

Pharaoh's Cage: Environmental Circumscription and Appropriability in Early State Development ¹

LAURA MAYORAL

Instituto de Análisis Económico (CSIC), Barcelona GSE and University of Gothenburg

OLA OLSSON

University of Gothenburg

April 2020

Abstract

What explains the origins and survival of the first states around five thousand years ago? In this research, we focus on the role of productivity shocks for early state development in a single region: ancient Egypt. We introduce a model of extractive state consolidation that predicts that political instability should be low whenever environmental circumscription is high, i.e., whenever there is a large gap between the productivity of the area under state control (core) and that of the surrounding areas (hinterland). In these periods the elite can impose high levels of taxation that the population will be forced to accept as exit is not an attractive option. In order to test this hypothesis, we develop novel proxies for historical productivities on the basis of high-resolution paleoclimate archives. Our empirical analysis then investigates the relationship between proxies of the productivity of the Nile banks and of the Egyptian hinterland on the one hand, and political outcomes such as ruler and dynastic tenure durations and the intensity of pyramid construction on the other, during 2685 - 750 BCE. Our results show that while both too high or too low Nile floods are associated with a greater degree of political instability, periods with a greater rainfall in the hinterland (and hence a lower degree of environmental circumscription) are associated with an immediate rise in military and pyramid construction activity but also with a delayed increase in political instability, since the decline in effective circumscription provides the farming population with an outside option in the hinterland.

¹We are grateful for comments from Ingela Alger, Paolo Buonanno, Giacomo Benati, James Fenske, Bishnupriya Gupta, Debraj Ray, David Schönholzer, and seminar participants in Barcelona, Bilbao, Gothenburg, NES Moscow, Santiago de Chile, IAST in Toulouse, ASWEDE workshop in Uppsala, Oslo and Warwick. Mayoral gratefully acknowledges financial support from the Generalitat de Catalunya, and the Ministry of Economy and Competitiveness Grant number ECO2015-66883-P, and La Caixa Foundation. The research was supported by Vetenskapsrådet, grant number 2017-03171.

1. INTRODUCTION

In the later part of the fourth millennium BCE, the first two states in history arose in Mesopotamia and Egypt. During the subsequent millennia, pristine states would independently emerge also in the Indus Valley, China, Mesoamerica and Peru (Spencer, 2010; Borcan et al, 2018). The transition from small-scale societies to states brought about fundamental transformations such as urbanization, taxation, public architecture, writing, and bureaucratic structures that would drastically alter the conditions of human existence.

A very large literature has dealt with the question why these original states appeared where and when they did. A less often studied question concerns the economic factors influencing the subsequent early development and consolidation of these states, a development that arguably worked as models for numerous later states in history. All of the early states were characterized by highly intensified agricultural production systems, a (relatively) dense population, and large-scale public goods. But why did centralized political power sometimes prosper for centuries in some places but then stagnate, fragment and even collapse?

In this paper, we focus on the dynamics of early state development in response to exogenous climate shocks. We contribute to the literature on early state development by analysing systematic time series data from historical and archaeological records on political stability and on climate shocks from one of the longest-lasting early states in history: Ancient Egypt. In the spirit of Carneiro (1970), we propose that the more *environmentally circumscribed* a territory is, the greater the likelihood of the presence of a state in that area. A territory is environmentally circumscribed whenever there is a large productivity gap between its core and the surrounding periphery, in such a way that it can be difficult to sustain an alternative livelihood in the latter area. In contrast to extant papers, we do not study state formation but *state instability*.

We introduce a model of extractive state consolidation where individuals can choose between two activities; one that is easily taxable and another one that is difficult to tax. To match our empirical investigation, we interpret this second activity as the possibility of evasion through out-migrating from the territory controlled by the state to another territory free from state control.² A key novelty of our approach is that we consider the degree of effective circumscription (or,

²Our model is however open to a wider interpretation. For instance, the non-taxable activity could be growing tubers, as in Mayshar et al (2020), or any other activity that is difficult to tax, and the main implications of the model would still apply. As convincingly argued by Scott (2017) and Mayshar et al (2020), the set of non-taxable activities at the time when early states were formed was very large and it possibly contains all activities different from growing cereal crops.

more broadly, the productivity gap between the taxable and non-taxable activity) to be time-varying, rather than a fixed characteristic. In periods when the area under state control (the core) is substantially more productive than the surrounding territories (the hinterland), circumscription is strong, implying that populations in the area are pulled towards the state in the core. In those periods, the extracting elite can impose high levels of taxation that the population will be forced to accept as exit is not an attractive option. In times when the hinterland instead is relatively productive, an outside option is introduced to the peasant population who will either exit from the state or force a reduction in taxation. Either option will lead to a decrease in the tax collection which has the potential of destabilizing the state. The dynamic nature of our model highlights new aspects that the standard static version cannot predict: While the effect of a shock in productivity in the core is *direct* (it raises the level of tax collection right away), an improvement of the conditions in the hinterland has an *indirect* effect, characterized by people departing from the state (or posing a threat to do so), which is likely to be a delayed adjustment.

Our empirical investigation analyses how asymmetric historical weather shocks in both the core and the hinterland led to changing levels in effective circumscription that, in turn, affected political instability and state capacity in Ancient Egypt during a period of approximately 2 millenia.³ It is widely recognized that the success of the Egyptian state is due to the peculiar geography of the Nile valley (Allen, 1997), characterized by the sharp contrast between the productivity of the Nile banks and that of the surrounding areas. Our choice to focus on ancient Egypt is motivated by three reasons. First, the ancient Egyptian state is characterized by an exceptional level of continuity, which allows us to study its evolution over a long period of time.⁴ Second, the Egyptian civilization has attracted a huge amount of attention from archaeologists and historians and, as a result, historical records (often validated by radio-carbon studies, see below) are much more precise than those pertaining to other civilizations (Shaw, 2000). This allows us to undertake quantitative analysis with a reasonable degree of confidence on our data.

A third reason for choosing to study Ancient Egypt is that changes in the degree of circumscription in Egypt are driven by two largely independent weather systems: the Mediterranean “winter” rainfall and the African monsoon “summer” rainfall in the Ethiopian highlands. The Mediterranean precipitation system provides winter rains to the Nile delta, to the northern parts of the Sinai desert, and to the neighboring lands in Southern Levant. The amount of rainfall

³More specifically, we analyze the period 2686-740 BCE which covers the beginning of the classical period of Ancient Egypt, from the Old Kingdom to the end of the New Kingdom.

⁴Unlike most other territories where early states were formed, the classical period in ancient Egypt arguably offers an object of study that is as close to an isolated and undisturbed ‘historical lab’ as one can find. It should be recognized that the foreign Hyksos, who ruled Egypt around 1600 BCE, was an important exception to the general isolation of Egypt from outside intervention.

in the hinterland surrounding the Nile valley determines to a large extent the type of activities that can be undertaken and the amount of population that can be sustained there. On the other hand, precipitation in the Ethiopian highlands was the main source of the summer floods of the Nile, which in turn determined the productivity of the Nile banks. The extent of these monsoon rains is determined by the latitudinal positioning of the Intertropical Convergence Zone (ITCZ), a weather system in the Indian Ocean. The further north that the ITCZ reaches, the greater the amount of summer rainfall in the Ethiopian highlands and the greater the seasonal floods of the Nile.

It is widely documented that the extent of Nile inundations and the amount of rainfall in the areas surrounding the Nile valley have experienced large variations over time (Said, 1993). This, together with the fact that the above-mentioned two weather systems are basically uncorrelated (see Figure C.9), makes the Egyptian territory a perfect case for testing the appropriability/circumscription hypothesis. As proxy variables for historical Nile inundations and hinterland rainfall, we introduce two paleoclimatological data sources (isotopic composition of cave stalagmites) that have not previously featured in econometric analyses of ancient Egypt and which we validate extensively using other standard sources. Thus, our identification strategy exploits the impact of two exogenous and time-varying orthogonal shocks on the degree of effective circumscription in Egypt, which will allow us to make causal claims about the relationship between state consolidation and the degree of circumscription.

In order to measure the degree of *political stability* over time, we use historical records on the continuity of the Egyptian state and tenure durations of individual rulers and dynasties. Our *state capacity* variables capture the temporal variation in the intensity of massive public monument construction (pyramids) and the fluctuating extent of Egyptian territorial expansion in its hinterland. Our results show that political stability increased in times of high circumscription, that is, in periods of intermediate Nile floods and dry conditions in the hinterland. The size of both effects is large and their magnitudes are comparable, which supports the hypothesis that the ability to expropriate output not only depends on producing a large output but also on *the ability of retaining the population and force them to accept a high tax*.

We argue that our paper contributes to the existing literature by focusing on the relatively unexplored question of *state consolidation and survival*. More specifically, our contribution is three-fold. First, we introduce a simple model of extractive state consolidation which incorporates the novelty that the relative productivities of the taxable and non-taxable activities can change over time. Second, we test the model predictions by exploiting two exogenous time-varying proxies

for the productivities associated to the above-mentioned activities and show that political stability and state capacity are positively related to increases in the productivity of the taxable activity and decreases to the productivity of the non-taxable activity. Finally, we introduce new data on climate shocks, political instability and state capacity for the Egyptian state covering almost 2 millenia.

Our work is related to several important strands of research in the literature. Our paper is most strongly related to Scott (2017), Schönholzer (2019) and Mayshar et al (2017, 2020) who emphasize the crucial importance of the transparency and appropriability of the main taxable resource for a strong state to arise. Since our results are obtained using a very different empirical strategy from those papers, our study gives additional support to the appropriability/circumscription hypothesis. Our theoretical model shares the assumption of Lagerlöf (2020) that a ruler in an early state can choose between spending tax revenue on own consumption and public goods like defense.⁵ Our work is broadly related to the literature on *state and fiscal capacity*, pioneered by Besley and Persson (2009) and reviewed by Johnson and Koyama (2017). The key reference on environmental circumscription is Carneiro (1970). Our dynamic version of the model has some similarities to Olson's (1993) famous account of 'roving vs stationary bandits', but our interpretation of circumscription is closer to that suggested by Allen (1997), who emphasises the role of circumscription as a cage in a very underpopulated world.⁶

A second important tradition is research on the historical impact of climate shocks on human societies.⁷ In recent years, there has further been an explosion of work in the science literature on *paleoclimatology* that has increased our understanding of how historical weather conditions are correlated with serious political instability and state collapse.⁸ However, researchers in this

⁵In Lagerlöf's framework (2020), there is an additional choice to invest in productive capacity like granaries for storing taxed grains.

⁶Mann (1986) famously referred to Ancient Egypt as a 'social cage'.

⁷Diamond (2005) reviews several historical cases of societal collapse, including the civilizations of the Maya. Harper (2017) offers convincing new evidence that detrimental climate shocks and accompanying rounds of infectious disease significantly contributed to the relatively rapid and unexpected decline of the Roman world. In economics, well-known contributions include Miguel et al (2004) on rainfall and conflict in Africa, Hidalgo et al (2010) on how negative rainfall shocks led to peasant invasions of farm land in Brazil, Dell et al (2012) on temperature shocks and economic growth, and Anderson et al (2017) on the relationship between temperature and Jewish persecutions in Europe. In a number of articles, Solomon Hsiang and coauthors (Hsiang et al, 2013; Hsiang, 2016; Carleton and Hsiang, 2016) have pioneered the econometrics of climate change and the study of its social effects.

⁸Such frequently discussed crisis events include the "4.2 ky event" (2250 BCE), featuring the collapse of the Akkadian empire (Cullen et al, 2010), population decline in the Harappan civilization (Giosan et al, 2012), the systemic Bronze Age collapse in the Mediterranean around 1200 BCE (Kaniewski et al, 2015), or the disintegration of the Maya civilization around 900 CE (Hodell et al, 1995; Kennett et al, 2012).

tradition rarely employ standard time series econometrics, as in our study, and neither do they study the interplay between different weather systems in a core and a periphery.

A third central strand of research is the large literature on historical Egypt.⁹ Allen (1997) studies the nature of agricultural production for the rise and consolidation of the Egyptian civilization, as discussed above. Two other highly related contributions on historical Egypt are Chaney (2013) and Manning et al (2017).¹⁰ Although our approach shares important similarities to Chaney (2013) in the sense that one of main explanatory variables is a proxy for annual Nile floods and his outcome variables also measure political stability, our study differs from the existing literature in important ways. Chaney's (2013) work concerns the Mamluk period in the 12-14th centuries CE whereas Manning et al (2017) deals with the Ptolemaic Era (332-30 BCE). We contribute by analyzing state development in a much longer and earlier period in Egypt's history when Egypt was relatively unaffected by neighboring powers. Second, unlike any previous analysis of historical Egypt that we are aware of, in our econometric analysis we rely on paleoclimatic data from natural archives that were unaffected by humans, whereas both Chaney (2013) and Manning rely on historical, man-made records of Nile floods, in part collected for tax reasons. Third, as far as we know, our paper offers the first test of a dynamic environmental circumscription model that considers the interplay between the intertemporal variation in hydrological conditions prevailing in a core and a hinterland.¹¹

The paper is structured as follows. In Section 2, we review the related literature on early state development and historical Egypt. In Section 3, we lay out our dynamic model of environmental circumscription. Section 4 gives a historical background, and Section 5 describes the main variables employed in the empirical analysis. Section 6 describes the empirical strategy and the empirical specifications while Section 7 presents our empirical results. Section 8 concludes the paper. Additional materials are presented in Appendices A to D.

2. RELATED LITERATURE

In this section 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

⁹Influential general overviews include Butzer (1976), Shaw (2000) and Wilkinson (2010).

¹⁰See Appendix for a more extensive account of these works and how they relate to our research design.

¹¹Schönholzer (2019) has recently presented a grid cell level cross-sectional empirical analysis of the importance of environmental circumscription for early state emergence across the world.

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. 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.¹² 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.

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 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 (2020) 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). Creating a panel of 44 major Mesopotamian polities between 3050 and 1750 BCE, Benati et al (2019) find that more inclusive political institutions tend to arise in periods when harvests were deficient and output less transparent.

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

¹²See Bentzen et al (2016) for an empirical test of this hypothesis.

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.

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).¹³ 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.¹⁴ 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

¹³Mann (1986) even referred to Egypt as a 'social cage' where the desert inhibited the peasants from out-migration.

¹⁴Our 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.

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 theoretical section below, we will present a formal model of such an intertemporal interaction between exogenous environmental circumscription and political stability.

3. A MODEL OF EARLY STATE CONSOLIDATION

This section presents a simple 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.

Our model is inspired by the work in Carneiro (1970) and Allen (1997) and related to the model introduced in Schönholzer (2019). These papers analyze how the degree of circumscription affects the probability of state *formation*. Instead, we focus on the impact of *changing levels* of circumscription on the stability and the capacity of the state and, hence, our emphasis is placed on *state consolidation*. From an empirical perspective, our approach will allow us to study the role of changes in circumscription over time, rather than in a cross-section - as in most of the extant literature in this area - which will be helpful in order to (at least partly) reduce concerns about omitted variable bias.

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).¹⁵ 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

¹⁵Thus, in our analysis we take the “core” and the “hinterland” territories as given, being exogenously determined by geographic and climatic conditions.

can be interpreted as growing tubers, as in Mayshar et al. (2020), 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 .

3.1. Individuals. There is a population of size 1 that at time t 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.¹⁶ 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 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 $A_t^j = \rho^j A_{t-1}^j + r_t^j$, with $\rho^j \in (0, 1)$, $j = c, h$.¹⁷ 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 \bar{r}^c, \bar{r}^h , respectively.

As an illustration consider the case of Ancient Egypt, where the productivity of the Nile valley depends on the extension of the yearly floods, so that A_t^c is high when floods are favorable (that is, not too low and not too high). The valley is surrounded by an area that displays much lower productivity, that is also subject to fluctuations due to weather shocks that make it more or less inhabitable. When rainfall in the area surrounding the valley is (relatively) high, the living conditions become less hostile and activities such as cattle raising, hunting and gathering, etc., can be sustained. Thus, productivity in the hinterland increases whenever climatic conditions are favorable and vice versa. As illustrated in Section 5 and Appendix B.1, climate patterns in these areas have had a large variability over time, which surely translated into important changes in the productivity of the Nile Valley and its surrounding areas and, hence, in the circumscription of the Egyptian state.

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 $F(\cdot)$.

¹⁶Typically, the main activity in the core is cereal cultivation. As convincingly argued by Scott (2017) and Mayshar et al (2020), 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.

¹⁷Under these assumptions, A_j is a stationary, mean-reverting process.

All the individuals in c are subject to a tax rate τ on their produced output which is set by the state on a yearly basis (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 A_t^h at the beginning of t , 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 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 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,

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

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

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

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

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

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

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

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

3.2. 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).¹⁸ The tax revenue at t is given by

$$(3) \quad TR_t = \tau E_t^c(\alpha) A_t^c L_t^c,$$

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 at the core *at time* t , A_t^c , 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* $t - 1$ or a shock that decreases the productivity of c in the same period will have a (negative) *indirect* effect on TR_t through its effect on the population at the core at t , L_t^c . This is due to the fact that, *ceteris paribus*, an increase in A_{t-1}^h (or a decrease in A_{t-1}^c) will reduce $\bar{\psi}_{t-1}$ and, therefore, will have a negative impact on the population at the core at t .

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 $NR_t = TR_t - \kappa$.¹⁹ 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., $NR_t > 0$. Mayshar et al. (2020) 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.

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.

¹⁸Alternatively, 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.

¹⁹More 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.

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.²⁰ 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.²¹ 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 , $i = fc, hg$. 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 $(1 - \eta(a_t))$, the incumbent ruler is defeated and state's consumption P_t and utility is zero.²² 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 maximised, where the latter is given by²³

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

²⁰See 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.

²¹The 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.

²²This 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.

²³For 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.²⁴ In these cases, the optimal size of the army, a_t , is

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

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.²⁵ An alternative scenario would be one in which the aggressor is primarily interested in conquering TR_t , in which case D_t and TR_t will be positively related.

A simple way to model this would be 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. For the sake of brevity, the following propositions focus on the case where λ is assumed to be equal to zero. The more general case where λ is allowed to be different from zero is discussed in Appendix A, where it's shown that similar conclusions are achieved.

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^c or an increase in A_t^h , will increase political instability at t .*

²⁴Provided $TR_t - \kappa \leq 0$, a corner solution ($a_t^* = 0$) will be obtained if d_t is zero. In this case political instability is trivially 0.

²⁵In the case of Egypt, for instance, the pharaoh was believed to be responsible for the Nile floods. Then, instances of famines derived from extreme Nile floods he was often directly blamed for them (Wilkinson, 2000).

Proposition 1 is the main result of the paper and implies that *ceteris paribus* a decrease in state revenue will lead to an increase in the probability that the ruler is defeated. See Appendix A for the proof of this proposition. Additionally, it's possible to establish some predictions about the evolution of P_t as a function of the productivity shocks.

Proposition 2. *Under the assumptions above, a weather shock that increases TR_t will increase $E(P_t)$. More specifically, *ceteris paribus*, an increase in A_t^c , A_{t-1}^c , or a reduction in A_{t-1}^h will increase $E(P_t)$.*

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 . See Appendix A for the proof of this proposition.

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.

4. BACKGROUND ON ANCIENT EGYPT

This section provides a brief overview of some of the key features of the history, economy, and political organization in the classical period of Ancient Egypt, which is the period that we study in the empirical section. See Appendix B for a more extensive account.

4.1. Historical overview. The practice of agriculture was widespread in Egypt around 4000 BCE. Shortly afterwards, around 3600 BCE, the Upper Egypt proto-state emerged. Only in Egypt the formation of the state followed so quickly the adoption of farming. It seems that the process of state formation was facilitated by climate change: the Neolithic wet phase, a period unusually wet in northern Africa was coming to an end, which implied migrations from the areas surrounding the Nile valley to the valley itself. Figure B.1 in Appendix B.1 shows the aridification process of the areas around the Nile valley from 6000 BCE to 3000 BCE approximately, as reflected by our proxy for rainfall described in Section 5. It shows that the creation of the Egyptian state coincides with the driest spell in the last 8000 years.

It is generally believed that Egypt was unified into a full state by King Menes around the year 3000 BCE. Figure B.2 in Appendix B shows a basic chronology of the main historical periods, relying primarily on dates provided by Shaw (2000). The Early Dynastic Period (3000-2686 BCE), with an undetermined chronology of rulers, was superseded by the Old Kingdom-era (2686-2160 BCE), which hosted dynasties 3-8 and a more precise dating of ruler tenure periods. Due to this improved data availability, our study period begins with the onset of the Old Kingdom in 2686 BCE. Construction activity reached its peak during the 4th dynasty (2613-2494 BCE) when enormous pyramids were erected at Giza to celebrate the reigns of different pharaohs.

After the reign of the 6th dynasty, a period of decline set in, culminating in the First Intermediate Period (henceforth IMP I) in 2160-2055 BCE. During intermediate periods, centralized political power broke down and several competing and short-lasting regional or external rulers fought for control of portions of the Nile Valley. Egypt was however once again unified during the Middle Kingdom (2055-1650 BCE) during which there was a resurgence in monumental building activity, military expeditions abroad and mining in the Sinai. Also the Middle Kingdom was punctuated by a period of decline in the Second Intermediate Period (IMP II) in 1650-1550 BCE, characterized by a divided Egypt with the people known as the Hyksos holding power in the north, Egyptian rule at Thebes in the center of the country, and Nubians ruling in the south. A final high point in the classical period of Ancient Egypt was the New Kingdom (1550-1069 BCE) when Egypt was a local superpower and extended its direct influence far into the Near East. Also this era ended with decline and foreign invasions at the onset of the last millennium BCE in the Third Intermediate Period (IMP III, 1069-664 BCE). Our extended study period ends in 750 BCE.

4.2. Weather shocks and agricultural production. The cornerstone of the ancient Egyptian civilization was its highly efficient system of artificial irrigation agriculture. The system of irrigation required a coordinated labor force and available evidence suggests that this coordination was primarily organized on a local rather than on a centralized level (Butzer, 1976; Said, 1993). The rhythm of this kind of agriculture followed a distinct and highly predictable annual pattern. The Nile floods typically started to appear in July, reaching a peak in September, and then receded in October. The inundated basins became fertile fields where farmers first plowed the soil and then sowed crops like barley, wheat and flax. The crops were harvested in February and were often stored in communal granaries. Although the artificial irrigation agriculture in ancient Egypt was highly successful during most years, it was also highly sensitive to disruption. Abnormally high floods might seriously damage the infrastructure of dikes and canals, whereas very low floods might make fields too dry for cultivation.

The territory of Ancient Egypt was affected by two different precipitation regimes; Mediterranean “winter” rainfall and African monsoon “summer” rainfall in the Ethiopian highlands. The Mediterranean precipitation system provided winter rains to the Nile delta, to the northern parts of the Sinai desert, and to the neighboring lands in the Southern Levant. These rains on their own were not sufficient to allow for sedentary agriculture in northern Egypt but implied that both the Levant and the Sinai were more fertile with a higher population density. In times of relatively abundant rains, an ecological land bridge between Egypt and the Levant emerged that allowed migration between the two areas and economic activity in the Sinai hinterland.

Precipitation in the Ethiopian highlands is the main source of the summer floods of the Nile. During most of the year, the Nile in Egypt (Main Nile) gets most of its water from the White Nile, draining large areas of Equatorial Africa. Right before the summer floods start in June, the White Nile contributes with about 75 percent of the Main Nile waters. Then during the summer floods in July-Oct, almost 100 percent of the Main Nile’s water originates from the Blue Nile (Hassan, 1981). These floods are, in turn, derived from the monsoon precipitation that fall on the Blue Nile’s drainage area in the Ethiopian highlands. The extent of these monsoon rains are determined by the latitudinal positioning of the Intertropical Convergence Zone (ITCZ), a weather system in the Indian Ocean. The further north that the ITCZ reaches, the greater the amount of summer rainfall in the Ethiopian highlands and the greater the seasonal floods of the Nile (see also the presentation of our climate proxies below).

There are many historical accounts of how aberrant Nile floods led to famines, social unrest, and political instability (Chaney, 2013; Manning et al, 2017). It is generally believed that deficient Nile floods contributed significantly to the social disorder during IMP I-II (Shaw, 2000). See Appendix B for a more detailed account of existing evidence.

4.3. State organization. As was discussed above, the surpluses from irrigation agriculture were sufficient to feed a non-producing elite and their servants in Ancient Egypt. The harvested cereals might in turn be stored, saved or redistributed by powerful agents. The emergence of distinct social hierarchies in pre-dynastic times and the subsequent rise of a state, was most likely the indirect consequence of a forced redistribution of surplus food resources from the farming population to a ruling elite. As argued by Scott (2017) and Mayshar et al (2020), the Egyptian environment with fairly predictable and easily observable Nile floods, combined with the superior storability, transferability, and appropriability of cereals, explains to a large extent why Egypt and Mesopotamia could develop highly stratified societies. There are many indications that a system of taxation was present in Egypt soon after unification.

The ruling elite had two main instruments at their disposal in their pursuit of taxes; the measurement of Nile floods and a writing system to keep records. Nile floods were carefully monitored and recorded from early times with the help of “nilometers” around the country that measured the level of the floods in July-Oct. With this tool, tax collectors could easily approximate the level of available harvests and meter out an appropriate level of taxation from the farming villages along the Nile (Said, 1997; Mayshar et al, 2017).

The accumulation of tax resources in granaries and temples controlled by the ruler, also served as a natural source of attraction to bandits and nomadic peoples in the hinterland that were not integrated in Egyptian society. Such considerations necessitated the permanent presence of a military organization taking orders from the pharaoh. It is also known that the ruler had access to corvée labor for the construction of the many massive public monuments that were erected to honor the gods and glorify the reigns of pharaohs. Most likely, this labor was ordered to work during the flooding season in July-October when no work was possible on the fields. It is not clear exactly how the rulers managed to extract these labor services from the population, but most of the evidence seems to suggest that the laborers were conscripted regionally in labor gangs on a rotating basis (Lehner, 1997).

4.4. Pyramids. The largest type of public monuments were the pyramids. Lehner (1997) offers an exhaustive account of all known ancient pyramids in Egypt. The first major pyramid was constructed at Saqqara around 2650 BCE by Djoser during the 2nd dynasty. Pyramid construction reached its historical peak during a few decades in the 2500s BCE with Khufu’s Great Pyramid at Giza, erected around 2570 BCE. Khufu’s limestone pyramid had a height of 146 meters, a volume 2.6 million cubic meters, and its construction demanded a work force of roughly 20,000-30,000 men at a time during at least twenty years. After the 4th dynasty, pyramids were not nearly as large. The last really large pyramid was built around 1810 BCE (Lehner, 1997). After the IMP II, pyramid construction ceased. Figure 1 shows the time distribution of pyramid construction during 2800-1600 BCE.

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

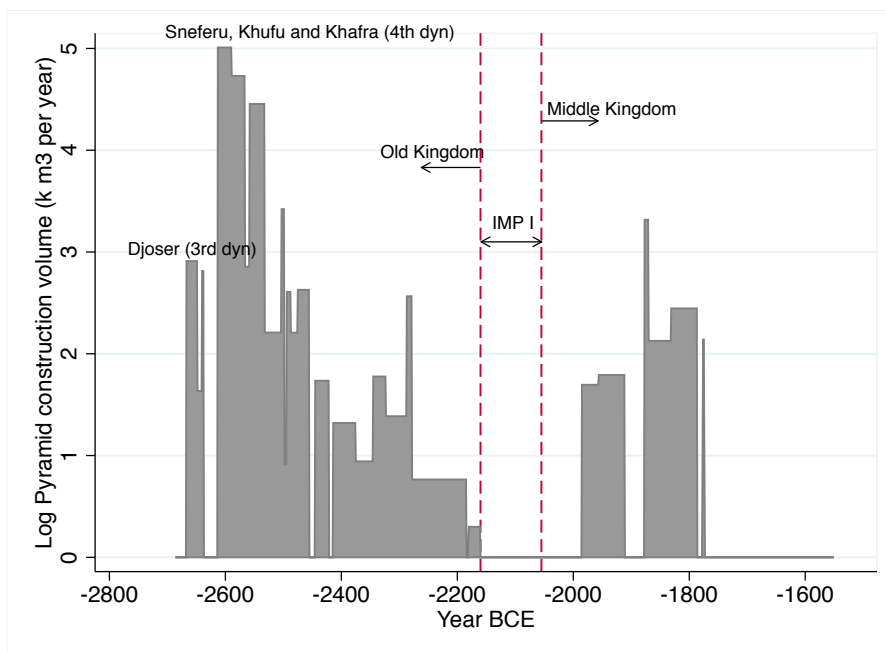


Figure 1. VOLUME OF PYRAMID CONSTRUCTION

Notes: This graph shows the average annual volume of pyramid construction in k cubic meters per year. Source: Constructed on the basis of data from Lehner, 1997.

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).

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 a permanent army capable of military excursions and even occupation of the Levantine hinterland. In the empirical section, we will argue that both pyramid construction and military conquest of territory are proxy variables for *state capacity* in a broader sense (Besley and Persson, 2009).

5. DATA

In this section we describe the main dependent and independent variables employed in our empirical exercise: the political stability and fiscal capacity outcome variables on the one hand, and the Nile flood data and the local rainfall variables, which provide the source of exogenous variation in our empirical analysis, on the other. A description of other variables employed in

the analysis is provided in Section 6. More detailed definitions of all variables, some preliminary analysis as well as a table of summary statistics are provided in Appendix D.

5.1. Political instability. Our main dependent variable aims to capture political instability over the period 2686 BCE – 750 BCE, which covers the Old, Middle, and New Kingdoms, as well as the first and second intermediate periods (see Figure B.2 in the Appendix for a summary of the Egyptian chronology). Consistent with the theory presented in Section 3 we consider that a period is politically stable if there exists a single king ruling over a centralised state. In line with this interpretation, a period is considered to be politically unstable whenever there are two or more rulers in a period, either due to ruler replacement or because there is no central rule due to political fragmentation and several kings are ruling simultaneously. Periods of absence of centralised rule were relatively infrequent in Ancient Egypt and, as documented by Shaw (2000), they are characterized by many short-lasting local kings co-existing simultaneously over the Nile (see Section 4.1 for a more detailed description).²⁶

We code our main dependent variable, POL INSTABILITY, as follows. First, we divide the sample period (2685-750 BCE) in 389 5-year intervals. Second, we assign a value equal to 1 if there is at least one ruler replacement over the 5-year interval (which happens in 20.6 of the periods) and a value equal to 2 if it's a period of no central rule (18% of the observations). The remaining periods (61.4%) are considered as politically “stable” and are coded as zero. The choice of these values aims to reflect the fact that the breakdown of the centralised power is a more severe instance of political instability than the replacement of a ruler. For the sake of robustness, several variations of this basic definition have also been considered, see Table 2.

To code ruler replacements, we use the ruler tenure lists from Shaw (2000), which is generally considered as providing the most accurate chronology for the tenure regimes of Egyptian kings and pharaohs. By carbon-dating museum samples of artefacts associated with particular rulers, Bronk Ramsey et al (2010) constructed a chronology of ruler periods during 2650-1100 BCE, i.e. during the classical period we study in the paper. When alternative ruler tenure lists were compared, Bronk Ramsey et al (2010) found that their radiocarbon dates were in strong agreement with Shaw (2000) and referred to it as the ‘consensus chronology’.²⁷

The variable POL INSTABILITY most likely will be able to capture instances of high political instability leading to pharaoh replacement and/or to the breakdown of centralized rule. In spite of this, it also presents important limitations regarding both its measurement and its interpretation.

²⁶Due to their weakness and limited power these rulers have typically left very scarce historical records.

²⁷See Appendix B.3 for details about the dating of rulers and dynasties in Ancient Egypt.

Firstly, although Shaw's tenure list is widely recognized as the most accurate chronology, there are still a number of dating issues (see Appendix B.3 for a summary) that imply that the exact annual dates provided by Shaw need to be taken with caution. For this reason and as a way of mitigating measurement error, we have decided to work with 5-year periods rather than with annual observations. For the sake of robustness alternative period lengths have been considered obtaining similar results (see Table 2).

Secondly, it is clear that ruler changes is a crude proxy for political instability. It surely happened that an appointed ruler during a period of relative stability might pass away after a short tenure due to natural, random health reasons or because the person was elected at an old age. Such a short duration might then spuriously be coded as reflecting political instability in the country. To address this limitation we consider an alternative indicator of political instability which only codes the most severe cases of ruler replacement: those that also involved the change of the whole dynasty to which the ruler belonged to. The replacement of a dynasty in Egypt was frequently the outcome of political turmoil and was typically associated with major political change, often reflected in new agendas regarding foreign policy or in the construction of domestic public monuments (Wilkinson, 2010). Appendices B.3 and D provide additional details about the definition of dynasties and dating issues. We define a new variable, (POL INSTABILITY (DYNASTIES)), in a similar way as before: we divide the our sample period in 5 year intervals and assign a value equal to 1 if there is a dynasty change in any of the years of the period (around 4% of the observations in the sample); a value of 2 in periods with no centralized power where at least 2 or more dynasties coexisted along the Nile river (12% of the observations), and a value of 0, otherwise (84% of the observations).

5.2. State Capacity. As proxies for government/ruler consumption or state capacity, we use two indicators; pyramid construction activity and the geographical extent of the Egyptian state. Starting with the pyramids data, we recognize that the presence of massive monuments are commonly used as a proxy for state capacity in archeology and anthropology. Pyramid construction required massive amounts of capital, labor and social coordination. Only a state with a strong capacity could construct pyramids like those in Egypt. However, pyramid construction reflects both state capacity *and* the burial preferences of the times. Since during the New Kingdom pyramid construction was no longer the preferred burial practice for rulers we exclude this period from the analysis.

We use data from Lehner (1997) on the volume of pyramid construction as well as on the pharaoh that built it. The exact timing of construction over the lifetime of the pharaoh is unknown. Therefore, to compute our proxy for pyramid construction, PYRAMIDS, we calculate average

construction per year by dividing the volume constructed over the lifetime of a pharaoh by the years he was in office. We then average over decades and compute the log of the resulting quantity.

An obvious drawback of these data is their lack of time resolution that will surely lead to attenuation bias. Nevertheless, since the average tenure of a pharaoh in our sample is 1.5 decades, it is reasonable to expect that our climate indicators will still be useful in predicting pyramid construction, despite the attenuation bias.

Figure 1 shows the time distribution of pyramid construction activity. Activity peaked already around 2600 BCE during the 4th dynasty. During IMP I, all large-scale monument construction ceased, until construction activity increased again during the Middle Kingdom-era, reaching a peak during the reign of Amenemhat III during the 12th dynasty. During the chaos of IMP II (1650-1550), many pyramid tombs and other burial sites were looted, a practice which accelerated after the fall of the 12th dynasty (Wilkinson, 2010). These experiences contributed to the adoption of much less conspicuous burial practices and pyramid construction ceased permanently already around 1700 BCE.

As a second indicator of government capacity, we use a variable measuring the geographical extent of the Egyptian state. As discussed by for instance Wilkinson (2010), it appears that domestic pyramid and temple construction to some extent was substituted during later history by the establishment of a permanent army capable of military excursions and even occupation of the Levantine hinterland. We use data from Geacron (2017) on the geographical extent of the Egyptian state as a proxy for such military capability (see Figure B.3). We thus consider empire-building as an alternative type of government consumption P_t to building pyramids.²⁸

5.3. Productivities and Weather shocks. To proxy our main independent variables, i.e., the productivities in the core and in the hinterland, A_t^c and A_t^h , we use weather shocks to those productivities.

5.3.1. Productivity in the core. Productivity in the core, A_t^c , is a non-linear function of the extent of Nile floods, as discussed in Section 4.2. What would be an ideal direct measure of Nile inundations? Egyptians have recorded Nile floods over hundreds of years with the help of Nilometers (such as the Roda Island Nilometer, used by Chaney, 2013, and Manning et al, 2017). Unfortunately, reliable Nilometer records for ancient Egypt do not exist.

²⁸It might be argued that empire-building also extended the area of the hinterland. While this is true, we refrain from considering the implications of such an assumption.

Instead, we use an indirect proxy of Nile inundation levels. The source of the Blue Nile (and of the smaller Atbara River) - the “pacemaker” of the late summer Nile floods in Upper and Lower Egypt - lies close to Lake Tana in the Ethiopian Highlands. Figure C.4 provides an overview of the geography of the area. Precipitation in this area is a result of the African Monsoon weather system and, more specifically, the location of the Intertropical Convergence Zone (ITCZ). When this zone moves far north, there is more monsoon rain in the Ethiopian Highlands, and vice versa (Said, 1993). The ideal “paleoclimate archive” should thus be located in this region and have observations with a high time resolution far back in time.

The closest natural archive to the Ethiopian Highlands that is subject to similar patterns of monsoon precipitation and that has good resolution, is found in Qunf cave in Southern Oman (Fleitmann et al, 2003). Southern Oman is exposed to the same patterns of monsoon precipitation as the Ethiopian Highlands, as reflected in Figure C.4. The time series at Qunf is derived from speleothem data, measuring the oxygen isotope content ($\delta^{18}\text{O}$) in a stalagmite from the cave. In the scientific literature, $\delta^{18}\text{O}$ -levels have been frequently used as proxies for historical rainfall. As explained by Fleitmann et al (2003, 2007), the data from Qunf cave should provide a reliable indicator for monsoonal precipitation in the northwestern Indian Ocean during the period. It contains 1412 observations from the time period 8608 BCE – 1650 CE with a long break 750 BCE – 638 CE. The time resolution of observations (i.e. years between observations) during the most relevant period 3000 – 1000 BCE is 1-14 years with an average of 3.85 years. Since the $\delta^{18}\text{O}$ -numbers do not have a natural interpretation in our setting, we create a *Nile flood index*, normalized to the range [0,10], where higher numbers indicate more monsoon rains, which should, in turn, imply higher levels of Nile inundations. We interpolate for missing years and also calculate 5-year averages for the study period in the analysis below and apply a Butterworth filter in order to smooth out the influence of extreme individual years (see Section 6.2 below for a discussion about the filters employed). Appendix D contains further details about the construction of this variable.

We have carried out an extensive validation of these data as a proxy for Nile inundations (see Appendix C). Despite the fact that this is an indirect proxy for Nile inundation levels, the data also offers some advantages over the use of Nilometer data. A key advantage of our measurement is that we can be assured that our $\delta^{18}\text{O}$ -time series has not in any way been contaminated or manipulated by man, either consciously or by a change of measurement over time.²⁹

Our validation exercise compares our proxy for Nile inundations to other indirect proxies such as sediment levels in the Nile delta, in lakes in the Ethiopian highlands and in the Mediterranean.

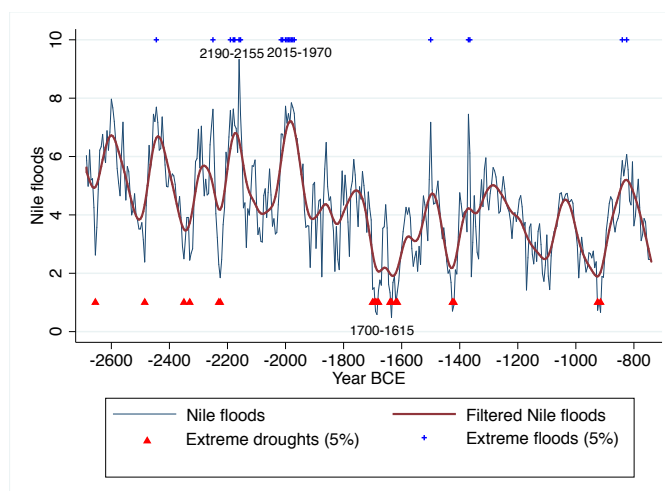
²⁹We also discuss other issues regarding the use of Nilometers in the Appendix.

We have also compared our measure to direct measurements of Nile inundations (i.e., Nilometer records). Although consistent Nilometer measurements are only available for a much later period (starting in 622 CE), the Qunf cave data is available until 1650 CE (with some long gaps in between), which implies an overlap between the two series exists and that a direct comparison is possible. Appendix C shows that there is a close correspondence between our proxy and other measurements of Nile inundations, see that appendix for additional details.

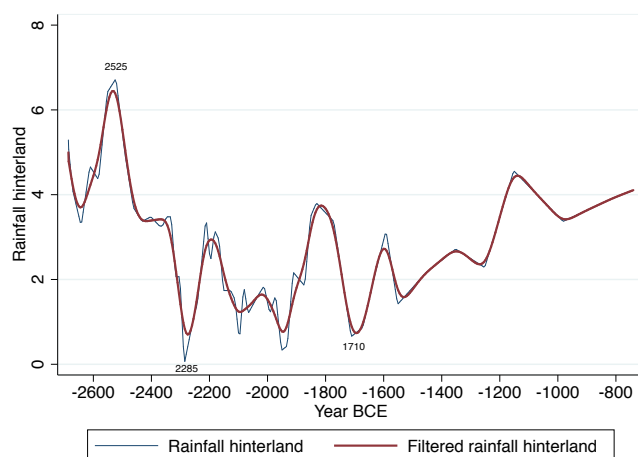
Panel (a) in Figure 2 shows the *Nile floods* index for the 2685-750 BCE period. The thin line shows the actual estimated index observations for 5-year intervals whereas the medium thick line shows the resulting Butterworth filter, using a 20-period filter. Blue-plus signs indicate the upper five percent observations (extreme floods) and the red triangles periods with the lower five percent periods (extreme droughts). A number of observations stand out. First, the precipitation patterns are cyclical with peak-to-peak cycles lasting on average about two centuries. The two most extreme flood periods happened 2190-2155 BCE and 2015-1970 BCE, i.e. right before and right after IMP I. Third, by far the worst drought happened during 1700-1615, i.e. in the build-up to and during IMP II. Fourth, the general declining trend in the MPI implies that Nile inundations gradually diminished during the classical period.

As mentioned before, productivity in Ancient Egypt was a non-linear function of the extent of Nile inundations, as productivity peaked when inundations were not too large and not too small. Lacking direct information to map Nile inundations into productivities, in our baseline specification we capture these non-linearities by considering a quadratic polynomial of *Nile floods*. We also employ alternative specifications, such as dummies for periods of extreme floods.

5.3.2. Productivity in the hinterland. To proxy productivity in the hinterland, we use rainfall in the area surrounding the Nile Valley. During wet periods, Mediterranean winter rainfall levels implied a relatively green and habitable hinterland between Egypt and the Southern Levant whereas dry periods gave rise to desertification. Rainy years should also have been associated with greater human activity in the hinterland drylands in general, including mining and pastoralism. To our knowledge, there is no detailed climate archive available from the Sinai, which would be the ideal location for such data. As a proxy for Mediterranean winter rainfall in northern Sinai, we use speleothem data on the $\delta^{18}\text{O}$ isotope in stalagmites from the Soreq cave, about 18 km west of Jerusalem and 60 km northeast of Gaza in contemporary Israel (Bar-Matthews and Ayalot, 2003, 2011). Throughout history, Gaza has been the Levantine outpost towards Sinai and Egypt (see Figure C.6 in Appendix C.2 for a map of the area). Appendix C provides a validation



(a) Nile Floods (Qunf data)



(b) Rainfall hinterland (Soreq data)

Figure 2. CLIMATE PROXIES: NILE FLOODS AND RAINFALL HINTERLAND. The figures show our filtered and non-filtered measures of Nile floods (Qunf data) and Rainfall hinterland (Soreq data) with extreme values and years indicated in the graph.

exercise and shows that there is indeed a strong correspondence between our proxy of rainfall and other proxies available for this area.

The data collected from this cave spans a very long period of 200,000 years and has been frequently used in the science literature as a reliable benchmark series. The average time resolution is not as fine for the 2685 – 750 BCE as in the Qunf cave and has an average of 28.1 years with

a maximum of 100. For missing years, we conduct an identical linear interpolation procedure as with the Qunf-data and then calculate 5-year averages levels and smooth the resulting series with a Butterworth filter. We also create a similar rainfall index as above, normalized to the range [0,10]. Panel (b) in Figure 2 shows the *rainfall hinterland*-index based on the data from Soreq. Also here, we see a declining trend in precipitation but it reaches a minimum around 1700 BCE and thereafter increases again. Precipitation reached a peak in 2525 BCE and then dropped dramatically to its lowest point just 240 years later in 2285 BCE. Also this hinterland series indicate a severe drought in 1710 BCE, right before IMP II.

As opposed to the previous case, productivity and the amount of rainfall in the hinterland are likely to be linked in a monotonic way since this area is particularly dry and rainfall is rarely excessive. In the empirical exercise we also allow for non-linearities in a similar way as in the case of Nile inundations and we do not find support for such a relationship.

Lastly, a key assumption in our theoretical model is that the weather systems affecting the core and the hinterland are uncorrelated. Table C.2, column (5) and Figure C.9 report the correlation coefficient and 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 that figure and the correlation coefficient (equal to 0.008) show, there is zero correlation between the two climate during the period considered in the empirical analysis.

6. POLITICAL INSTABILITY IN ANCIENT EGYPT: EMPIRICS

This section describes the sample period, our empirical specifications as well as the key features of our empirical strategy, while Section 7 presents our main results.

6.1. Sample. Our empirical analysis examines the evolution of political instability and government capacity in ancient Egypt over the period 2685 BCE – 750 BCE, which covers the Old, Middle, and New Kingdoms, as well as the First, Second and part of the Third Intermediate periods, see Section 4.1 for details about the chronology of Ancient Egypt. Due to data limitations in some of our key variables, however, most of our regressions focus on the period 2686 BC to 1140 BC.³⁰ Given the antiquity of our data, measurement error and, in particular, the precision in the dating of our variables is an important concern.³¹ For this reason, we have aggregated

³⁰The reason why we focus on this shorter period is that the level of resolution of the proxy for rainfall in the hinterland deteriorates considerably after 1140 BC. However, we show as well that our results hold when the whole period is considered.

³¹See Appendix B.3 for details on the dating of our historical, archaeological and paleoclimatic variables.

our variables in five year-periods.³² Thus, unless otherwise noted, the unit of observation in our regressions is five-year periods. We also show that results are robust to considering alternative levels of aggregation. In spite of this, measurement error cannot be discarded and most likely will lead to attenuation bias.

6.2. Empirical specifications and empirical strategy. Due to the difficulties in obtaining historical time-varying data on early state formation and consolidation, a good part of the literature investigating early state formation is forced to adopt a cross-sectional approach.³³ This approach assumes that after conditioning on observables, cross-sectional units are identical and, therefore, all differences in observed conditional outcomes can be attributed to differences in the treatment. A well-known weakness of this approach is its vulnerability to omitted variable bias. Instead, our identification strategy exploits exogenous time-varying shocks to the productivity of the Nile valley as well as its neighboring areas and investigates their effect on political instability and state capacity. By considering just one unit (Ancient Egypt), our approach eliminates the effect of unobservable differences across units, and thus eliminates a potential source of omitted variable bias. Our main identifying assumption is that, conditional on observables, the different time periods are comparable, so that changes in outcomes can be attributed to the effect of the treatment variables.

As argued by Hsiang (2017), climate events can have both direct and indirect effects on outcomes. *Direct* effects refer to the contemporaneous impact of weather conditions, whereas *indirect* effects allude to the impact of climate on individual's beliefs which may affect their decisions and, therefore, the resulting outcomes. This is particularly relevant in our case, as our theory involves both direct and indirect effects of climate. For instance, a bad harvest in the core has a direct effect on contemporaneous tax collection but also an indirect one on future collection, that operates through the decision of farmers to migrate to/from the core. This decision, in turn, is made based on beliefs about the evolution of climate. These beliefs might need some time to be formed and have an impact on outcomes. It is widely assumed that agents facing some weather event for long periods of time will update their beliefs about the weather, whereas individuals facing those events during a short period will not alter their beliefs.

³²For the sake of robustness, other frequencies have been also been employed, including annual and decadal observations, see Table 2.

³³As summarized in Sections 1 and 2, other papers have used a time series approach in a historical context. One example is Chaney (2013), but his study focuses on a much later period of Egyptian history and that cannot be considered as a time of early state consolidation.

In order to capture both direct and indirect effects, papers using climatic time series often filter the data so that the latter captures the low-frequency component (see Zang et al. 2007, Tol and Wagner 2010). We also follow that approach here, and apply a low-pass filter to our climate proxies. More specifically, we apply the Butterworth low-pass filter, see Pollock (2000) for details. This approach also has a downside. As noted by Hsiang (2017), identification can become more problematic since the use of filtered data typically requires the comparison of populations over long periods of time, which can undermine our main identifying assumption. To be able to cope with this threat, our models contain time trends as well as period dummies.

Additionally, our empirical models allow for non-linear effects (by considering quadratic terms of the independent variables), as well as for dynamic effects (by allowing for lags of the independent and dependent variables). We estimate models of the form,

$$Y_t = \beta_0 + f_1(\text{NILE FLOODS}) + f_2(\text{RAINFALL HINTERLAND}) + X_t\beta + \epsilon_t,$$

where f_1 and f_2 are polynomials of the climate proxies and their lags, X_t is a vector of controls (that includes time trends, period fixed effects, lags of the dependent variables and other control variables, see Section 7 for a description), and ϵ_t is an uncorrelated residual term. Estimation is typically conducted by OLS with robust standard errors except when static models are considered, in which case Newey-West standard errors are used instead, since errors are autocorrelated in that case. Ordered logit specifications are also employed as several of our dependent variables are discrete.

7. RESULTS

This section describes our empirical results. Section 7.1 presents the main analysis, which focuses on the relationship between weather shocks and political instability. Section 7.2 provides additional evidence by analyzing the relationship between proxies of state capacity (in particular, pyramid construction and area under control of state) and the weather shocks. Variable definitions, summary statistics as well as a preliminary analysis that examines the time-series properties of the data are provided in Appendix D.

7.1. Political Instability. Table 1 presents our main results relating political instability and the proxies for the weather conditions in the Nile and surrounding areas. Two types of models have been estimated. Columns 1–5 consider static models (i.e., models that only include contemporaneous values of the variables), while columns 6–9 consider dynamic models instead. More

Dependent Variable: Pol Instability									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NILE FLOODS _t	-0.166 (0.185)	-1.082** (0.017)	-0.698** (0.028)	-0.691** (0.031)	-0.690** (0.033)	-0.228** (0.033)	-0.166 (0.049)	-0.233** (0.032)	-0.176* (0.039)
NILE FLOODS _t ²		0.101** (0.034)	0.062* (0.065)	0.061* (0.076)	0.061* (0.077)	0.026** (0.016)	0.020* (0.017)	0.027** (0.015)	0.022* (0.013)
RAINFALL HINTERLAND _t				-0.035 (0.582)	-0.052 (0.772)	-0.506** (0.004)			
RAINFALL HINTERLAND _t ²					0.003 (0.906)				
RAINFALL HINTERLAND _{t-1}						0.558** (0.002)	0.058** (0.001)	-0.467* (0.014)	
RAINFALL HINTERLAND _{t-2}								0.528** (0.007)	0.065** (0.018)
<i>t</i>	0.011** (0.043)	0.010* (0.067)	0.009** (0.021)	0.007 (0.130)	0.007 (0.130)	0.002 (0.253)	0.002 (0.219)	0.002 (0.155)	0.002 (0.148)
<i>t</i> ²	-0.000** (0.036)	-0.000* (0.054)	-0.000** (0.017)	-0.000* (0.089)	-0.000* (0.092)	-0.000 (0.364)	-0.000 (0.286)	-0.000 (0.239)	-0.000 (0.203)
TENURE _t			-0.048*** (0.000)	-0.048*** (0.000)	-0.048*** (0.000)	0.066*** (0.000)	0.067*** (0.000)	0.067*** (0.000)	0.067*** (0.000)
TENURE _t ²			0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
TENURE _{t-1}						-0.042*** (0.000)	-0.043*** (0.000)	-0.042*** (0.000)	-0.042*** (0.000)
TENURE _{t-1} ²						0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
<i>c</i>	0.863 (0.224)	2.823** (0.013)	2.402*** (0.002)	2.601*** (0.001)	2.614*** (0.001)	-0.128 (0.671)	-0.316 (0.243)	-0.166 (0.570)	-0.338 (0.197)
<i>R</i> ²	0.159	0.215	0.393	0.395	0.395	0.757	0.753	0.757	0.754
Obs	310	310	309	309	309	308	308	308	308
Lags Dep. variable	-	-	-	-	-	×	×	×	×
Infor. Crit (BIC)	729.20	713.81	644.02	649.89	654.79	394.65	395.02	394.10	392.97
Cumby-Huizinga test (p-val)	0.00	0.00	0.00	0.00	0.00	0.138	0.143	0.134	0.1492
Standard errors	HAC	HAC	HAC	HAC	HAC	HAC-[HC]	HAC-[HC]	HAC-[HC]	HAC-[HC]

Table 1. POLITICAL INSTABILITY AND CLIMATE SHOCKS. Dependent variable is POL INSTABILITY. Estimation is by OLS and p-values (based on heterokedasticity and autocorrelation consistent (HAC) or heteroskedasticity consistent [HC] standard errors are reported in brackets in columns 1–5 and 6–9, respectively. The period covered is 2685 BCE – 1140 BCE and the unit of analysis is 5 year periods. The table reports the value of the Bayesian information criterion (BIC) corresponding to each of the models considered as well as the p-value of the Cumby-Huizinga test of residual autocorrelation. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

specifically, in those columns we estimate autoregressive distributed lag models (ADLM). These models incorporate lags of both the dependent and the independent variables. As is customary in time series modelling, we use information criteria to choose the number of lags in the model. In particular, we use the Bayesian information criterion (BIC) and Table 1 reports its value for each of the estimated models.³⁴ When dynamic models are considered, autocorrelation in the residual term becomes a serious concern, as in that case lagged values of the dependent variable are likely to be correlated with the (autocorrelated) residuals. Thus, it is very important to make sure that the residuals are not autocorrelated. We achieve this by introducing lags of the relevant variables, as dictated by the BIC. Table 1 reports the value of the Cumby-Huizinga test of residual autocorrelation for each of its columns, and shows that the hypothesis of lack of autocorrelation in the residuals cannot be rejected in columns 6–9. Accordingly, for the latter columns we report two types of standard errors: HAC (in brackets) and heteroskedasticity consistent (HC) standard errors (in square brackets). Notice however that, since the hypothesis of autocorrelation in the residuals is rejected, there is no need of computing HAC standard errors (and in fact, they are in general smaller than HC ones in columns 6–9). For this reason (and given as well that using HC standard errors seems to be a more conservative option), in the following tables we report HC standard errors exclusively when dynamic models are considered.³⁵ All models in Table 1 are estimated by OLS but, for the sake of robustness, Table D.4 in Appendix D.4 reestimates that table by maximum likelihood in an ordered logit specification obtaining identical conclusions.

Column 1 regresses ruler instability on NILE FLOODS and on a quadratic time trend that aims to capture systematic trends in the data. Since the residual term is heavily autocorrelated in the absence of dynamic terms (see the autocorrelation test at the bottom of Table 1), heteroskedasticity and autocorrelation consistent (HAC) standard errors are computed.³⁶ Column 1 shows that the coefficient of NILE FLOODS is negative but not statistically different from zero. Column 2 allows for a non-linear relationship between rulers instability and NILE FLOODS by introducing the square of the latter variable. Allowing for a non-linear relationship delivers estimates that are more precisely estimated, and now both NILE FLOODS and its square are significant at the 5% level. While the former keeps the negative sign, the square term is positive, suggesting that ruler instability is lowest for intermediate Nile floods. This result bears resemblance to that reported

³⁴See Neath and Cavaugh (2012) for details about this IC. Results are also robust to considering alternative information criteria, such as Akaike.

³⁵All results are robust to using HAC standard errors instead.

³⁶More specifically, the reported p-values are based on Newey-West standard errors, where the bandwidth is chosen following the criterion suggested by Lazarus et al (2018). Results are very robust to alternative values of the bandwidth.

in Chaney (2013) for a much later period (1169–1425 CE), although that paper mostly focuses on the absence of head judge replacements in periods of aberrant Nile floods.³⁷

Column 3 introduces $TENURE_t$ and its square to account –in a flexible way– for the possibility that ruler replacement becomes more likely as the ruler gets old. This variable measures the number of years from the ruler’s first year in office to the first year of the current period. Both variables are highly significant and indicate that ruler instability is highest the first years of the period in office and after a long tenure. The results in column 2 are not affected by the inclusion of these variables. Column 4 introduces the contemporaneous value of $RAINFALL_HINTERLAND$, our proxy for rainfall in the areas surrounding the Nile valley. Its associated coefficient is negative, indicating that more rain tends to decrease political instability, although it is far from being statistically significant (and its introduction increases the value of the BIC). To explore whether there is a non-linear relationship between pharaoh instability and $RAINFALL_HINTERLAND$, column 5 adds the square of the latter variable, which turns out to be insignificant as well. The lack of significance of the rainfall variables in columns 4 and 5 is unsurprising in the light of our theory, as lagged rainfall is not included in the model and this omission is likely to bias upwards the coefficient of contemporaneous rainfall, given the positive correlation between contemporaneous and lagged rainfall.

To examine the intertemporal effects of our variables, columns 6–9 consider dynamic models. Column 6 allows for the introduction of one lag of each of the variables in Column 5, and presents the model that minimizes the BIC.³⁸ It also includes lags of the dependent variable so that uncorrelated residuals are obtained (two lags are enough to achieve so). Introducing dynamics improves substantially the fit of the model, as shown by the rise in R^2 (that increases from 0.39 in column 5 to 0.76 in column 6), and by the decrease in the BIC (that almost halves). Column 6 shows that the effect of Nile floods on political instability is similar as in previous tables and is concentrated in the contemporaneous period, as lags of the level and the square of Nile floods are not significant (and introducing them in the model results in an increase in the BIC).

³⁷See Section 2 for a discussion of Chaney’s paper.

³⁸This specification doesn’t include the square term of $RAINFALL_HINTERLAND$ or its lags as they were never statistically significant and their inclusion in the model considerably increased the BIC. Models including lagged $NILE_FLOODS$ are not reported as lags were never significant and their inclusion created considerable multicollinearity. We also considered regressions that allow for the inclusion of two lags of the variables, in general, they resulted in higher values of the BIC.

While the conclusions regarding the role of NILE FLOODS on political instability remain unchanged, allowing for lagged regressors modifies substantially those related to the role of RAINFALL HINTERLAND. Column 6 shows that the lagged value of RAINFALL HINTERLAND has a positive and significant effect on ruler instability. This result is consistent with our hypothesis that a reduction in the circumscription of the Egyptian state increases future ruler instability, as it triggers the reallocation of the population outside the state's borders (or the threat to do so). On the other hand, the contemporaneous value of RAINFALL HINTERLAND is now negative and significant, implying that higher contemporaneous rainfall can bring about a short-term strengthening of the pharaoh's power by making the nearby areas more easily exploitable.

Notice that an improvement in the weather conditions in the areas surrounding the Nile valley could make the state more vulnerable to attacks from neighboring peoples. However, this would contribute *contemporaneously* to an increase in instability, and not the opposite, as we find. Our theoretical model suggests that the main impact of RAINFALL HINTERLAND on political instability is through (future) migration pressures, which implies that the main effect will be driven by lagged values of RAINFALL HINTERLAND. To examine this prediction, column 7 drops the contemporaneous value of RAINFALL HINTERLAND. Dropping the contemporaneous level of RAINFALL HINTERLAND leads to a decrease in the BIC, meaning that under this criterion this model is preferred. The coefficient of RAINFALL HINTERLAND in $(t - 1)$ is positive and significant with a magnitude that is roughly the sum of the contemporaneous and the lagged coefficients of RAINFALL HINTERLAND in column 6. This regression confirms that an increase of RAINFALL HINTERLAND has a positive and significant impact on ruler instability but that the effect needs some time to become apparent. Columns 8 and 9 are similar to columns 6 and 7 but the rainfall variables are lagged one more period. This improves the specification of the model, as shown by a decrease in the BIC. In particular, the model with the smallest BIC is that in column 9 (where only $\text{RAINFALL HINTERLAND}_{t-2}$ is introduced) and, thus, this is our preferred specification. The conclusions in columns 8 and 9 are very similar to those in columns 6 and 7. Summarizing, the models in Table 1 explain a large fraction of the variance of POL INSTABILITY (around 75%). Dynamic models (columns 6 to 9) provide a much better fit and show that political stability in Ancient Egypt was highest when Nile floods were not too low and not too high and when the conditions in the neighboring areas in previous periods were rough.

Finally, we provide a sense of the magnitudes of the coefficients using the estimates provided in our preferred specification (column 9). The effects of weather conditions in the Nile and surrounding areas on political instability are large and not very different in magnitude. In the case of NILE FLOODS, since the effect is non-linear we provide a graph depicting the marginal effect for all values of NILE FLOODS (see Figure D.10 in Appendix D.4). This graph shows that the

marginal effect is negative (positive) for small (large) values of NILE FLOODS. We also provide here a couple of examples for a small and for a large value of NILE FLOODS. An increase of NILE FLOODS from the first to the 25% percentile (around 1.5 standard deviations) decreases ruler instability by .14 (which amounts to a decrease of 17% of the standard deviation of RULERS). An increase of NILE FLOODS from the 75% to the 99% percentile (around 1.5 standard deviations) produces an increase in the dependent variable of .17, (which corresponds to an increase of 21% of the standard deviation of RULERS). An increase of RAINFALL HINTERLAND_{t-2} of 1.5 standard deviations leads to an increase in RULERS_t of .12, which amounts to 15% of the standard deviation of POL INSTABILITY.

Tables 2 examines the robustness of the results in Table 1. Dynamic models selected using the BIC are considered in all columns, but the coefficients of the lags of the dependent variable are not reported to save space. Column 1 is similar to column 9 in Table 1 but captures the non-linear relationship of Nile floods and ruler stability by defining two dummy variables, EXTREME FLOOD and EXTREME DROUGHT, that are equal to 1 when the corresponding observation of NILE FLOODS is in the upper or bottom 5% of the distribution of NILE FLOODS, respectively. The coefficient of NILE EXTREME DROUGHT is positive and highly significant while that of NILE EXTREME FLOOD is close to zero and insignificant, suggesting that droughts had a more severe impact on ruler instability.

Column 2 considers whether there is an interaction effect between the two weather shocks. To examine whether the effect of RAINFALL HINTERLAND is heterogeneous depending on whether Nile floods are extreme or not, we define a new dummy, NILE EXTREME, that is 1 whenever Nile floods are extremely low or extremely high (i.e., whenever either NILE EXTREME DROUGHT or NILE EXTREME FLOOD are equal to 1). Column 2 shows that both NILE EXTREME_t and RAINFALL HINTERLAND_{t-2} are positive and significant, as expected, but the interaction of the two is negative and also significant. Furthermore, using an F-test we cannot reject that the sum of the coefficients of RAINFALL HINTERLAND_{t-2} and the interaction term is equal to zero.³⁹ This suggests that Nile floods had a first-order effect on ruler instability and that whenever floods were extreme, instability followed, irrespective of the conditions in the hinterland. These conditions, however, became important in “normal” Nile periods (where “normal” here refers to the central 90% of Nile flood values), when the state was not under the effect of extreme Nile shocks. This result is reasonable since as it is well-known that Egypt’s economy heavily hinged on the extent of Nile floods and whenever those severely failed, famine and unrest followed.

³⁹Using a similar F-test, we do reject that the sum of the coefficients of NILE EXTREME_t and the interaction term is zero.

Dependent Variable: Pol Instability										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NILE EXTREME FLOOD _t	-0.066 (0.546)									
NILE EXTREME DROUGHT _t	0.167*** (0.003)									
NILE FLOODS _t			-0.193* (0.061)	-0.256* (0.050)	-0.228** (0.025)	-0.179* (0.079)	-1.148 (0.104)	-0.263** (0.030)	-1.112** (0.022)	-0.045* (0.072)
NILE FLOODS _t ²			0.021* (0.051)	0.027* (0.080)	0.024** (0.043)	0.022* (0.069)	0.143* (0.065)	0.029** (0.031)	0.121** (0.039)	0.006** (0.043)
RAINFALL HINTERLAND _{t 2}	0.058** (0.038)	0.061** (0.031)	0.073*** (0.006)	0.080** (0.033)	0.042 (0.101)	0.062** (0.013)	0.385** (0.031)	0.052* (0.077)	0.211** (0.016)	
RAINFALL HINTERLAND _{t 10}										0.018** (0.012)
NILE EXTREME _t		0.251** (0.032)								
NILE EXTREME _t RAINFALL HINTERLAND _{t 2}		-0.100* (0.054)								
TENURE _t	0.066*** (0.000)	0.067*** (0.000)	0.065*** (0.000)	0.003 (0.630)	0.044*** (0.000)	0.070*** (0.000)	0.323*** (0.000)	0.064*** (0.000)	0.046*** (0.006)	0.038*** (0.000)
TENURE _t ²	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.000 (0.958)	-0.000*** (0.000)	-0.001*** (0.000)	-0.004*** (0.000)	-0.001*** (0.000)	-0.001*** (0.003)	0.0003*** (0.000)
TENURE _{t 1}	-0.045*** (0.000)	-0.044*** (0.000)	-0.041*** (0.000)	-0.004 (0.533)	-0.031*** (0.000)	-0.040*** (0.000)	-0.200*** (0.002)	-0.048*** (0.000)	-0.024 (0.138)	-0.166 (0.109)
TENURE _{t 1} ²	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.000 (0.997)	0.000*** (0.000)	0.001*** (0.000)	0.003*** (0.000)	0.001*** (0.000)	0.0003** (0.015)	0.000 (0.017)
<i>t</i>	0.002 (0.360)	0.001 (0.463)	0.004* (0.098)	0.006** (0.028)	0.002 (0.208)	0.002* (0.090)	0.010 (0.375)	0.001 (0.637)	0.023 (0.148)	0.000 (0.182)
<i>t</i> ²	-0.000 (0.384)	-0.000 (0.503)	-0.000 (0.139)	-0.000** (0.028)	-0.000 (0.235)	-0.000* (0.070)	-0.000 (0.486)	0.000 (0.544)	-0.000 (0.347)	-0.000 (0.254)
<i>R</i> ²	0.754	0.753	0.762	0.631	0.366	0.710	0.494	0.767	0.666	0.926
Obs	308	308	296	299	308	388	308	307	150	1536
Lags Dep. variable	×	×	×	×	×	×	×	×	×	×
Inf. Crit (BIC)	393.6	393.9	375.8	367.9	490.4	497.9	377.5	403.3	288.6	-47.76
Cumby-Huizinga test (p-val)	0.112	0.162	0.142	0.637	0.144	0.941	-	0.438	0.842	0.149
Period dummies	-	-	-	-	-	-	-	×	×	-
Estimation	OLS	OLS	OLS	OLS	OLS	OLS	OLOGIT	OLS	OLS	OLS

Table 2. POLITICAL INSTABILITY AND CLIMATE SHOCKS: VARIATIONS. Dependent variable is POL INSTABILITY except in columns 4 and 5 where POL INSTABILITY(MET) and POL INSTABILITY(DUMMY) were employed instead. Estimation is by OLS except in column 7 where maximum likelihood in an ordered logit specification was employed. The unit of analysis is 5 year periods, except in columns 9 and 10 where the unit of analysis is 10 year and 1 year periods, respectively. To save space lags of TENURE and TENURE² have been omitted, see the main text. The sample period is 2686 BCE – 1140 BCE except in columns 5, where it is 2686 BCE – 740 BCE. All columns contain lagged values of the dependent value as dictated by the BIC. P-values based on heterokedasticity consistent (HC) standard errors are reported in brackets. The table reports the p-value of the Cumby-Huizinga test of residual autocorrelation (first lag). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The remaining columns in Table 2 provide additional variations: Column 3 uses an alternative procedure to filter the weather data. In particular, the Baxter King filter is employed in an specification otherwise identical to that in column 9, Table 1. Column 4 uses a different chronology

provided by the Department of Egyptian Art of the Metropolitan museum to code the political instability variable, POL INSTABILITY(MET).⁴⁰ Column 5 considers an alternative definition of the political instability variable employed in Table 1, denoted as POL INSTABILITY(DUMMY), where all the strictly positive values in POL INSTABILITY are coded as 1 (i.e., ruler replacements and centralization break down are both coded as 1). Column 6 considers a longer time period than in our baseline regressions, more specifically from 2686 BCE to 750 BCE.⁴¹ Since the dependent variable is discrete, column 7 considers an ordered logit specification. Column 8 introduces dummies for the different periods in which Egyptian chronology is typically divided (Old, Middle and New Kingdoms as well as the first and second intermediate periods).⁴² As the probability of transition from one period to another is likely to be affected by climate changes, these variables can be considered as “bad controls”. However, by including them in the regression we can check whether the results in Table 1 are driven by the highly unstable first and second intermediate periods or whether they still hold when only within period variation is considered. Finally, columns 9 and 10 replicate our preferred specification (column 9 in Table 1 using different units of analysis. In particular, columns 9 and 10 consider 10 year and 1 year periods, respectively, as unit of analysis.⁴³ shows that results are robust to considering a different level of temporal aggregation. In particular, the unit of analysis employed in this column is 10-year periods.⁴⁴ In all of these variations very similar conclusions as those in Table 1 are obtained.

Tables 1 and 2 show that political instability is associated to reductions in the degree of circumscription in Ancient Egypt, and that this conclusion is robust to a number of variations, including alternative chronologies, the introduction of period fixed effects, the use of alternative definitions of the climate variables and the use of alternative estimation methods, among others. A common concern associated to these tables, however, is the fact that rulers can be replaced for reasons that do not necessarily involve social unrest and political instability, such as natural death or disease. To address for this concern, Table 3 considers an alternative proxy of political instability, POL INSTABILITY(DYNASTIES), which only considers those ruler replacements that also led to the removal of a dynasty. This variable presents some advantages with respect to POL INSTABILITY,

⁴⁰See http://www.metmuseum.org/toah/hd/phar/hd_phar.htm (October 2002).

⁴¹We don't consider this interval in Table 1 because the level of resolution of RAINFALL HINTERLAND becomes very crude after 1140 BCE.

⁴²See Appendix B.3 for details about the Egyptian chronology.

⁴³In column 10, the model preferred by the BIC contains three lags of TENURE and its square that for the sake of brevity have been omitted from Table 2.

⁴⁴As in all other columns, we employed the BIC for model selection and 5 lags of the dependent variable were needed to obtain uncorrelated residuals, whose estimates are not reported in Table 2 to save space.

Dependent Variable: Pol Instability (Dynasties)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
NILE FLOODS _t	-0.001 (0.937)	-0.207* (0.065)	-0.223* (0.063)	-0.224* (0.058)	-0.224* (0.053)	-5.993*** (0.001)	
NILE FLOODS _t ²		0.023* (0.064)	0.025** (0.046)	0.025** (0.044)	0.025** (0.040)	0.674*** (0.000)	
NILE EXTREME FLOOD _t							2.722*** (0.008)
NILE EXTREME DROUGHT _t							2.884** (0.015)
RAINFALL HINTERLAND _{t-2}					0.002 (0.878)	0.622* (0.078)	0.480* (0.051)
TENURE DYNASTY _t			0.002 (0.297)	0.002 (0.309)	0.002 (0.300)	0.015 (0.188)	0.013 (0.214)
TENURE DYNASTY _{t-1}			-0.002 (0.462)	-0.002 (0.474)	-0.002 (0.464)	0.012 (0.238)	0.006 (0.518)
<i>t</i>				0.000 (0.714)			
<i>t</i> ²				-0.000 (0.616)			
<i>c</i>	0.029 (0.723)	0.467* (0.068)	0.414 (0.155)	0.420 (0.150)	0.413 (0.161)		
<i>R</i> ²	0.840	0.843	0.849	0.850	0.849	0.702	0.694
Obs	308	308	308	308	308	308	308
Lags Dep. variable	×	×	×	×	×	×	×
Inf. Crit (BIC)	135.28	134.21	133.11	144.4	139.73	151.216	155.80
Cumby-Huizinga test (p-val)	0.48	0.42	0.54	0.57	0.53	–	–
Estimation	OLS	OLS	OLS	OLS	OLS	OLOGIT	OLOGIT

Table 3. POLITICAL INSTABILITY(DYNASTIES) AND CLIMATE SHOCKS. Dependent variable is POL INSTABILITY(DYNASTIES). P-values based on HC standard errors are reported in brackets. The table reports the value of the Bayesian information criterion (BIC) as well as the p-value of the Cumby-Huizinga test of residual autocorrelation. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the most important one being that the fall of a ruler *and* a dynasty is likely to be the consequence of highly turbulent events, and thus, it could be a better proxy for political instability.

On the negative side, however, the classification of rulers in dynasties dates back to Manetho's *Aegyptiaca* (3rd century BC) and the criteria followed to elaborate that list is not always transparent, as explained in more detail in Appendix B.3. Having this limitation in mind, we have explored whether our climate proxies have any power in predicting dynasty replacement. Table

3 presents the results, obtained in a set up otherwise very similar to that in the previous tables.⁴⁵ Column 1 in Table 3 regresses DYNASTIES on NILE FLOODS and lags of the dependent variable (so that the residuals become uncorrelated), and shows that there is not a statistically significant (linear) relationship between the former variables.

However, by adding the square of NILE LEVEL to the former specification as in column 2, the results change. Column 2 shows that higher dynastic instability is associated with too low or too high Nile floods. We used the BIC to test whether lagged values of the Nile variables should be introduced in the regression and the BIC was lowest when only contemporaneous values of Nile water levels are considered. Column 3 introduces TENURE DYNASTY, which measures the number of years a dynasty has been in office up to the first year of the current time period, and its first lag, again selected using the BIC.⁴⁶ Both variables are significant but the results in column 2 are unaffected by the introduction of these variables. Column 4 adds a quadratic trend that doesn't change the conclusions of column 3. Both terms of the trend are insignificant and since their introduction increases the BIC, they are removed from subsequent specifications. Column 5 considers the introduction of (at most two lags of) rainfall in the hinterland. The model that minimizes the BIC is one with only one lag, RAINFALL HINTERLAND_{*t* - 2}, as in the previous tables. The coefficient of this variable is positive but not statistically significant (while the coefficients of the Nile floods variables remain very similar). Column 6 re-estimates column 5 by maximum likelihood in an ordered logit specification. The precision of the estimates in this case improves and RAINFALL HINTERLAND_{*t* - 2} has a positive and significant coefficient, suggesting that higher rainfall leads to political instability in the following periods. Finally, column 7 captures aberrant Nile floods by two dummies reflecting extreme floods and droughts. Both variables are positive and significant in the specification, suggesting that serious political instability was more likely in periods of extreme Nile behaviour.

Summarizing, the results in Tables 1 to 3 show that political stability in Ancient Egypt is associated with high productivity in the Nile valley *together with* low productivity in the nearby areas. Furthermore, the size of both effects is large and comparable in magnitude. The following section presents additional evidence supporting our claim that periods of high circumscription helped the consolidation of the state.

⁴⁵Table 3 focuses exclusively on dynamic models as they provide a much better fit for the data and there is evidence that some of our key variables have a delayed response. Similar conclusions can be obtained, however, if static models with HAC standard errors are considered instead.

⁴⁶Quadratic terms of TENURE DYNASTY and its lags were also considered but they increased the BIC –without affecting the conclusions about the Nile variables –.

7.2. Additional evidence: State capacity. The theory outlined in Section 3 posits that periods of high circumscription are associated with a higher tax collection and a greater consumption of the state and the ruler. This section explores the impact of changes in circumscription on our proxies for the latter.

Our first proxy for the strength of the state is the area under state control. We consider the variable AREA which is the log of the state controlled area (in km²).⁴⁷ Although admittedly this variable is likely to be a rough proxy for the true area, Figure B.3 shows that the area under state control varied widely over time. The largest surface was achieved during the New Kingdom where the state controlled territories up to the current Syria. On the other hand, the minimum area was reached during the intermediate periods, where central authority collapsed. In our view, conquered areas were primarily used for resource extraction and hence served to increase state consumption.

Our main hypothesis is that good harvest years should lead to an increase in the size of the area under the control of the state. In addition, in line with the circumscription hypothesis, wetter conditions in the surrounding areas would weaken the power of the state in the medium and long-run, leading therefore to a reduction in the area under the state's control. It is also worth noticing that the existence of the latter medium/long-term relationship does not preclude the existence of a (potentially different) short-term response to changes in weather conditions in the hinterland. It is reasonable to expect that in the short run more favorable conditions in the areas surrounding the Nile valley would increase the relative value of those areas and at the same time could facilitate army incursions in them. This could bring about a short-term positive effect of rainfall in the hinterland.

Table 4 examines these predictions. As in previous tables, we allow for the inclusion of lags of both the dependent and the independent variables and make sure that the residual term is uncorrelated (see the values of the Cumby-Huizinga test at the bottom of Table 4). The number of lags to be included in the regression is selected using the BIC and only the preferred models are reported. Column 1 regresses AREA on NILE FLOODS and its square. The sign of the coefficients are as expected, suggesting that intermediate floods are associated to a larger area. Nevertheless, standard errors are large and the coefficients are not significant. Column 2 controls in a flexible way for the number of years the pharaoh has been in office. We choose the model with a smallest BIC and introduce in the regression the contemporaneous and lagged values of TENURE and its square. These variables are not significant but since they decrease the value of the BIC, we maintain them in the remaining columns. Their inclusion also decreases the p-values of the

⁴⁷The data come from Geacron, <http://geacron.com/home-es/?lang=es>.

Dependent Variable: Log of Area controlled by the State							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
NILE FLOODS _t	0.063 (0.173)	0.059 (0.110)	0.066* (0.079)	0.095** (0.044)	0.125** (0.027)		0.142** (0.021)
NILE FLOODS _t ²	-0.008 (0.170)	-0.007 (0.113)	-0.008* (0.065)	-0.011** (0.036)	-0.013** (0.023)		-0.014** (0.021)
NILE EXTREME DROUGHT _t						-0.107** (0.020)	
NILE EXTREME FLOOD _t						-0.001 (0.969)	
RAINFALL HINTERLAND _t				0.203* (0.055)	0.170* (0.078)	0.187* (0.067)	0.193* (0.077)
RAINFALL HINTERLAND _{t-1}				-0.211* (0.053)	-0.185* (0.073)	-0.203* (0.067)	-0.202* (0.082)
TENURE _t		0.002 (0.442)	0.002 (0.474)	0.002 (0.567)	0.002 (0.463)	0.002 (0.484)	0.003 (0.384)
TENURE _t ²		-0.000 (0.251)	-0.000 (0.251)	-0.000 (0.267)	-0.000 (0.242)	-0.000 (0.273)	-0.000 (0.193)
TENURE _{t-1}		0.001 (0.706)	0.001 (0.777)	0.001 (0.773)	0.001 (0.679)	0.001 (0.671)	0.001 (0.419)
TENURE _{t-1} ²		-0.000 (0.605)	-0.000 (0.837)	-0.000 (0.817)	-0.000 (0.683)	-0.000 (0.586)	-0.000 (0.333)
TENURE DYNASTY _t			0.001 (0.228)	0.000 (0.439)	0.001 (0.262)	0.001 (0.221)	0.001 (0.181)
TENURE DYNASTY _t ²			-0.000 (0.223)	-0.000 (0.339)	-0.000 (0.238)	-0.000 (0.206)	-0.000 (0.174)
TENURE DYNASTY _{t-1}			-0.001 (0.215)	-0.001 (0.163)	-0.000 (0.800)	-0.000 (0.427)	-0.000 (0.988)
TENURE DYNASTY _{t-1} ²			0.000 (0.288)	0.000 (0.225)	0.000 (0.926)	0.000 (0.400)	-0.000 (0.918)
<i>t</i>					-0.001 (0.315)	-0.001 (0.324)	0.000 (0.972)
<i>t</i> ²					0.000 (0.205)	0.000 (0.235)	0.000 (0.538)
<i>c</i>	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	0.174 (0.117)	0.167** (0.047)	0.062 (0.609)	0.061 (0.630)	0.374* (0.053)	0.639** (0.013)	0.382** (0.046)
<i>R</i> ²	0.960	0.962	0.963	0.963	0.964	0.963	0.964
Obs	309	308	308	308	308	308	308
Lags Dep. variable	×	×	×	×	×	×	×
BIC	-85.52	-81.25	-62.25	-55.03	-47.53	-46.031	-35.99
Cumby-Huizinga test (p-val)	0.899	0.986	0.675	0.563	0.755	0.713	0.581
Period dummies	-	-	-	-	-	-	×

Table 4. AREA CONTROLLED BY THE STATE AND CLIMATE SHOCKS. Dependent variable is AREA. P-values based on HC standard errors are reported in brackets. The table reports the value of the Bayesian information criterion (BIC) as well as the p-value of the Cumby-Huizinga test of residual autocorrelation. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Nile variables, which are now close to being significant (p-values are .11). Column 3 is similar to column 2 but controls in a flexible way for TENURE DYNASTY (the number of years that the current dynasty has been in office) and its square. The inclusion of these variables improves the precision of the estimates and NILE FLOODS_t and its square are now significant at the 10%. We next consider the impact of RAINFALL HINTERLAND on the area under state control and column 4 reproduces the specification with the smallest BIC, one that incorporates the current and the lagged value of RAINFALL HINTERLAND. In line with the circumscription hypothesis, the coefficient of the first lag is negative, suggesting that wetter years in the hinterland are associated with future reductions in the area under control of state. On the other hand, the contemporaneous value of RAINFALL HINTERLAND has a positive and significant coefficient, indicating that higher contemporaneous rainfall has a positive effect on the area controlled by the state. This result is not at odds with the circumscription hypothesis, which only predicts the behavior of lagged rainfall values. This result seems reasonable since higher levels of rainfall might increase the value of the surrounding areas and at the same time, it could facilitate army incursions. Column 5 introduces a linear trend and its square which turn out to be insignificant, but since their introduction reduces the BIC both terms are left in the regression. The precision of the estimates improves but otherwise the results are very similar. Column 6 uses two dummy variables to capture the nonlinear association between Nile floods and the area under state control. These dummies capture extreme droughts and extreme floods (corresponding to the bottom and the top 5% observations, respectively). Both coefficients are negative but only the former is significant. Finally, column 7 introduces period dummies and shows that results are robust to considering within-period variation exclusively.

As our second proxy for government consumption we consider PYRAMIDS, (the log of) the volume of pyramid construction. Our hypothesis is that in periods when the tax revenue was large more resources could be devoted to pyramid construction. Thus, we expect intermediate Nile floods to be associated with higher tax revenues and therefore, more intense pyramid construction. Similarly, we expect lower rainfall in the hinterland in previous periods to be associated with larger pyramid volumes for the same reason. As in the case of AREA, the relationship between *contemporaneous* values of RAINFALL HINTERLAND and PYRAMIDS could be stronger than in previous tables. The reason is that the materials employed to build the pyramids (limestone, granite and minerals) were often found in the desert surrounding the Nile valley. Thus, favorable contemporaneous weather conditions in those areas could facilitate mining expeditions and, thus, could have a positive short-run effect on pyramid construction.

To test these hypotheses we use data on pyramid construction from Lehner (1997), who provides information on the total volume in cubic meters constructed over the lifetime of a pharaoh. An

obvious drawback of these data is their lack of time resolution, as only total volume by pharaoh is provided. This lack of resolution will surely lead to attenuation bias.

We code the pyramids variable as follows: first, we consider a different unit of analysis (10-year periods) to previous tables since, as discussed above, the time resolution of the background data is poor. Given that the average tenure of a pharaoh in our sample is 1.5 decades, it is reasonable to expect that our climate indicators will still be useful in predicting pyramid construction, despite the attenuation bias. Second, we calculate average construction per year by dividing the volume constructed over the lifetime of a pharaoh by the years he's been in office. Finally, we average over decades and compute the log of the resulting quantity. The resulting variable, denoted as PYRAMIDS, is depicted in Figure 1.

Table 5 presents our results. As mentioned before, the unit of analysis in this case is 10 year periods. In addition, since pyramid construction had stopped by the beginning of the New Kingdom, our analysis focuses on the period 2686 BC –1550 BC.⁴⁸ As in previous tables, we allow for the inclusion of lags of both the dependent and the independent variables. The number of lags to be included in the regression is selected using the BIC and only the preferred models are reported.

Column 1 regresses PYRAMIDS on NILE FLOODS_t, which has a positive and significant coefficient, indicating that higher Nile floods are associated to more pyramid construction. As in previous tables, column 2 adds to the specification in column 1 the square of NILE FLOODS_t, which now is not significant (and increases the BIC). Since there is no evidence non-linearities, the inclusion of the square term increases the BIC and makes the estimation of NILE FLOODS_t noisier, we do not include it in subsequent specifications (but see column 7 below for further evidence on non-linearities). Column 3 adds the contemporaneous level of RAINFALL HINTERLAND, which now has a positive and significant coefficient. The positive effect of wetter contemporaneous conditions in the neighbouring areas is not difficult to account for: higher contemporaneous rainfall surely facilitated the access to minerals and other construction materials that were typically found in the desert surrounding the Nile banks. Column 4 introduces the lag of the rainfall variable, which has a negative and significant coefficient, suggesting that higher rainfall decreases next period's pyramid construction.⁴⁹ This result parallels that found in Tables 1 to 3 and is consistent with the circumscription hypothesis. Columns 5 and 6 add TENURE and TENURE DYNASTY, respectively, to the specification, and the results are robust to including those variables. As in previous tables, column 7 introduces two dummy variables that capture instances

⁴⁸This period covers from the Old Kingdom up to the beginning of the New Kingdom, see Appendix B.3 for details about Egyptian's chronology.

⁴⁹Different combinations of lags of RAINFALL HINTERLAND were employed and the one reported in this table was the one preferred by the BIC.

Dependent Variable: Pyramids								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NILE FLOODS _t	0.124*	-0.420	0.114*	0.138*	0.123	0.134*		0.126
	(0.099)	(0.365)	(0.083)	(0.052)	(0.110)	(0.061)		(0.120)
NILE FLOODS _t ²		0.064						
		(0.296)						
NILE EXTREME FLOOD _t							1.196***	
							(0.002)	
NILE EXTREME DROUGHT _t							-0.314*	
							(0.080)	
RAINFALL HINTERLAND _t			0.198*	1.390*	1.377*	1.478**	1.174*	1.426*
			(0.089)	(0.058)	(0.070)	(0.047)	(0.076)	(0.061)
RAINFALL HINTERLAND _{t-1}				-1.157*	-1.128*	-1.242*	-0.932	-1.177*
				(0.070)	(0.095)	(0.055)	(0.116)	(0.073)
TENURE _t					0.022**			
					(0.039)			
TENURE _t ²					-0.000**			
					(0.016)			
TENURE DYNASTY _t						-0.002		
						(0.550)		
TENURE DYNASTY _t ²						0.000		
						(0.573)		
c	-0.373	0.694	-0.679*	-0.797*	-0.848**	-0.704	-0.244	-0.722
	(0.161)	(0.380)	(0.082)	(0.052)	(0.041)	(0.103)	(0.114)	(0.132)
R ²	0.759	0.762	0.776	0.790	0.801	0.790	0.824	0.794
Obs	113	113	113	113	113	113	113	113
Lags Dep. variable	×	×	×	×	×	×	×	×
Period dummies	-	-	-	-	-	-	-	×
BIC	239.331	242.6	235.7	233.17	236.1	242.2	218.0	244.9
Cumby-Huizinga test (p-val)	0.632	0.766	0.1706	0.1706	0.172	0.248	0.943	0.172

Table 5. PYRAMID CONSTRUCTION AND CLIMATE SHOCKS. Dependent variable is PYRAMIDS, i.e. the log of the volume of pyramid construction associated with a particular ruler. The unit of analysis is the decade and the sample covers the period 2686 BC–1550 BC. P-values based on HC standard errors are reported in brackets. All columns incorporate lagged values of the dependent variable (1 lag) chosen according to the BIC. The table reports the value of the Bayesian information criterion (BIC) as well as the p-value of the Cumby-Huizinga test of residual autocorrelation. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

of aberrant behaviour of the Nile river in the form of extreme floods and extreme droughts. In line with the results obtained in column 2 above, there is no evidence of a non-linear effect

of Nile floods. While droughts are associated with less pyramid construction, extreme floods are associated with more construction. These results are consistent with the linear specification employed in the other columns in this table. Finally, column 8 introduces period dummies and shows that results are also robust to considering within-period variation although the coefficient of NILE FLOODS, which remains very stable across specifications, is now less precisely estimated (p-value .12).

In summary, the results in this section are consistent with the importance of the circumscription hypothesis as they show that the area under state control and the volume of pyramid construction deteriorated in years of aberrant Nile floods (particularly so when periods were abnormally dry) as well as in times following wet conditions in the hinterland.

8. CONCLUSIONS

In this paper, we have presented a dynamic model of environmental circumscription in which productivity shocks in a core and a hinterland affect the effective level of circumscription for a farming population. In periods of beneficial weather in the core and poor weather in the hinterland, people are pulled towards the core, implying a greater population to tax, higher revenues and political stability. When the reverse situation holds, people are pushed towards the hinterland, tax revenues fall, and the ruling elite is more likely to be overthrown.

The predictions from the model are then applied to an empirical investigation of weather shocks and political stability in Ancient Egypt during 2686-740 BCE. Developing novel paleoclimatic proxies for Nile floods and hinterland rainfall, we show that political instability is highest when floods are either extremely high or extremely low. With a time lag, favorable conditions in the hinterland will cause political instability in the core, as predicted by the model.

In future research, we would like to explore whether our dynamic environmental circumscription model has external validity for other time periods and settings. For instance, it would be interesting to investigate further whether rainfall shocks might explain the frequent rises and collapses of different polities in Mesopotamia during antiquity or the sudden disappearance of the Maya culture on the Yucatan peninsula around 900 CE. Is it even possible that lessons from these historical cycles might have some bearing on the risk of political collapse in our contemporary societies? We hope to address these issues in future work.

9. REFERENCES

- Acemoglu, D. and J. Robinson (2006) *Economic Origins of Dictatorship and Democracy*. Cambridge University Press.
- Allen, R.C. (1997) "Agriculture and the Origins of the State in Ancient Egypt" *Explorations in Economic History* 34, 135-154.
- Anderson, R.W., N.D. Johnson and M. Koyama (2017) "Jewish Persecutions and Weather Shocks: 1100-1800" *Economic Journal* 127(602), 924-958.
- Bar-Matthews, M. et al (2003) "Sea-land isotopic relationships from planktonic foraminifera and speleothems in the Eastern Mediterranean region and their implications for paleorainfall during interglacial intervals" *Geochimica et Cosmochimica Acta*, 67(17), 3181-3199.
- Bar-Matthews, M. and A. Ayalot (2011) "Mid Holocene climate variations revealed by high-resolution speleothem records from Soreq Cave, Israel, and their correlation with cultural changes" *The Holocene*, 21(1), 163-171.
- Bell, B. (1971) "The dark ages in ancient history: The first dark age in Egypt" *American Journal of Archaeology* 75(1, Jan), 1-26.
- Benati, G., C. Guerriero, and F. Zaina (2019) "The Rise of Inclusive Political Institutions and Stronger Property Rights: Time Inconsistency vs Opacity" University of Bologna
- Bentzen, J., N. Kaarsen, and (2016) "Irrigation and Autocracy" *Journal of the European Economic Association*.
- Besley, T. and T. Persson (2009) "The origins of state capacity: property rights, taxation and politics" *American Economic Review*, 99 (4). pp. 1218-1244.
- Borcan, O., O. Olsson and L. Putterman "State History and Economic Development: Evidence from Six Millennia" *Journal of Economic Growth* 23(1), 1-40.
- Bronk Ramsey, C. et al (2010) "Radiocarbon-based Chronology for Dynastic Egypt" *Science* 328, June 18, 1554-1557.
- Burke, P.J and A. Leigh (2010) "Do Output Contractions Trigger Democratic Change?" *American Economic Journal: Macroeconomics*. 2(4), Oct, pp. 124-57.
- Butzer, K. (1976) *Early Hydraulic Civilization in Egypt*, University of Chicago Press.

Carleton, T. and S. Hsiang (2016) "Social and Economic Impacts of Climate" *Science* Vol 353, Issue 6304, 9 Sep.

Carneiro, R.L. (1970) "A Theory of the Origin of the State" *Science* 169, 733-738.

Chaney, E. (2013) "Revolt on the Nile: Economic Shocks, Religion, and Political Power" *Econometrica* 81(5), 2033-2053.

Cullen, H.M. et al (2000) "Climate change and the collapse of the Akkadian empire: Evidence from the deep sea" *Geology* 28, 379-382.

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

Dell, M. B.F. Jones, and B. Olken (2012) "Temperature Shocks and Economic Growth: Evidence from the Last Half Century" *American Economic Journal: Macroeconomics* 4(3), 66-95.

Diamond, J. (2005) *Collapse: How Societies Choose to Fail or Survive*

Fleitmann, D. et al (2003) "Holocene forcing of the Indian Monsoon recorded in a stalagmite from Southern Oman", *Science*, vol 300, June 13, 1737-1739.

Fleitmann, D. et al (2007) "Holocene ITCZ and Indian monsoon dynamics recorded in stalagmites from Oman and Yemen (Socotra)" *Quaternary Science Reviews* 26, 170-188.

Geacron (2017) <http://geacron.com/home-en/>

Giosan, L. et al (2012) "Fluvial landscapes of the Harappan civilization", *PNAS*, May 29, E1688-E1694.

Harper, K. (2017) *The Fate of Rome: Climate, Disease and the End of an Empire*. Princeton University Press.

Hassan, F. (1981) "Historical Nile floods and their implications for climatic change" *Science* 212, June 5, 1142-1145.

Hidalgo, D.F., S. Naidu, S. Nichter, N. Richardson (2010) "Economic Determinants of Land Invasions", *Review of Economics and Statistics*, vol 92(3), 505-523.

Hodell, D.A., J.S. Curtis and M. Brenner (1995) "Possible role of climate in the collapse of classic Maya civilization", *Nature*, vol 375 (1 June), 391-394.

Hsiang, S. (2016) "Climate Econometrics" *Annual Review of Resource Economics*, 8, 43-75

Hsiang, S. M. Burke, and E. Miguel (2013) "Quantifying the Influence of Climate on Human Conflict" *Science* 341, 13 Sep.

Johnson, N. and M. Koyama (2017) "States and Economic Growth: Capacity and Constraints" *Explorations in Economic History*, 2017, vol. 64, issue C, 1-20

Kaniewski, D. (2015) "Drought and societal collapse 3200 years ago in the Eastern Mediterranean: A review" *Climate Change*

Kennett, D.J. et al (2012) "Development and Disintegration of Maya Political Systems in Response to Climate Change" *Science* 338, 9 Nov, 788-791.

Lagerlöf, N. (2020) "Multiple Steady Statehood: The Roles of Productive and Extractive Capacities" York University, *mimeo*.

Lazarus, E., D.J. Lewis, J.H. Stock, and M. W. Watson, 2018, "HAR Inference: recommendation for practice", *Mimeo*.

Lehner, M. (1997) *The Complete Pyramids*, London: Thames & Hudson.

Mann, M. (1986) *The Sources of Social Power*, Vol 1, Cambridge University Press.

Manning, J. et al (2017) "Volcanic suppression of Nile summer flooding triggers revolt and constrains interstate conflict in ancient Egypt" *Nature Communications*, 8:900.

Mayshar, J., O. Moav, and L. Pascali (2020) "Cereals, Appropriability and Hierarchy" forthcoming in *Journal of Political Economy*.

Mayshar, J., O. Moav and Z. Neeman (2017) "Geography, Transparency and Institutions" *American Political Science Review* 111(3), 622-636.

Miguel, E., S. Satyanath, and E. Sergenti (2004) "Economic Shocks and Civil Conflict: An Instrumental Variables Approach" *Journal of Political Economy* 112(4), 725-753.

Neath, A. A. and Cavanaugh, J. E. (2012), "The Bayesian information criterion: background, derivation, and applications". *WIREs Comp Stat*, 4 pp.199-203.

Olson, M. (1993) "Dictatorship, Democracy, and Development" *American Political Science Review* 567-576.

Pollock, D.S.G (2000) "Trend estimation and de-trending via rational square-wave filters", *Journal of Econometrics* 99, p. 317-334.

Said, R. (1993) *The River Nile: Geology, Hydrology and Utilization*. Oxford: Pergamon Press.

- Sanchez de la Sierra, R. (2020) "On the Origin of States: Stationary Bandits and Taxation in Eastern Congo," *Journal of Political Economy*, 128.
- Schönholzer, D. (2019) "The Origin of the State: Incentive Compatible Extraction under Environmental Circumscription" UB Berkeley, *mimeo*.
- Scott, J.C. (2017) *Against the Grain: a deep history of the earliest States*. Yale University Press, New Haven.
- Shaw, I. (2000) *The Oxford History of Ancient Egypt*, Oxford University Press.
- Spencer, C.S. (2010) "Territorial Expansion and Primary State Formation" PNAS 107(16), 7119-7126.
- Tol, R.S.J and S. Wagner (2010) "Climate Change and Violent Conflict in Europe over the Last Millennium" *Climate Change* 99(1-2), 65-79.
- Weber (1917) *Politics as a Vocation*.
- Wilkinson, T. (2010) *The Rise and Fall of Ancient Egypt*. New York: Random House.
- Wittfogel, K. (1957) *Oriental Despotism; a Comparative Study of Total Power*, Yale University Press.
- Zang D., P. Brecke, H. F. Lee, Y. He, Jane Zhang (2007) "Global climate change, war, and population decline in recent human history" *Proceedings of the National Academy of Sciences* 104, pp. 19214-19219

APPENDIX A. THEORY: PROOFS AND EXTENSIONS

A.1. **Proofs.** This appendix 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_t^c or A_t^{c-1} or because it decreases A_t^h). Thus, $TR_t^0(w_t) > 0$. 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

$$(5) \quad \frac{\partial(1 - \eta(a))}{\partial w} = \frac{D_t[D_t^0(TR_t - \kappa) - D_t TR_t^0]}{2^{3-2}}$$

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

If $D_t^0 = 0$ (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 $D_t^0 < 0$ (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 $D_t^0 > 0$, (i.e., if $\lambda > 0$), then (5) is negative provided

$$\frac{D_t^0/D_t}{TR_t^0/(TR_t - \kappa)} < 1.$$

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,

$$\frac{D_t^0/D_t}{TR_t^0/(TR_t - \kappa)} = \frac{\lambda TR_t - \lambda \kappa}{\lambda TR_t + \chi_t} < 1,$$

since λ , κ and χ_t 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_t^c or A_{t-1}^c or because it decreases A_t^h). We consider here the case where $\lambda=0$ and discuss at the end the general case. Under the assumptions of Proposition 2 this derivative is given by

$$(6) \quad \frac{\partial E(P_t)}{\partial w} = \eta(a_t)^0 (TR_t - \kappa - a_t) + \eta(a_t) (TR_t^0 - a_t^0)$$

The first term in equation (6) is positive, since $\eta(a_t)^0$ 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 $TR_t^0 > a_t^0$. Notice that

$$(7) \quad TR_t^0 - a_t^0 = TR_t^0 \left(1 - \frac{d_t}{2(d_t^2 + d_t N R_t)^{1/2}} \right) > 0$$

The last inequality follows because $TR_t^0 > 0$ and $d_t N R_t$ is also positive, which implies that $d_t / 2(d_t^2 + d_t N R_t)^{1/2}$ 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 $\lambda = 0$ 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.

APPENDIX 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.

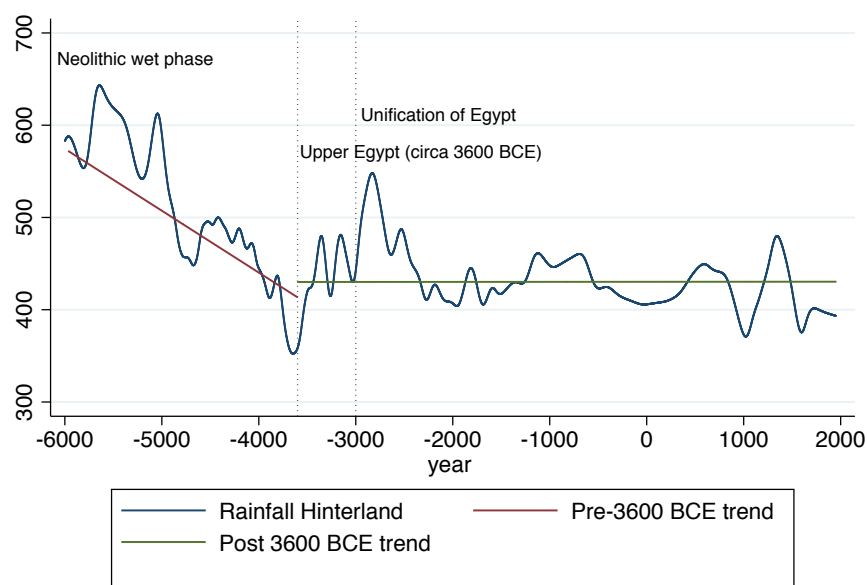


Figure B.1. NEOLITHIC WET PHASE AND EGYPTIAN UNIFICATION .

B.2. 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 Abbasid 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.⁵⁰ 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 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.3. 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 3rd 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

⁵⁰In 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.

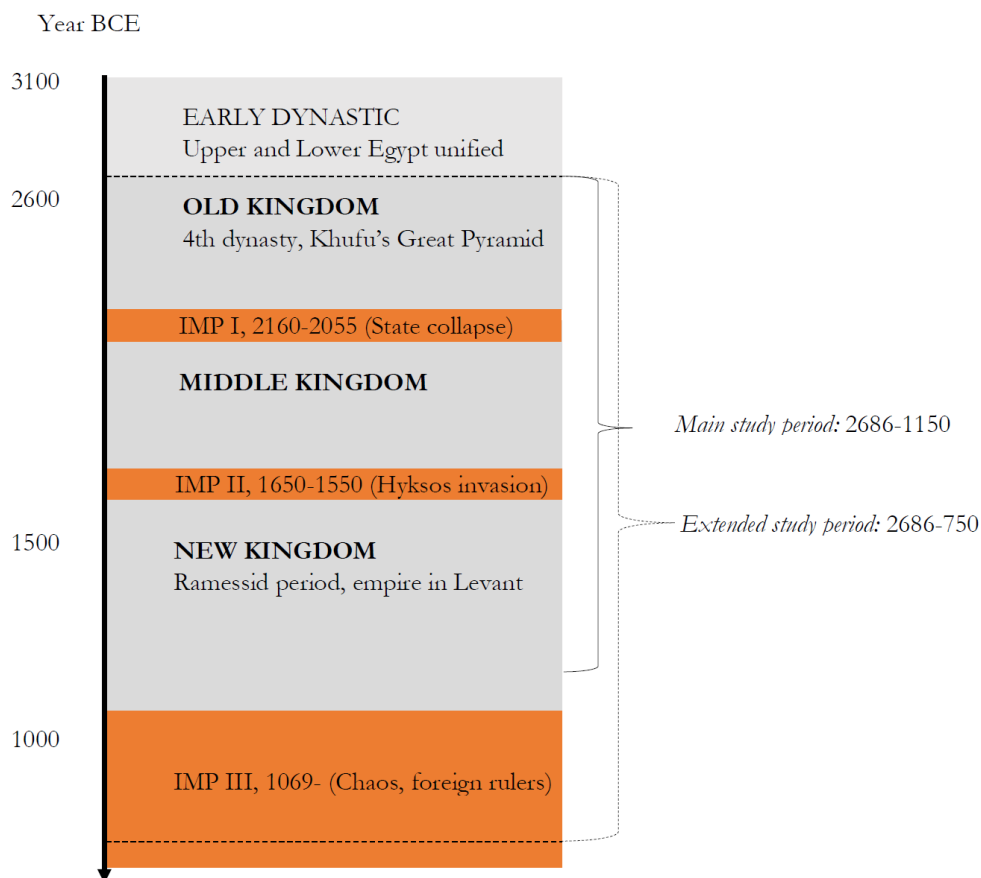


Figure B.2. ANCIENT EGYPT CHRONOLOGY. This graph depicts the main periods in Ancient Egypt's chronology. Source: Shaw (2000).

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 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).

In order to further check the robustness of our results in the empirical section, we also employ a second ruler tenure list from Department of Egyptian Art of the Metropolitan Museum.⁵¹ The basic tendency in the regressions remain intact when we use this variable instead.

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 last until about 1100 BCE.

B.4. 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.

APPENDIX C. 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.

C.1. Nile floods. As described in Section 5, our main proxy for Nile floods is a high resolution paleoclimatic archive ($\delta^{18}\text{O}$ 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

⁵¹See http://www.metmuseum.org/toah/hd/phar/hd_phar.html (October 2002).

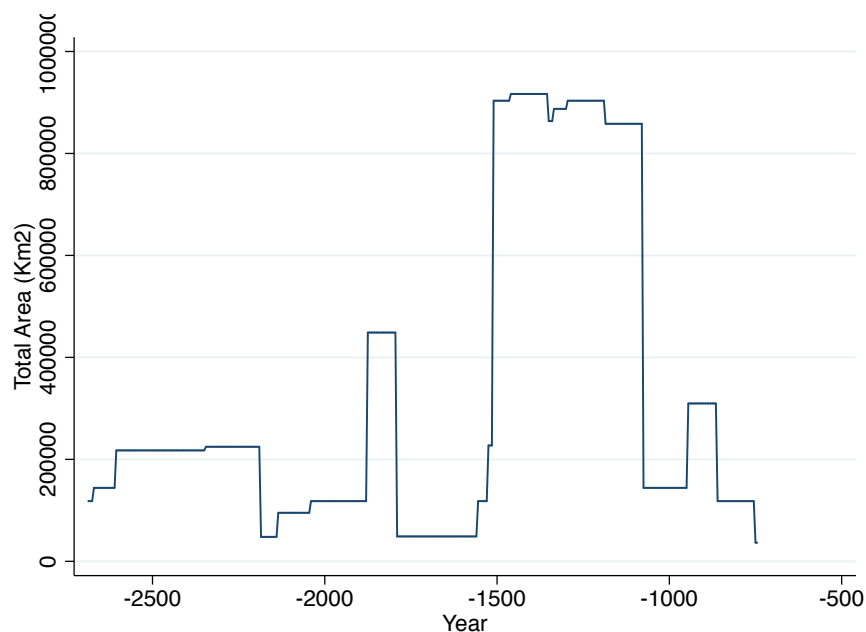


Figure B.3. AREA (IN KM²) OF ANCIENT EGYPT. The figure shows the total area controlled by the Egyptian state during antiquity. Source: Geacron

of the Intertropical Convergence Zone (ITCZ). Figure C.4 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.

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

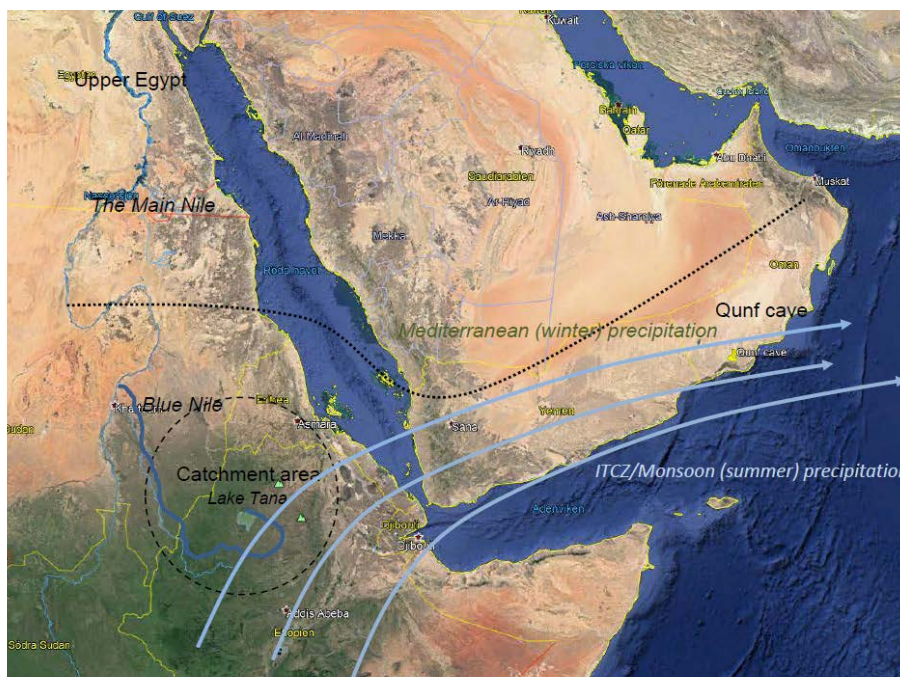


Figure C.4. INTERTROPICAL CONVERGENCE ZONE (ITCZ). This graph shows the location of Qunf's cave, the Nile catchment area in the Ethiopian Highlands, the Mediterranean belt of winter precipitation, and the typical reach of the Intertropical Convergence Zone (ITCZ) during summer months.

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 750 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.

We would expect to find a positive relationship between our Nile flood (Qunf) observations and the Lake Tana Ti concentration. Figure C.5a shows the fitted binscatter relationship, indicating a positive association, as expected. The correlation coefficient (0.63) in Table C.1, column (1), suggests a strong positive relationship, implying that more precipitation at Qunf is indeed associated with greater rainfall also in the Ethiopian highlands.

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 C.1. VALIDITY TESTS FOR NILE FLOODS. Note: 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

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 C.5b, 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 C.1 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 C.5c shows the binscatter for 6,805 interpolated, annual observations and Table C.1 shows that the correlation coefficient is almost 0.68.

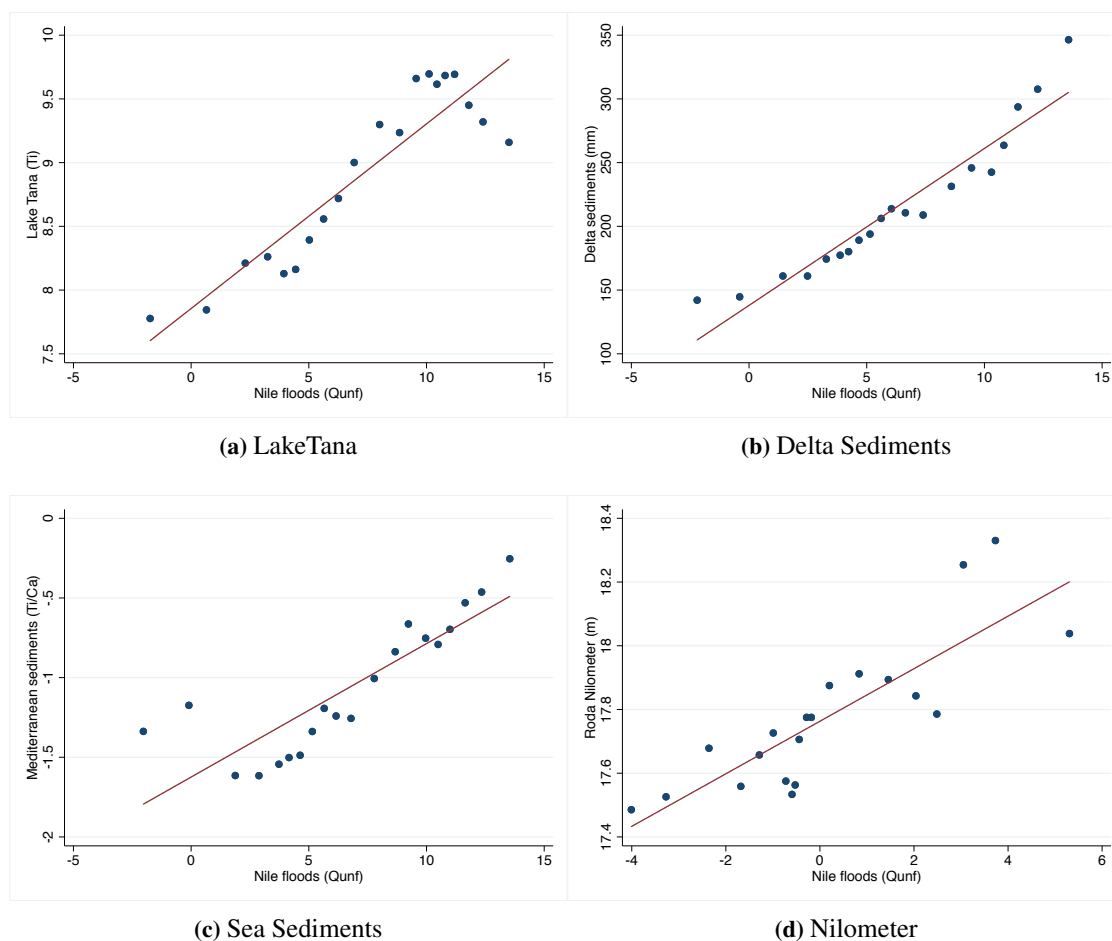


Figure C.5. BINSCLATTER PLOTS BETWEEN NILE FLOODS AND ALTERNATIVE PROXIES. Note: The figure shows the fitted binscatter relationship between our Nile floods (Qunf) index and various other climate indicators in the same order as specified in Table C.1.

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 C.5d shows the binscatter relationship and Table C.1, 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.

C.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 $\delta^{18}\text{O}$ -levels from Soreq cave, 18 km west of Jerusalem. Figure C.6 shows a map of the region, which also includes the location of the Roda Island Nilometer (in Cairo).



Figure C.6. SOREQ CAVE LOCATION

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 C.7a 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

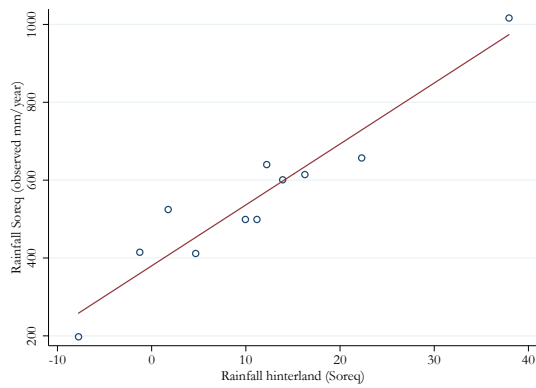
positive fit and the correlation coefficient in Table C.2, column (1), is 0.95, demonstrating that higher levels of our rainfall hinterland index are indeed associated with more observed precipitation.

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	$\delta^{13}C$	Si/Al	index
Observation period	2000,2010	1900,2012	-5000,-1150	-6686,-1150	-2685,-750
Observations	11	1,344	3,849	5,536	388

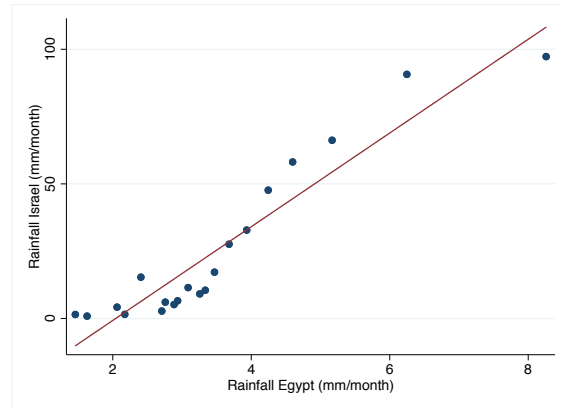
Table C.2. VALIDITY TESTS FOR RAINFALL HINTERLAND. Note: The table shows the Pearson correlation between our annual, filtered *Rainfall hinterland* (Soreq) index and combinations of various other climate indicators. Correlated variables in columns (1)-(2) are actual, annual and monthly observations whereas 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

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 C.7b, we show the binscatter of 1,344 monthly precipitation observations in Israel and Egypt over the time period 1901-2012 from World Bank (2019). The levels are very different in levels with Egyptian precipitation rarely exceeding 8 mm per month. The fitted regression line indicates a strong positive relationship and the correlation coefficient in Table C.2, column (2), is 0.82. The main take-away from the figure is that changes in rain levels in Israel also reflect changes in Mediterranean rains over Egypt.

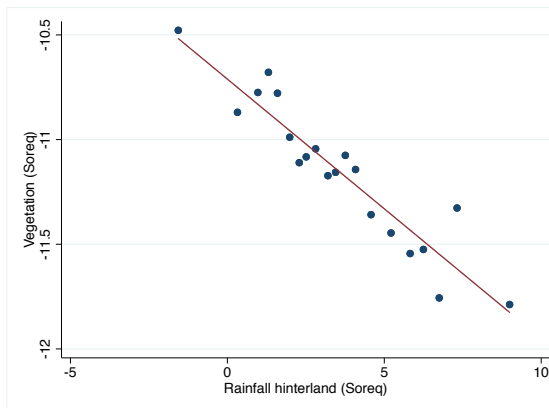
A third validation issue concerns the correspondence between the $\delta^{18}O$ isotope-levels at Soreq with other paleoclimatic proxies from the period. Bar-Matthews and Ayalot (2003, 2011) also obtained data from a second climate archive inside Soreq cave - the $\delta^{13}C$ isotope - often considered to be a proxy for levels of vegetation in the vicinity of the cave (lower levels of $\delta^{13}C$ indicate *more* vegetation). In Table C.2, column (3), and in Figure C.7c, we show that there is,



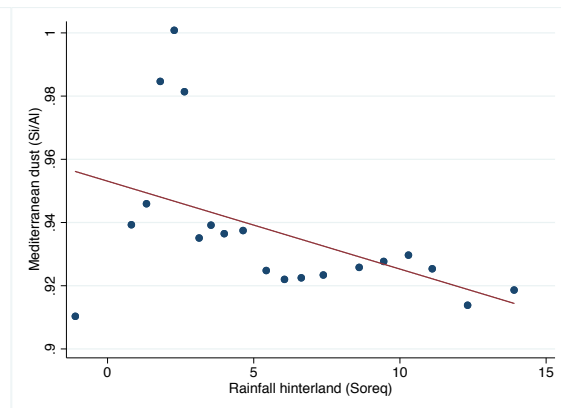
(a) Observed rainfall versus Rainfall Hinterland



(b) Observed Rainfall: Israel vs Egypt



(c) Vegetation vs Rainfall Hinterland



(d) Mediterranean Dust versus Rainfall Hinterland

Figure C.7. RAINFALL IN THE HINTERLAND VERSUS ALTERNATIVE PROXIES. Note: The figure shows the fitted relationship between our Rainfall hinterland (Soreq) index and observed rainfall outside Soreq cave in Figure C.7 (a), whereas Figures C.7 (b-d) show the binscatter relationship between variables in the order as specified in Table C2.

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 C.2, column (4), and in Figure C.7d, 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 C.7d 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 C.8 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 C.8 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 C.2, column (5) and Figure C.9 shows 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.

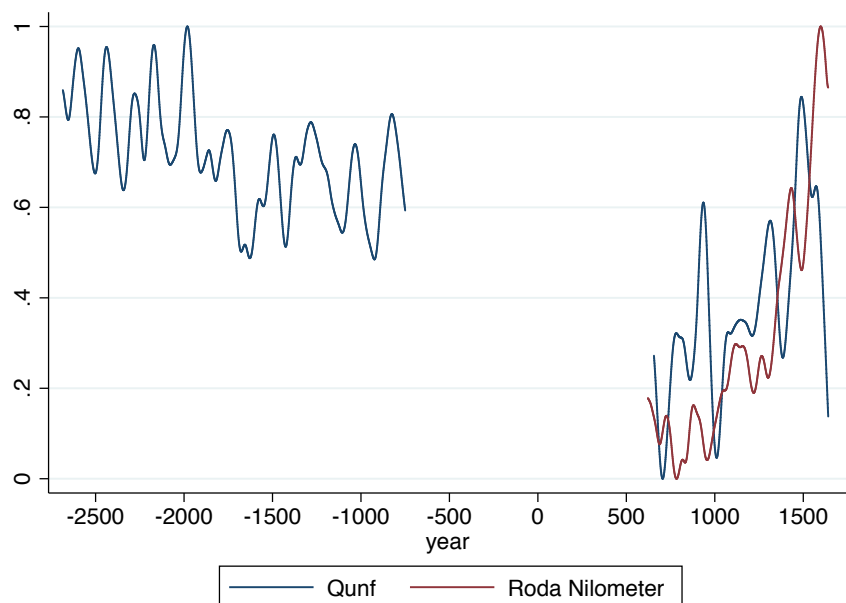


Figure C.8. NILE FLOODS (QUNF) VERSUS RODA NILOMETER DATA. Note: The figure shows the evolution of the Nile floods data obtained from the Qunf cave as well as direct measurements of the Nile river obtained from the Roda Nilometer. Both series have been filtered using a Butterworth index and to facilitate the comparison, the resulting variables have been normalized to be in the interval $[0,1]$.

APPENDIX D. ADDITIONAL INFORMATION

This section provides definitions of the variables employed in the empirical analysis (Section D.1) and a table of summary statistics (see Section D.2). Section D.3 describes a preliminary analysis that examines the time series properties of the variables employed in the empirical analysis. Section D.4 present additional results mentioned in the main text.

D.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];⁵³ 2 if any of the years in the five-year period $(t, t + 1)$

⁵³See 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),

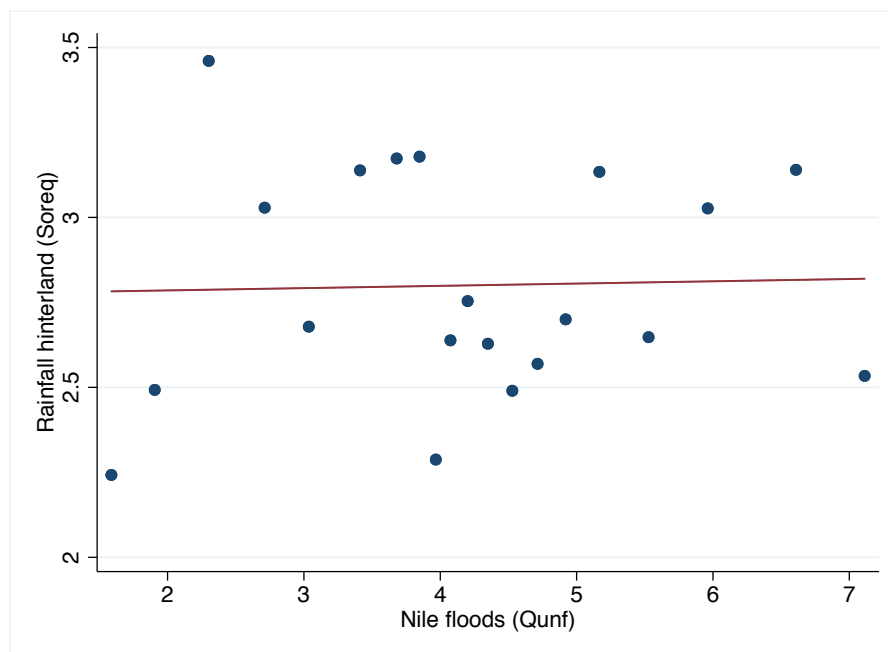


Figure C.9. NILE FLOODS VERSUS RAINFALL IN HINTERLAND. Note: The figure shows the fitted binscatter relationship between our Nile floods (Qunf) index and rainfall hinterland (Soreq) index as specified in Table C.2.

corresponds to periods of high political instability for which exact chronology is not available [around 17% of the periods],⁵⁴ and 0 otherwise.⁵⁵

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

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.

⁵⁴More 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).

⁵⁵Alternative 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".

and political instability, and 0 otherwise. See Appendix B1 for definitions and main historical sources of Egyptian dynasties.⁵⁶

AREA_{*t*}: Area under State Control. This variable captures the log of the area (in km²) under State control at the end of the five-year period starting in *t*. Source: Geacron.

PYRAMIDS_{*t*}: Log of the volume of pyramid construction averaged in 10 year periods. This variable measures the log of the volume of pyramid construction associated to the life of a pharaoh, divided in 10 year periods.

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 10) year periods. And third, we filter the resulting series using a Butterworth filter (see Pollock 2000, JoE, for details). Figure 2 (upper panel) plots the 5-year averaged Qunf data versus the filtered data (correlation is .85).

RAINFALL 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 2 (lower panel) plots the original versus 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.

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.

⁵⁶As 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.

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.

D.2. **Summary statistics.** Table D.3 presents summary statistics of the variables in our empirical analysis.

Table D.3. 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
NILE FLOODS (BK)	367	4.80	1.35	2.30	7.95
RAINFALL (BK)	367	3.20	1.38	0.95	7.11
AREA	390	327365.17	316089.89	36561.00	916795.00
PYRAMIDS	114	1.21	1.33	0	5.01
TENURE	390.00	14.24	15.43	1.00	93.00
TENURE DYNASTY	390.00	76.89	57.17	1.00	255.00

D.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 variables employed in the empirical analysis. More specifically we have applied DF-GLS unit-root tests, see Elliot et al. (1996). We were able to reject the unit root hypothesis in all cases.⁵⁷

D.4. **Additional Results.** Table D.4 reestimates Table 1 using an ordered logit specification, since our main dependent variable only takes three values.

⁵⁷Results are available upon request.

Dependent Variable: Pol Instability									
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
NILE FLOODS _t	-0.403*** (0.001)	-2.860*** (0.000)	-2.269*** (0.000)	-2.261*** (0.000)	-2.259*** (0.000)	-1.273* (0.072)	-1.124 (0.113)	-1.299* (0.066)	-1.148 (0.104)
NILE FLOODS _t ²		0.268*** (0.000)	0.205*** (0.002)	0.203*** (0.003)	0.203*** (0.003)	0.152** (0.050)	0.139* (0.073)	0.155** (0.046)	0.143* (0.065)
RAINFALL HINTERLAND _t				-0.053 (0.740)	-0.094 (0.826)	-2.922* (0.059)			
RAINFALL HINTERLAND _t ²					0.007 (0.919)				
RAINFALL HINTERLAND _{t-1}						3.244** (0.035)	0.341* (0.058)	-2.685* (0.088)	
RAINFALL HINTERLAND _{t-2}								3.050* (0.052)	0.385** (0.031)
<i>t</i>	0.024*** (0.000)	0.022*** (0.000)	0.023*** (0.000)	0.021** (0.040)	0.021** (0.041)	0.007 (0.549)	0.008 (0.469)	0.009 (0.434)	0.010 (0.375)
<i>t</i> ²	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.017)	-0.000*** (0.018)	-0.000 (0.723)	-0.000 (0.580)	-0.000 (0.596)	-0.000 (0.486)
TENURE _t			-0.136*** (0.000)	-0.137*** (0.000)	-0.137*** (0.000)	0.324*** (0.000)	0.322*** (0.000)	0.324*** (0.000)	0.323*** (0.000)
TENURE _t ²			0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	-0.004*** (0.000)	-0.004*** (0.000)	-0.004*** (0.000)	-0.004*** (0.000)
TENURE _{t-1}						-0.199*** (0.001)	-0.201*** (0.002)	-0.199*** (0.001)	-0.200*** (0.002)
TENURE _{t-1} ²						0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)
<i>R</i> ²	0.070	0.101	0.206	0.206	0.206	0.499	0.492	0.499	0.494
Obs	310	310	309	309	309	308	308	308	308
Lags Dep. variable (2 lags)	-	-	-	-	-	×	×	×	×
Infor. Crit (BIC)	590.0	576.7	522.4	528.04	533.76	380.31	378.50	380.01	377.54

Table D.4. POLITICAL INSTABILITY AND CLIMATE SHOCKS, ORDERED LOGIT SPECIFICATION.

Dependent variable is POL INSTABILITY. Estimation is by maximum likelihood in an ordered logit specification; p-values based on heterokedasticity consistent (HC) standard errors are reported in brackets. The period covered is 2686 BCE – 1140 BCE and the unit of analysis is 5 year periods. The table reports the value of the Bayesian information criterion (BIC) corresponding to each of the models considered. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure D.10 reports the marginal effect of NILE FLOODS on POL INSTABILITY, which has been computed using the coefficients in column 9 from Table 1.

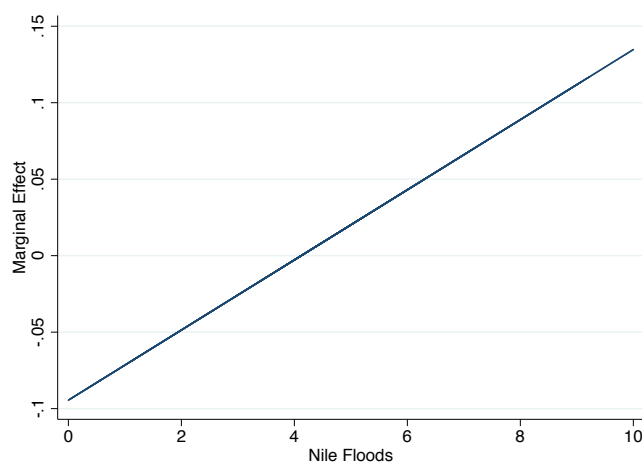


Figure D.10. MARGINAL EFFECT OF NILE FLOODS ON POL INSTABILITY

This graph shows the marginal effect of NILE FLOODS on POL INSTABILITY. The graph shows that increases in NILE FLOODS have a negative effect on political instability for moderate values of NILE FLOODS but the effect becomes positive for large values of NILE FLOODS. Estimates correspond to column 9, Table 1.

APPENDIX E. ADDITIONAL REFERENCES

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.

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.

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.

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