extreme floods and greater risk of sultan replacement (the main political authority) seems to be weaker in this period.

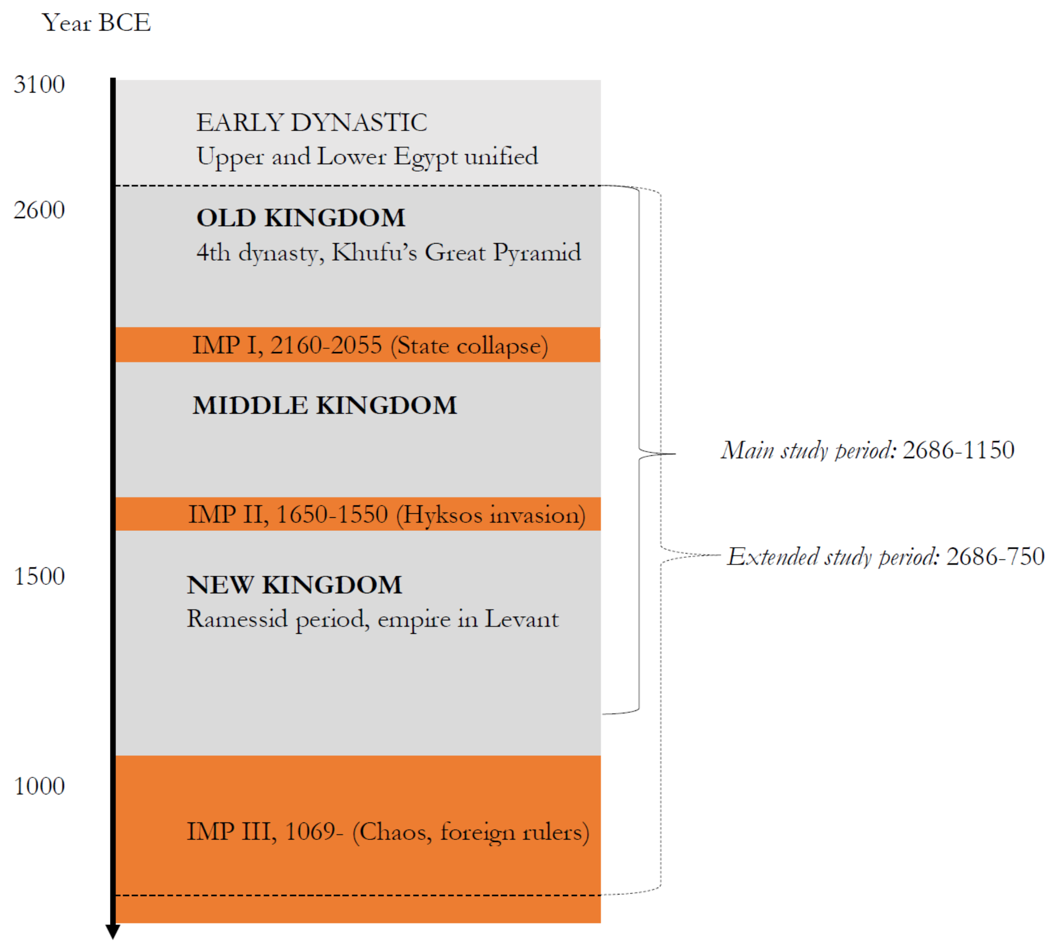
only successful in case of Head Judge support. In general, the results in Chaney (2013) support the hypothesis that extreme floods in the core cause famines and hardships, which in turn lead to social unrest and political instability. However, Chaney (2013) does not study the interaction between conditions in the core and in the hinterland.

In a recent study covering the Ptolemaic period (332-30 BCE), Manning et al (2017) analyze how volcanic eruptions outside Egypt affected monsoon rains negatively, which in turn led to lower Nile floods. The association between volcanic eruptions and low Nile floods is established using the Roda Island Nilometer data from 622-1902 CE. The authors use Ptolemaic era records on interstate warfare and the frequency of revolts, priestly decrees, and land sales, as political outcomes variables and analyze how the probability of such events are affected by volcanic eruptions during the Ptolemaic era (the authors have no quantitative Nile flood data for this period). Their results show that eruption years (and the associated, unobserved low Nile floods) imply a greater risk of revolt onsets, in particular two years after the eruption. Terminations of wars with Syria are more likely to happen 0-2 years after an eruption, indicating a weakening of state capacity. The greater use of priestly decrees in eruption years is interpreted as indicating a stronger need for the ruling class to reinforce state rule during crises, whereas a greater prevalence of land offered for sale suggests a greater economic stress. Just as in Chaney (2013), Manning et al (2017) focus on conditions in the Nile core regions.

B.4. Rulers and Dynasties: Chronologies and Dating issues. Our main source of information about ruler and dynasty tenure periods is Shaw (2000). This book, in turn, builds on numerous historical sources such as *Aegyptiaca* by the 3nd century BC historian Manetho, as well as inscriptions on monuments and other archeological artefacts such as the *Palermo stone* that offer chronologies of events during certain rulers. Modern scholars have further used astronomical observations, referred to in ancient texts, of lunar phenomena with known regular cycles in order to anchor ruler reigns to specific years. Despite these efforts, there are several remaining measurement issues with the exact dating of ruler tenures, including the recurring periods of fundamental uncertainty during intermediate periods when many state functions collapsed.

In a novel approach, Bronk Ramsey et al (2010) used 211 radiocarbon measurements made on museum samples from plants that could definitely be tied to a certain pharaoh. The samples included seeds, baskets, and plant-based textiles. The authors then performed standard radiocarbon measurement of the samples using accelerator mass spectrometry. In this way, the authors constructed radiocarbon-based chronology of ruler periods during 2650-1100 BCE, i.e. during the classical period we study in the paper. When the authors compared ruler tenure lists, they found that their radiocarbon dates were in strong agreement with Shaw (2000), whereas an alternative lists such as Hornung et al (2006) cited dates for ruler

Figure B.2. Ancient Egypt Chronology.



Notes: This graph depicts the main periods in Ancient Egypt's chronology. Source: Shaw (2000).

ascensions that were typically 40-80 years later, in particular during the Old Kingdom. Bronk Ramsey (2010, p 1556) refers to Shaw (2000) as the "consensus chronology" and we use it as the basis of our analysis of political outcomes. Our ruler tenure list starts in 2686 with Nebka (2686-2667 BCE), the first ruler of the 3rd dynasty during the Old Kingdom (2686-2160 BCE).

We also use data from Shaw (2000) on dynastic tenures, who in turn relies on Manetho. In Ancient Egypt, dynasties were normally defined by a sequence of leadership within a certain family. If no male heirs were available, children from the same ruling elite were sometimes adopted, as was later practiced also in Rome. Dynastic change were often periods of major political change when rulers embarked on new ambitious political agendas, involving for instance more intense pyramid construction or expansions into the Levant. The main time period of our study (2686-1150 BCE) covers dynasties 3-20 in the Egyptian chronology, all of which were native to Egypt. Later dynasties, such as the 22-23rd dynasties, were often of foreign descent.

Figure B.2 summarises the main periods in the history of Ancient Egypt. Upper and Lower Egypt were unified during the Early Dynastic Period around 3000 BCE. Our study period starts in 2686 BCE with The Old Kingdom, which ends in the chaos of IMP I around 2160 BCE. After the Middle Kingdom, IMP II starts around 1650 BCE, and the New Kingdom lasts until about 1100 BCE.

B.5. Area under Control of the State. Figure B.3 depicts the evolution of the area under control of the state. This variable is a proxy for state capacity since periods of a large Egyptian empire required substantial military resources to keep frontiers stable. Our main source of this data is Geacron.

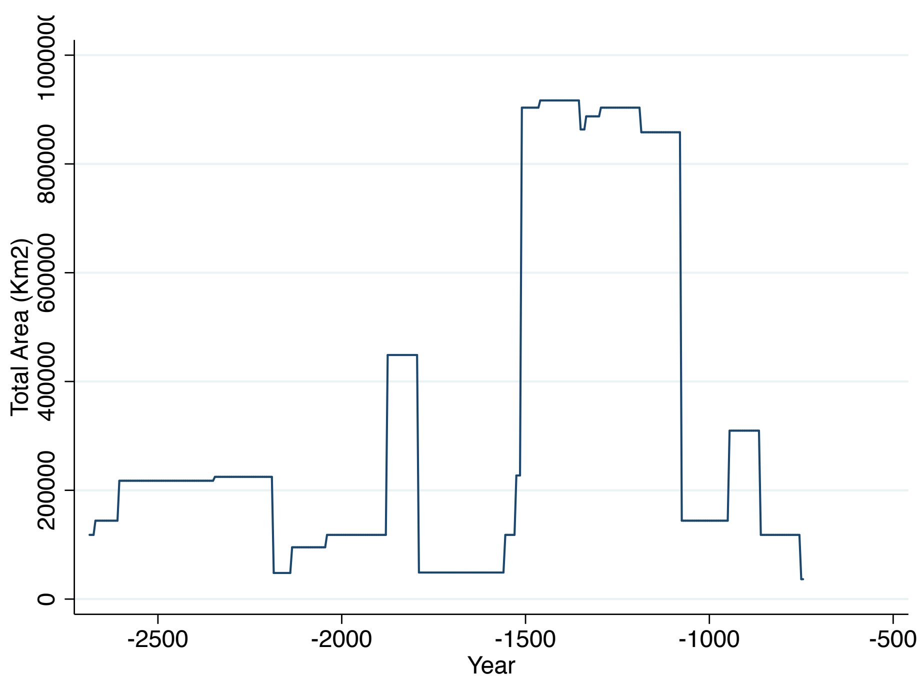


Figure B.3. Area (in Km^2) of Ancient Egypt. Notes: The figure shows the total area controlled by the Egyptian state during antiquity. Source: Geacron.

B.6. Pyramids. As mentioned in the main text, the largest type of public monuments were the pyramids. Figure B.4 shows the time distribution of pyramid construction during 2800-1600 BCE obtained from Lehner (1997). The time series uses information about 31 pyramids that we know were constructed by a particular ruler. 29 rulers built pyramids and one ruler (Sneferu) built three pyramids. We also know the volume (in k cubic meters) of each pyramid and the tenure period of each ruler. To illustrate the historical progression of pyramid construction, Figure B.4 presents the annualized rate of pyramid construction, calculated as the total volume of construction completed during a pharaoh's tenure divided by the number of years the pharaoh was in office.

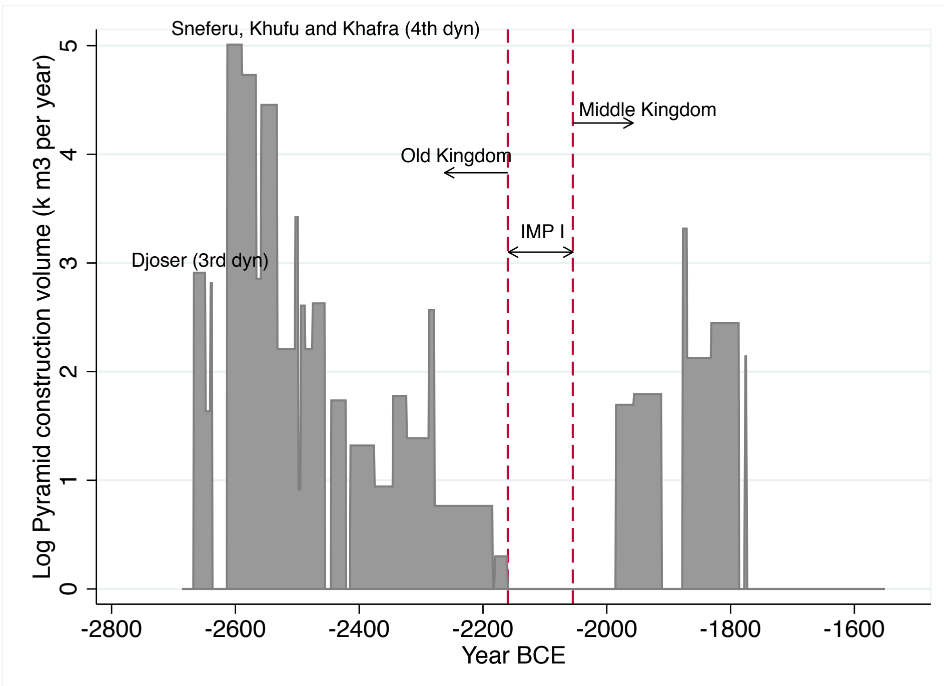


Figure B.4. VOLUME OF PYRAMID CONSTRUCTION. Notes: This graph shows the annualized rate of pyramid construction, calculated as the total volume of construction completed during a pharaoh's tenure in k cubic meters divided by the number of years the pharaoh was in office. Source: Constructed on the basis of data from Lehner, 1997.

It is believed that the pharaoh moved his whole court to the pyramid site, which thus in effect became the de facto capital of the state. The pyramids were primarily built as tombs and cult centers for the pharaohs who were worshipped as gods after they had passed away. As divine rulers, the pharaohs were believed to be able to influence the fortunes of life in Egypt, which in turn depended heavily on Nile floods (Bell, 1971). In this sense, the population probably accepted the toil and taxes as a sacrifice worth making as long as the pharaohs were associated with supernatural powers that guaranteed regular Nile floods and abundant harvests. However, apart from being a religious institution, specialists on Egypt believe that pyramid construction was a vital component in the broader project of forming and consolidating a unified Egyptian state in the Nile Valley (Lehner, 1997).

The construction of the pyramids required not only an abundance of manual labor but also massive amounts of construction material. The bulk of the construction material was limestone, often quarried nearby the building site. Recent findings of papyri rolls from the 26th year of the reign of Khufu document how an official named Merer led a crew of 200 men who transported limestone blocks, weighing 2-3 tons each, from a quarry across the river on a river barge up to the construction site of the Great Pyramid. The papyri sources indicate that the ancient Egyptians traveled on man-made canals from the Nile to the Giza plateau, unloading their stone blocks at a special port at the foot of the pyramid area (Sheisha, 2022; Smithsonian, 2015). The huge pyramid construction projects also required large amounts of copper, quarried in the eastern deserts, for the manufacturing of tools. It is further known that blocks of granite must have been transported on the Nile from the Aswan area in the far south (Mirsky, 1997). Such transportation was only possible during years when floods were large enough for the heavy barges to reach their destinations safely.

As discussed by for instance Wilkinson (2010), it appears that domestic pyramid construction was to some extent substituted during later history by the establishment of large temples and a permanent army capable of military excursions and even occupation of the Levantine hinterland. In the empirical section, we have argued that both pyramid construction and military conquest of territory are proxy variables for *state capacity* in a broader sense (Besley and Persson, 2009).

C. Variable definition, Summary Statistics and Preliminary Analysis

This section provides definitions of the variables employed in the empirical analysis (Section C.1) and a table of summary statistics (see Section C.2). Section C.3 describes a preliminary analysis that examines the time series properties of the variables employed in the empirical analysis.

C.1. Variable definitions. This section provides definitions for the variables employed in our empirical analysis.

- Pol instability$_t$; This variable takes on three values: 1 if there is a ruler change according to Shaw's chronology in the five year period $(t, t+1)$ [around 20% of the periods contain a ruler change];[48] 2 if any of the years in the five-year period $(t, t+1)$

[48]See Appendix B for a discussion about the chronologies and dating issues for rulers and dynasties. Thanks to the information provided by a Ramessid king-list written on a papyrus on the Museo Egizio in Turin (Turin cannon), there are few weak links in the order and dating of Old Kingdom rulers (exceptions are Menkaura 2532–2503 and Neferirkara 2475–2455). The end of the period is more obscure.

corresponds to periods of high political instability for which exact chronology is not available [around 17% of the periods];[49] and 0 otherwise.[50]

- Pol instability (dynasty)$_t$: This variable is defined in five-year periods and takes three values: 1 if there is a change of dynasty in the 5-year period $(t, t+1)$, 2 if two or more dynasties coexist in that period according to Shaw's chronology, as the fact that two or more dynasties coexist in the same period is a symptom of lack of centralization and political instability, and 0 otherwise. See Appendix B1 for definitions and main historical sources of Egyptian dynasties.[51]
- area$_t$: Area under State Control. This variable captures the log of the area (in km^2) under State control at the end of the five-year period starting in t. Source: Geacron.
- area$_t$(growth): Growth rate of the area under State control. This variable is the first difference of area$_t$. Source: Geacron.
- pyramids$_t$: Log of total pyramid construction volume (in cubic meters) over the lifetime of a pharaoh. Constructed on the basis of data from Lehner (1997).
- Nile floods$_t$. Defined in five- or ten-year periods, depending on the dependent variable employed. To compute this measure we first interpolate linearly the original Qunf's index; Second, we average the annual data in five- (or ten-) year periods. And third, we filter the resulting series using a Butterworth filter (see Pollock 2000, JoE, for details). Figure 5 plots the filtered data.
- rain hinterland: Defined in five- or ten-year periods, depending on the dependent variable employed. For symmetry, we have also filtered the data using a similar procedure as before. Since the level of resolution in Soreq's case is cruder (=less high frequency components), the original and the filtered series are very similar (correlation is .99). Figure 5 plots the filtered data. Since the resolution of this variable sharply deteriorates after 1140 BC, regressions containing this variable restrict the sample up to that year.
- nile extreme$_t$, 10%: Dummy variable equal to 1 if any of the years in the five-year period belongs to the 5 or to the 95 percentile of the distribution of the (filtered) annual Nile water level variable.

[49] More specifically, a 2 is assigned to i) the years of the last two dynasties of the Old Kingdom (the 7th and 8th dynasties, 2181–2160), in which around 17–20 ephemeral kings were in office (p.117 Shaw) but not exact chronology is available; ii) the first intermediate period (2160–2055), iii) the years of the 14th Dynasty (1773–1650), at the end of the Middle Kingdom, for which no exact chronology is available and many local rulers existed, and iv) the second intermediate period (1650–1550).

[50] Alternative codings of this variable were also considered and very similar empirical results were obtained. For instance, results are robust to considering a binary version of this variable, where periods of instability (including both pharaoh replacement as well as very instable periods with several rulers in power) were assigned a "1" .

[51] As opposed to rulers, Shaw provides dates on dynasty replacements even in intermediate periods. In addition, the dating in those unstable periods is likely to be less accurate.

- NILE EXTREME FLOOD$_t$, 5%: Dummy variable equal to 1 if any of the years in the five-year period belongs to the 95 percentile of the distribution of the (filtered) annual Nile water level variable.
- NILE EXTREME DROUGHT$_t$, 5%: Dummy variable equal to 1 if any of the years in the five-year period belongs to the 5 percentile of the distribution of the (filtered) annual Nile water level variable.
- TENURE$_t$. Number of years from ruler's first year in office up to the first year of the current 5 (or 10) year period.
- TENURE DYNASTY$_t$. Number of years from first ruler of the corresponding dynasty up to the first year (t) of the current 5 year period.
- PERIOD$_t$. Period in Ancient's Egypt chronology. There are six periods, as detailed in Appendix B. These periods are, in successive order: Old Kingdom, IMP I, Middle Kingdom, IMP II, New Kingdom, IMP III.

C.2. Summary statistics. Table C.1 presents summary statistics of the variables in our empirical analysis.

Table C.1. Summary Statistics

Variable	N	mean	sd	min	max
POL INSTABILITY	391	0.56	0.78	0.00	2.00
POL INSTABILITY (MET)	383	0.56	0.75	0.00	2.00
POL INSTABILITY (DYNASTY)	391	0.27	0.66	0.00	2.00
NILE FLOODS	391	4.27	1.25	1.89	7.21
RAINFALL HINTERLAND	391	2.91	1.25	0.70	6.44
NILE EXTREME FLOODS	391	0.05	0.21	0.00	1.00
NILE EXTREME DROUGHTS	391	0.07	0.25	0.00	1.00
AREA	390	327365.2	316089.9	36561	916795
TENURE	390	14.24	15.43	1.00	93.00
TENURE DYNASTY	390	76.89	57.17	1.00	255.00

C.3. Preliminary analysis. Since we are using time series data, a preliminary step is to investigate whether the variables employed in the empirical analysis are stationary. To that effect, we have applied unit root tests to each of the dependent variables employed in the empirical analysis. Table C.2 presents the results. The modified GLS Augmented Dickey-Fuller test (ADF-GLS, Elliot et al., 1996) was employed and lags were chosen using the Modified BIC. A constant and a trend were included in all cases.

	Unit root tests (ADF-GLS)
POL INSTABILITY	-4.455***
POL INSTABILITY (DUMMY)	-5.565***
POL INSTABILITY (DYNASTIES)	-3.290**
AREA	-1.869
AREA (GROWTH)	-13.431***
PYRAMIDS	-3.180 **

Table C.2. UNIT ROOT TESTS. This table presents the value of the ADF-GLS test for unit roots (Elliot et al, 1996) for each of the dependent variales employed in the empirical analysis. The null hypothesis is the existence of a unit root in the data. Large negative values indicate rejection of the null hypothesis. The number of lags was chosen using the Modified BIC. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

D. Validation Tests of Climate Indicators

In this section, we carry out a number of validation checks of our two main climate proxy variables: The *Nile flood*-index and the *Rainfall hinterland*-index.

D.1. Nile floods. As described in Section 4, our main proxy for Nile floods is a high resolution paleoclimatic archive (δ18O isotope-levels) from a stalagmite in Qunf Cave in Southern Oman (Fleitmann et al, 2003). We argue that it should provide a reliable indicator for the intensity of African Monsoon precipitation in the Ethiopian highlands and thus reflect the annual movements of the Intertropical Convergence Zone (ITCZ). Figure 2 shows the location of Qunf cave in relation to the movements of the ITCZ. Furthermore, we believe that there are several advantages of using this data including: i) The high time resolution with an average of 3.85 years during the relevant period; ii) Unlike data from historical Nilometers, the Qunf data have been uncontaminated by human manipulation (the data from the Roda Nilometer in Cairo was partly collected for tax collection purposes by different ruling elites during almost 1300 years); and iii) It is one of the most frequently cited works on the paleoclimate in the area from this period.[52]

However, there are also potential sources of measurement error in using the Qunf data as a proxy for historical Nile floods. In the section below, we perform a number of validation checks in order to ensure that our Nile floods-index has internal validity. Our main methodology for these checks is to use annual, interpolated and filtered observations of the same Nile floods-index as in the main analysis (see details about construction in text and below).We then correlate this index with other paleoclimatic proxies in order to analyze consistency. Alternative strategies, such as exact matching of observations for individual years with available actual observations or using filtered 5-year average observations, have also been tried and results are available upon request.

[52] As of February 2019, Fleitmann et al (2003) has been cited 1300 times according to Google Scholar and 882 times in Web of Science.

First, we argued above that the preferable climate archive for Blue Nile levels would have been located at Lake Tana in the Ethiopian Highlands. Marshall et al (2011) indeed present data on titanium (Ti) concentration in a sediment core from Lake Tana. The Ti concentration (in units of mg/g) is a proxy for rainfall because the mineral is common in local soils and a high concentration in the lake sediments means that heavy rainfall washed a lot of soil into the lake. The variable can be matched with available Qunf data from 8606 BCE to 1642 CE with a hiatus for the Nile floods data between 760 BCE-640 CE. Unfortunately, the time resolution for the Lake Tana Ti-measure is rather low - on average 19.76 years - which is the main reason why we only use it for a validation check rather than as the main proxy for Nile floods.

Validity tests for NILE FLOODS.				
	(1)	(2)	(3)	(4)
VARIABLES	Lake Tana	Delta sedim.	Medit. sedim.	Roda Nilometer
Nile floods (Qunf)	✓	✓	✓	✓
Pearson correlation ρ	0.628***	0.864***	0.676***	0.485***
Unit of measurement	Ti	mm	Ti/Ca	m
Observation period	-8608,1642	-6050,1642	-6686,1505	638,1642
Observations	8,864	6,306	6,805	1,005

Table D.3. VALIDITY TESTS FOR NILE FLOODS. NOTES: The table shows the Pearson correlation between our filtered *Nile floods* (Qunf) index and various other climate indicators. All variables are generated using a 100-year Butterworth filter of annual (interpolated and actual) observations. Unit of measurement for upper row variables and observation period are also shown where a negative number, for instance -8608, implies year 8608 BCE, whereas a positive number, for instance 1642, implies year 1642 CE. Characteristics and sources of variables are explained in the text. *** p<0.01, ** p<0.05, * p<0.1.

We would expect to find a positive relationship between our Nile flood (Qunf) observations and the Lake Tana Ti concentration. In Figure 3, Panel (a) shows the fitted binscatter relationship, indicating a positive association, as expected. The correlation coefficient (0.63) in Table D.3, column (1), suggests a strong positive relationship, implying that more precipitation at Qunf is indeed associated with greater rainfall also in the Ethiopian highlands.

A second method for assessing historical Nile floods exploits the thickness of sedimentation layers in the Nile delta and also out in the Mediterranean. Marriner et al (2012) collected sediment levels from several locations in the Nile Delta for successive centennial observations. In Figure 3, Panel (b), we show the relationship between interpolated, annual sediment levels and Nile flood levels from 6050 BCE - 1642 CE (6,306 observations). The figure shows a strong positive relationship, as expected, indicating that high monsoon rainfall at Qunf is associated with high floods and thick Nile Delta sediment layers. The coefficient in Table D.3 demonstrates that the correlation is very strong (0.86).

Third, also sediment levels measured off-shore out in the Mediterranean can be used to assess the historical quantities of the Nile discharge. Revel et al (2015) collected samples from sediment cores from the period 6686 BCE - 1505 CE with a time resolution during the relevant period of 7.87 years. The main proxy that we extract from their work is the isotope ratio

Ti/Ca, capturing a ratio of terrigenous (Ti from Blue Nile) to marine (Ca from Mediterranean) contents. As before, a higher relative Ti content of the sediment should originally derive from precipitation over the Ethiopian highlands that wash Ti-rich soils into the Blue Nile. Figure 3, Panel (c) shows the binscatter for 6,805 interpolated, annual observations and Table D.3 shows that the correlation coefficient is almost 0.68.

So far, we argue that we have demonstrated that our Nile floods (Qunf) data source has a strong relationship in the predicted direction with the closest alternative data sources on lake levels in the Ethiopian highlands, and with sediment levels in the Delta and in the Mediterranean. What we have not yet shown is the relationship between our Nile flood proxy and actual historical measurements of Nile levels.

As discussed above, the Roda Island Nilometer in Cairo has data for a much later time period than the one studied in the paper (622-1921 CE). In order to investigate the association between the Qunf data and the Roda Nilometer, we correlate the filtered, annual observations of the maximum Nile levels with our Nile floods-index. In total, we can use 1,005 yearly observations during 638-1642 CE. Figure 3, Panel (d), shows the binscatter relationship and Table D.3, column (4) the correlation coefficient. As expected, there is a clear positive association so that greater monsoon precipitation implies higher Nile floods. The correlation coefficient is 0.485. When the Nile flood-index increases from its lowest to its highest levels, the water level at the Roda Nilometer is predicted to increase by about 0.8 m.

D.2. Rainfall hinterland. In this subsection, we carry out a number of validation checks for our proxy variable for rainfall in the hinterland of Egypt. As discussed above, our main index is created from $\delta 18O$-levels from Soreq cave, 18 km west of Jerusalem. Figure D.5 shows a map of the region, which also includes the location of the Roda Island Nilometer (in Cairo).

A first and critical issue is how well our rainfall index, created from isotope observations from a stalagmite inside Soreq cave, captures actual observed rainfall outside the cave. Figure 4, Panel (a), shows the scatter plot of 11 actual, annual rainfall observations made during 2000-2010 by Bar-Matthews and Ayalot (2003), and our Rainfall hinterland index. The figure indicates a strong positive fit and the correlation coefficient in Table D.4, column (1), is 0.95, demonstrating that higher levels of our rainfall hinterland index are indeed associated with more observed precipitation.

A second question is how well rainfall in Israel also captures rainfall in Egypt's hinterland? Our circumscription model implies that "rainfall in hinterland" should describe conditions in an area away from the core but still close enough so that a return is possible when the weather improves. In Figure 4, Panel (b), we show the binned scatterplot of 1,344 monthly precipitation observations in Israel and Egypt over the time period 1901-2012 from World Bank (2019). Although the level of rainfall is different, with Egyptian precipitation rarely

Figure D.5. Soreq cave location

Validity tests for rainfall hinterland					
	(1)	(2)	(3)	(4)	(5)
VARIABLES	Rainfall Soreq	Rainfall Egypt	Veg. Soreq	Medit. dust	Nile floods (Qunf)
Rainfall hinterland (Soreq)	✓		✓	✓	✓
Rainfall Israel (mm)		✓			
Pearson correlation ρ	0.952***	0.820***	-0.627***	-0.240***	0.008
Unit of measurement	mm/year	mm/month	δ13C	Si/Al	index
Observation period	2000,2010	1900,2012	-5000,-1150	-6686,-1150	-2685,-750
Observations	11	1,344	3,849	5,536	388

Table D.4. Validity tests for rainfall hinterland. Note: The table shows the Pearson correlation between our annual, filtered *Rainfall hinterland* (Soreq) index and combinations of various other climate indicators. Correlated variables in columns (1)-(2) are actual, annual and monthly observations whereas correlated variables in columns (3)-(4) are generated using a 100-year Butterworth filter of annual (interpolated and actual) observations. Observations in column (5) are the same filtered, 5-year index observations as in our study. Unit of measurement for variables and observation period are also shown where a negative number implies year BCE, whereas a positive number implies year CE. Characteristics and sources of variables are explained in the text. *** p<0.01, ** p<0.05, * p<0.1.

exceeding 8 mm per month, the fitted regression line indicates a strong positive relationship between the two rainfall variables, and the correlation coefficient in Table D.4, column (2), is 0.82. The main take-away from the figure is that changes in rain levels is Israel also reflect changes in Mediterranean rains over Egypt.

A third validation issue concerns the correspondence between the δ18O isotope-levels at Soreq with other paleoclimatic proxies from the period. Bar-Matthews and Ayalot (2003, 2011) also obtained data from a second climate archive inside Soreq cave - the δ13C isotope - often considered to be a proxy for levels of vegetation in the vicinity of the cave (lower levels of δ13C indicate *more* vegetation). In Table D.4, column (3), and in Figure 4, Panel (c), we show that there is, as predicted, a negative and significant relationship for the time period 5000 BCE to 1150 BCE ($\rho = -0.63$).

Fourth, it is well known from ancient texts that a major threat to the Egyptian civilization during periods of drought was that the drying up of the hinterland led to the growth of massive sand dunes that sometimes invaded or completely swallowed cultivated fields near the Nile (Bell, 1971). The sediments layer collected by Revel et al (2015) out in the Mediterranean also includes data on the elemental ratio Si/Al, which derives from aeolian dust or sand, blown into the Nile from the nearby deserts. An important component of desert sand is quartz (SiO_2) and the Si/Al-ratio might thus serve as a proxy for the historical sandstorms and drought conditions. In Table D.4, column (4), and in Figure 4, Panel (d), we show that there is a moderately strong negative relationship between Rainfall hinterland and dust intensity in the sediment layers over the time period 6686-1150 BCE ($\rho = -0.243$). The binscatter in Figure 4, Panel (d), shows that the correlation is weak in lower levels of our Rainfall hinterland index, suggesting that in periods of very low rainfall levels, the desert climate becomes volatile and might either bring high or low dust storms.

Figure D.6 display the Qunf data together with data from the Roda Nilometer. To facilitate the comparison both indices have been normalized so that they move between [0,1]. To interpret the graph, recall that the resolution of the Qunf data for the period for which there is overlap with the Roda Nilometer data is much worse than that in our study period: while in the latter case there is one observation approximately every 3 years on average, for the period for which it overlaps with the Nilometer data, the average resolution is one observation every 18 years. In spite of this, Figure D.6 shows that both variables display quite similar patterns.

Lastly, we made it an explicit assumption in our theoretical model that shocks to the Mediterranean and Monsoon weather systems are uncorrelated. Table D.4, column (5) and Figure D.7 show the binscatter plot for the filtered 5-year average values of our Nile floods and Rainfall hinterland-variables for the time period of our study, 2685-750 BCE. As the figure and the correlation coefficient indicate, there is zero correlation between the two climate variables during the period considered in the empirical analysis. period.

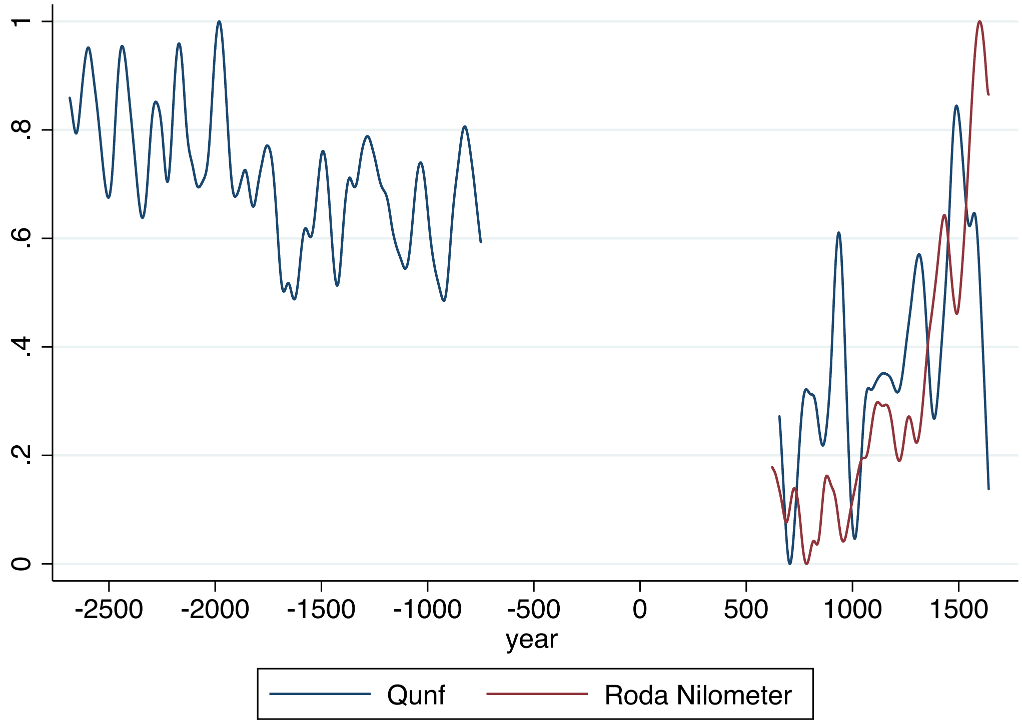


Figure D.6. NILE FLOODS (QUNF) VERSUS RODA NILOMETER DATA. NOTES: The figure shows the evolution of the Nile floods data obtained from the Qunf cave as well as direct measurements of the Nile river obtained from the Roda Nilometer. Both series have been filtered using a Butterworth index and to facilitate the comparison, the resulting variables have been normalized to be in the interval [0,1].

E. Additional references

Acemoglu, D. and J. Robinson (2006) *Economic Origins of Dicatorship and Democracy*. Cambridge University Press.

Burke, P.J and A. Leigh (2010) "Do Output Contractions Trigger Democratic Change?" *American Economic Journal: Macroeconomics.* 2(4), Oct, pp. 124-57.

Dal Bo, E., P. Hernández and S. Mazzuca (2016) "The Paradox of Civilization Pre-Institutional Sources of Security and Prosperity", *mimeo*.

Dube, O. and J.F. Vargas (2013) "Commodity Price Shocks and Civil Conflict: Evidence from Colombia", *The Review of Economic Studies* 80, 1384–1421.

Hornung, E. et al (2006)(eds) *Ancient Egyptian Chronology*. Leiden: Netherlands.

Marriner, N. et al (2012) "ITCZ and ENSO-like pacing of Nile delta hydro-geomorphology during the Holocene" Quaternary Science Reviews 45, 73-84.

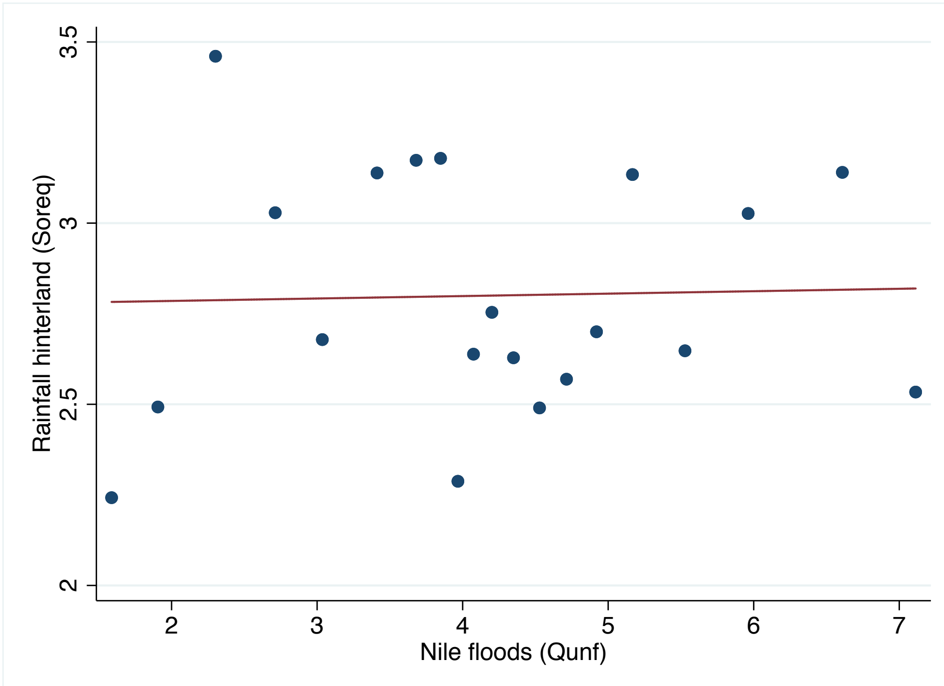


Figure D.7. Nile floods versus Rain hinterland. Notes: The figure shows the binned scatterplot relating our Nile floods (Qunf) index and rainfall hinterland (Soreq) index as specified in Table D.4.

Marshall, M. et al (2011) "Late Pleistocene and Holocene drought events at Lake Tana, the source of the Blue Nile" *Global and Planetary Change* 78, 147-161.

Mirsky, A. (1997) "Influence of Geologic Factors on Ancient Egyptian Civilization" *Journal of Geoscience Education* 45, 415-424.

Revel, M. et al (2015) "20,000 years of Nile River dynamics and environmental changes in the Nile catchment area as inferred from Nile upper continental slope sediments" *Quaternary Science Reviews* 130, 200-221.

Smithsonian (2015) "The World's Oldest Papyrus and What It Can Tell Us About the Great Pyramids" https://www.smithsonianmag.com/history/ancient-egypt-shipping-mining-farming-economy-pyramids-180956619.

World Bank (2019) *Climate Change Knowledge Portal*, https://climateknowledgeportal.worldbank.org/.