Floods, Droughts, and Environmental Circumscription in Early State Development: The Case of Ancient Egypt [1]

LAURA MAYORAL
Instituto de Análisis Económico (CSIC) and Barcelona School of Economics

OLA OLSSON
University of Gothenburg

July 2022

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 weather-related productivity shocks for early state development in ancient Egypt. We present a framework of extractive state consolidation predicting that political stability should be high 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 such periods, the elite can impose high levels of taxation that the population will be forced to accept as exit to the hinterland is not a feasible option. In order to test this hypothesis, we develop novel proxies for both the historical productivity of the Nile banks and of the Egyptian hinterland on the basis of high-resolution paleoclimate archives. Our empirical analysis then investigates the relationship between these proxies for environmental circumscription and political outcomes such as ruler and dynastic tenure durations and the area under state control during 2685 - 760 BCE. Our results show that while extreme Nile floods are associated with a greater degree of political instability, periods with a greater rainfall in the hinterland (i.e. a lower effective environmental circumscription) causes a decline in state capacity and a delayed increase in political instability.

[1] The paper was previously circulated as "Pharaoh's Cage: Environmental Circumscription and Appropriability in Early State Development". We are grateful for comments from Ingela Alger, Paolo Buonanno, Giacomo Benati, Joan Esteban, James Fenske, Bishnupriya Gupta, Debraj Ray, David Schönholzer, and seminar participants in Barcelona, Bilbao, Gothenburg, NES Moscow, Santiago de Chile, IAST Toulouse, ASWEDE workshop in Uppsala, Oslo and Warwick. Mayoral gratefully acknowledges financial support from the Generalitat de Catalunya, 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 and stability 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 early state formation but *early state political stability.*

We propose a framework of early 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 exiting from the core 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, 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. As a result, the tax collection will increase and state capacity will be reinforced. In times when the hinterland instead is relatively productive, an

[2]Our model is however open to a wider interpretation. For instance, the non-taxable activity could be growing tubers, as in Mayshar et al (2022), 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 (2022), 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.

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.

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 two 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 in 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 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.8), makes the Egyptian territory an ideal case for testing a dynamic circumscription hypothesis. As proxy variables for historical Nile inundations and hinterland rainfall, we introduce two paleoclimatological data sources (intertemporal variations in 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 strat-

[3]More 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.

[4]Unlike 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.

egy exploits the impact of two exogenous and time-varying orthogonal shocks on the degree of effective circumscription in Egypt.

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. The notion that a longer ruler tenure indicates a higher level of political stability, follows recent work comparing historical developments in the Christian and Muslim worlds (Blaydes and Chaney, 2013). Our *state capacity* variable captures 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 non-extreme Nile floods and dry conditions in the hinterland. The size of both effects is large and significant, 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 twofold. First, in order to capture the evolution of early state development in a controlled environment, we introduce new data on paleoclimatic shocks, political instability and state capacity for the Egyptian state covering almost 2 millenia. Second, we test the predictions from our model of dynamic environmental circumscription by exploiting two exogenous time-varying and independent proxies for productivity in the core and in the hinterland. We then show that political stability and strong state capacity are associated with well-behaved Nile floods (i.e., not too excessive or too scarce) and high agricultural productivity in the core, and with dry conditions and relatively low land productivity in the hinterland.

Our work is related to several important strands of research in the literature. Our paper is most strongly related to Schönholzer (2020) who studies the importance of environmental circumscription for the emergence of early states in a global cross-section of grid cells. The key sources of variation is the difference of land quality for agriculture in a particular cell compared to its neighboring cells.[5] Schönholzer finds strong cross-sectional evidence that environmental circumscription is an important determinant of state emergence. Our approach instead focuses on the time variation within a single well-documented environment.

Another important recent research is Allen et al (2020), who exploit random shifts to the course of the Euphrates and the Tigris in order to understand the dynamics of state formation in ancient Mesopotamia. Based on archaeological evidence, the authors find that (city) state formation is more likely to occur following the divergence of a river *away* from a particular site. On such an occasion, the necessity for coordinated irrigation creates a *demand* for a state to arise. In contrast, our paper and Schönholzer (2020) might be described as focusing on the rise or consolidation of states as a result of an *opportunity* to tax a concentrated population.

Scott (2009, 2017) discusses many examples of historical *state evasion* by marginal populations who prefer to exit from the core rather than being *domesticated* and dominated by a ruling elite. Further, Mayshar et al (2017, 2022) emphasize the crucial importance of the transparency and

[5] This paper appeared more or less simultaneously and independently of our paper. Schönholzer also studies this relationship in a sample of archaeological sites and when only comparing river valleys to each other.

appropriability of the main taxable resource for a strong state to arise. Our conceptual framework integrates these important insights. In addition, 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.[6] 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 extension 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 *social cage* in an underpopulated world (Mann, 1986).

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 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

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

[7] Diamond (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.

[8] Such 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 a, 2012).

[9] Influential general overviews include Butzer (1976), Shaw (2000), Wilkinson (2010) and Moreno Garcia (2020).

[10] See Appendix B.2 for a more extensive account of these works and how they relate to our research design.

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.

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 describe our conceptual framework and lay out the key testable hypotheses. A formal model capturing these ideas is presented in Appendix A.1. 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.2 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 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.[11] 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 (2022) argue that historical political entities in territories relying on cereals typically saw the emergence of more stratified states earlier than in territories where tubers and roots were the staple crop of the population. The key reason for this regularity was the greater *appropriability* for taxation of cereals compared to tubers. A similar analysis is offered by Scott (2017).

Two recent papers focus on early state development in Mesopotamia. Allen et al (2020) exploit random shifts to the course of Euphrates and Tigris and show that city state formation is more likely to emerge when a dependable source of water diverges away from a site. The authors argue

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

that this tendency should be understood as a response to the demand for coordinated irrigation facilities in already established communities. Benati and Guerriero (2021) argue that a state was formed in Mesopotamia when adverse production conditions pushed the existing elite to establish a state by giving strong political and property rights to non-elite groups that possessed complementary skills.

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

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

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

The theory of the importance of environmental circumscription for original state formation can easily be extended into a dynamic model of early state development.[12] Effective circumscription

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

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

3 Early State Consolidation: A Conceptual Framework

In this section, we outline an intuitive conceptual framework that is intended to give structure to the empirical research design below. For a complete formal model along the same lines, please see Appendix A.1. Our framework is inspired by the work of Carneiro (1970) and Allen (1997), and shares some features with Schönholzer (2020). These papers analyze how the degree of environmental circumscription in a core and hinterland affects the probability of *state formation*. However, unlike these papers, we focus on the impact of the *intertemporal variation* in levels of circumscription on the stability and the capacity of the state. In this sense, our emphasis is placed on *state consolidation* rather than on original state formation.

We start by assuming a given territory with a highly productive *core* (c) and a less productive *hinterland* (h). In the case of Egypt, the core might be thought of as the fertile lands directly adjacent to the Nile whereas the hinterland are the less fertile areas to the east and west. A state has been formed in the core and the distinctive feature of this early state is that its ruling elite (the pharaoh and his enablers) have the capacity to levy taxes on the sedentary farming population residing in the core. The ruling elite also has a capacity to use these tax resources to produce large-scale monuments and defensive fortifications. Taxes cannot be levied on the population in the less accessible hinterland and the elite only has a limited capacity to constrain people from evading the state by exiting to the hinterland.

There is a total population of fixed size that at time t either lives in the core or in the hinterland. For simplicity, we consider that there are just two time periods, $t = \{1, 2\}$ and the world ends after period 2. Although artificial, this two-period approach allows us to highlight the main points of interest avoiding unnecessary complications. In the core, agricultural output is produced, whereas the hinterland is primarily suitable for activities like pastoralism and hunting game. We denote the time-varying (average) land productivity in the core as A_t^c and (average) productivity in the hinterland as A_t^h, where $A_t^c > A_t^h$ always holds for $t = \{1, 2\}$. Average productivity in both regions depends on annual weather variability so that A_t^j will be high (low) whenever weather in t is good (bad) in $j = \{c, h\}$. In Egypt, good weather in the core might be thought of as abundant but not extreme Nile floods that inundated the irrigated fields. Good weather in the hinterland might be thought of as abundant local rainfall that increased biodiversity and the vegetational cover. Weather variability within an area is typically autocorrelated but we assume that variations between the core and in the hinterland are independent and mostly uncorrelated, since they are the outcome of different weather systems. We will discuss these assumptions of our framework in some detail in the empirical section below.

The key feature of our dynamic environmental circumscription model is the interaction between the levels of A_t^c and A_t^h. The degree of environmental circumscription at time *t* is simply defined by the comparative level of $A_t^c/A_t^h(> 1)$. The larger the value of this ratio, the higher the degree of environmental circumscription and vice versa. Thus, all else equal, we propose that if A_1^c is high, a greater share of the population will be attracted to be settled in the core in the second period, as weather shocks are autocorrelated.[13] Conversely, when productivity is relatively high in the hinterland, a lower share of the total population will choose to reside in the core in period 2 and instead relocate to the hinterland.

But why would anyone choose to live in the hinterland when conditions for producing food are always better in the core? The main reason, which is emphasized for instance by Carneiro (1970) and Scott (2009), is the fact that the core is dominated by an elite that extracts taxes from the population there. Such taxes might take the form of corvée labor for public projects or direct taxes on agricultural production. The nature and intensity of taxation typically depended critically on the transparency and appropriability of agricultural output, as emphasized by Mayshar et al (2017, 2022). In our model of early state development, we simply assume that the elite has the capacity to tax the population in the core but not in the hinterland. Hence, due to taxation, some parts of the population with a low marginal product might be better off in the hinterland than in the core. It follows that the larger the productivity gap between the core and the hinterland at time 1, the greater the share of the population that will be attracted to the core, and vice versa. At each point in time, there is a part of the population that actively considers whether it is best to remain in the core or to exit to the hinterland.

The assumption above of an exit option in the hinterland, is true to Carneiro's (1970) original model. We recognize however that also another scenario might have existed in early state development: That in times of deficient floods in core and relative good conditions in the hinterland, some marginal populations might remain based in the core but actively attempt to evade the state by subsisting on non-transparent or non-taxable food stuffs and by sometimes taking temporary refuge in the hinterland from the government's tax officers (Scott, 2009). When Nile levels are good and the hinterland very dry, such *state evasion* is less likely.

We assume that the total amount of taxes collected by the elite depends on two factors: The size of the population producing crops and paying taxes in the core and the weather shocks that determine agricultural productivity in the core. Positive weather shocks at time 1 in the core increase tax revenues in two ways: Directly, by immediately increasing the level of taxable agricultural yields in time 1, and indirectly, by attracting a greater population to the core that accept to pay taxes. The latter, indirect mechanism is likely to act more slowly than the immediate productivity

[13] More specifically, in the model we assume that individuals differ in ψ_i, the individual relative productivity in the core vs the the hinterland. Thus, everything else equal, an increase in A_1^c will make the core more attractive than the hinterland for a larger number of people.

improvement. Thus, to capture this delayed effect, we assume that population displacement or revolts happen in the second period. Beneficial weather shocks in the hinterland at $t = 1$ will, ceteris paribus, decrease the tax-paying population in the core at $t = 2$ and, thus, decrease total taxes collected in the second period.

Total tax revenues collected from farmers are used by the ruler for public goods provision but also for elite private (conspicuous) consumption such as pyramids, festivals and luxury goods obtained by trade. In an extended model in Appendix A, we differentiate between elite consumption and investments in coercive/defensive capacity. Here, we simply assume that both types of goods increase *state capacity* and hence makes it more likely for the ruler and his dynasty to remain in power. Conversely, if total tax revenues fall, state capacity will weaken and increase the risk of internal insurrection or an external attack that might dethrone the ruler and his dynasty.

There is ample support in the literature that economic downturns (that in our framework translate into a reduction in the tax collection) increase the probability of conflict and political instability (Acemoglu and Robinson 2006, Burke and Leigh 2010, Chaney 2013, etc.). Periods of economic stress facilitate the organization of internal revolts, as factions of the elite can take advantage of the weakness of the state (and thus, the smaller size of mobilized defensive resources) and the discontent in the population (due to scarcity and famine) to organize an insurrection and replace the leader. This is the so-called "opportunity cost" effect. An alternative scenario would be one in which the aggressor is primarily interested in robbing the state, in which case the probability of attack and the amount of collected taxes will be positively related (the so-called "rapacity" effect, see Dube and Vargas, 2013).

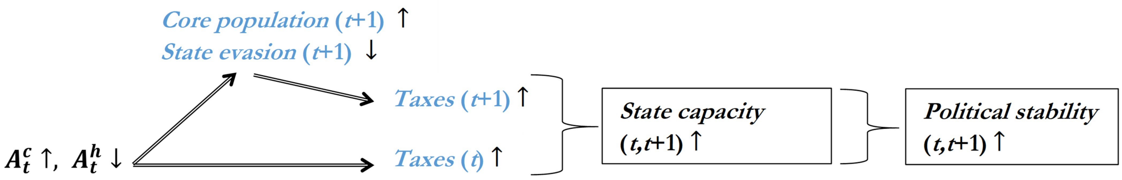
In the case of Egypt, we argue that a negative correlation between total tax collection and the probability of attack is the most relevant situation. The pharaoh was believed to be responsible for the Nile floods. Then, in case of famines derived from extreme Nile floods, he was often directly blamed for them (Bell, 1971). This could have led to internal revolts, whereby factions of the elite could have taken advantage of the discontent in the population and/or of the weakness of the army to organize an insurrection to depose the ruler. In addition, the classical period in ancient Egypt offers an object of study that is as close to an isolated and undisturbed "historical lab" as one can find, where the lack of neighboring powers that could pose a serious threat to the stability of the Egyptian state was the general rule.[14]

The proposed causal mechanism from the framework above is summarized in Figure 1 below. An increase in environmental circumscription will have a direct positive effect on production and tax levels at time *t*. If A_t^c rises and/or A_t^h falls, so that effective environmental circumscription rises, then the tax-paying population in the core will increase in the next period and state evasion tendencies will be low, indirectly boosting total tax revenues in *t*+1 and state capacity. A higher state capacity should increase political stability by making successful insurrections less likely. The reverse happens if A_t^c falls and A_t^h rises. In our empirical analysis, we have proxy data for A_t^c, A_t^h, state capacity (military territorial extension), and political instability (ruler and dynastic

[14] As mentioned earlier, the Hyksos, who ruled Egypt around 1600 BCE, was an important exception to the general isolation of Egypt from outside intervention.

turnover), whereas population, state evasion and tax levels (in blue) are unobserved.

Figure 1: HYPOTHESIZED CAUSAL RELATIONSHIPS



NOTE: The figure is a summary of the hypothesized causal relationships, from left to right, in the theoretical framework above. Factors in blue are unobserved empirically.

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 3600 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.

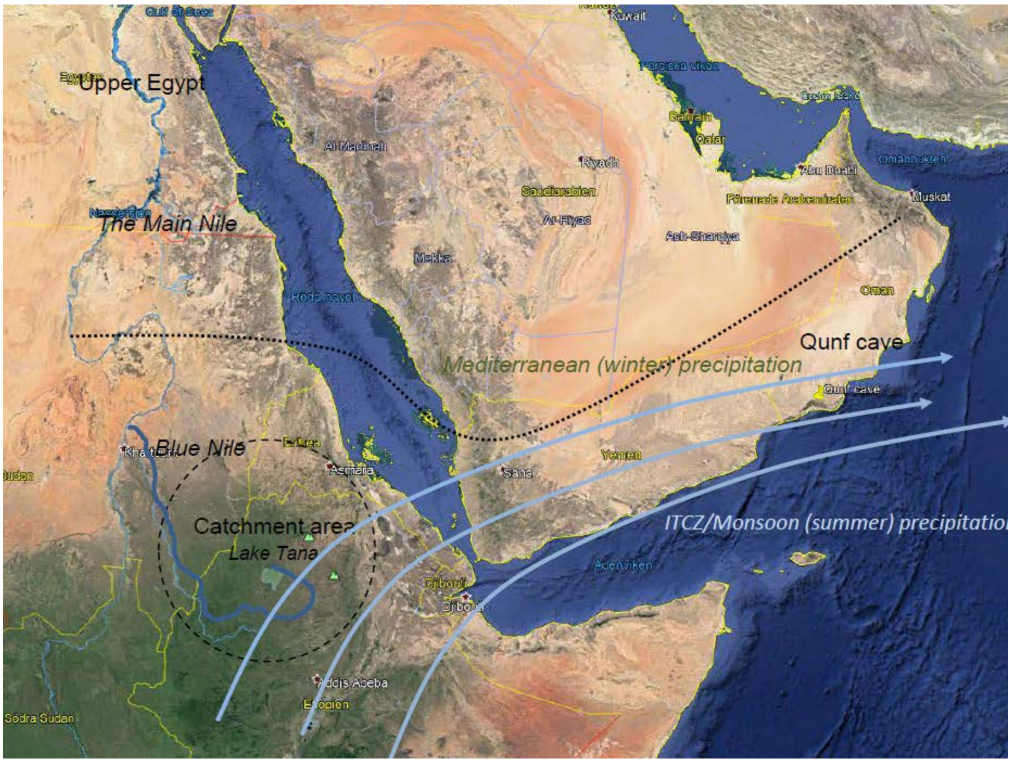
4.2 Weather shocks and agricultural production

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. Abnornally 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, as shown in Figure 2. 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).

Figure 2: INTERTROPICAL CONVERGENCE ZONE (ITCZ)



NOTE: This figure shows the location of Qunf 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.

Agricultural output relied heavily on the extent of Nile floods. Deviations from optimal inundation levels had a severe impact on land productivity. Aberrant floods could occur for two reasons. On the one hand, when floods were insufficient, the land lacked the much-needed water and nutrients. On the other hand, excessive floods could damage the complex irrigation networks that were essential to ensure efficient distribution of the Nile waters.

There are many historical accounts of how periods of 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.

An important and well-known characteristic of Nile river floods (which is also shared by other climate-related variables) is the fact that they are very persistent. More specifically, Nile floods are characterized by long cycles where prolonged periods of abundant floods are followed by lengthy periods of deficient floods (Hurst, 1951). This highly persistent correlation pattern is important in our argument. A single year of aberrant floods possibly wouldn't be enough to severely weaken state capacity and/or trigger migration pressures. However, the existence of extended periods where copious or scarce floods tend to cluster together could surely have a big impact on the viability and consolidation of the Egyptian state.

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 (2022), 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).

From early dynastic times, the ruler had access to corvée labor who worked at centrally organized production centers where crops, textiles, wine and other necessities were produced by a specialized work force (Moreno Garcia, 2020). 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 eventually necessitated the permanent presence of a military organization taking orders from the pharaoh. Corvée labor was also used 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).

5 Data

In this section we describe the main dependent and independent variables employed in our empirical exercise: the political stability and state 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 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 – 760 BCE, which covers the Old, Middle, and New Kingdoms, as well as the first and second intermediate periods (see Figure B.2 in Appendix B 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).[15]

We code our main dependent variable, POL INSTABILITY, as follows. First, we divide the sample period (2685-760 BCE) into 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 is a period of no central rule (18% of the observations). The remaining periods (61.4%) are considered as politically "stable" and are coded as 0. 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, variations of this basic definition, including a binary version of it, 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'.[16]

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. 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

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

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

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 our sample period into 5-year intervals and assign a value equal to 1 if there is a dynastic change in any of the years of the period (around 3% 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 a proxy for state capacity, we use a variable measuring the geographical extent of the Egyptian state. 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). [17]

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 two sources of data on intertemporal variation in rainfall.

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.

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 2 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

[17] It 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.

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 2. The time series at Qunf is derived from speleothem data, measuring the oxygen isotope content (δ18O) in a stalagmite from the cave. In the scientific literature, δ18O-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 in 760 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 δ18O-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 δ18O-time series has not in any way been contaminated or manipulated by man, either consciously or by a change of measurement over time.[18]

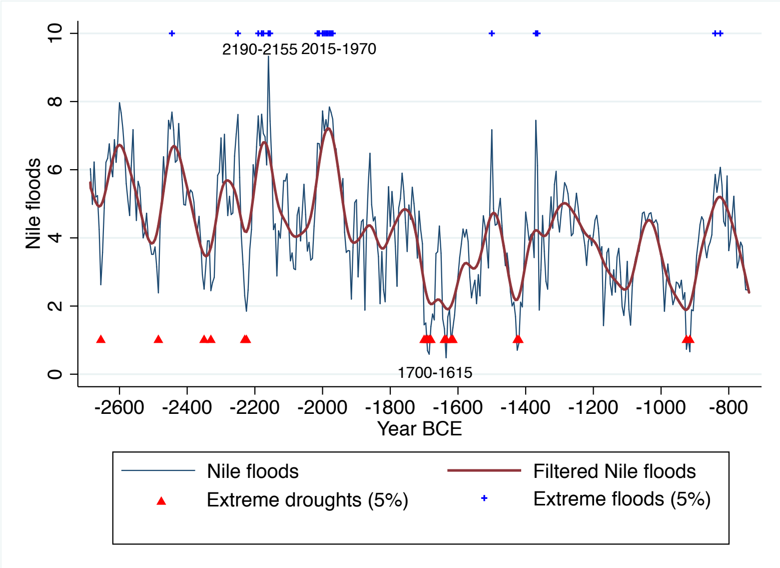
Our validation exercise compares our proxy for Nile inundations to other indirect proxies such as sediment levels in the Nile delta, in lakes in the Ethiopian highlands and in the Mediterranean. We have also compared our measure to direct measurements of Nile inundations (i.e., Nilometer records). Although 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 3 shows the *Nile floods* index for the 2685-760 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

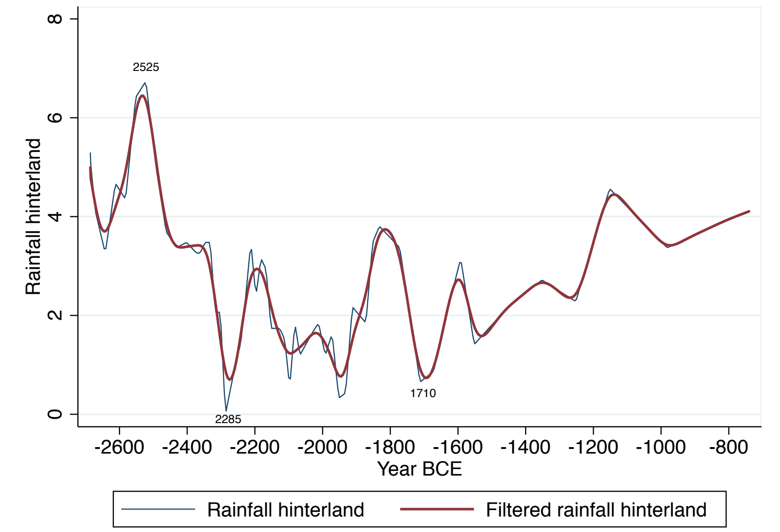
[18] We also discuss other issues regarding the use of Nilometers in the Appendix.

Figure 3: CLIMATE PROXIES: NILE FLOODS AND RAIN HINTERLAND.

(a) Nile Floods (Qunf data)



(b) Rain hinterland (Soreq data)



NOTES: 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.

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 the Nile floods-index implies that Nile inundations gradually diminished during the classical period.

As mentioned before (see Section 4.2), 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 $\delta 18O$ 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.5 in Appendix C.2 for a map of the area). Appendix C provides a validation 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 – 760 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 3 shows the *Rain hinterland*-index based on the data from Soreq. Also here, we see a declining trend in precitation 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, consistent with our prediction, 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.8 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, 2685-760 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, the key features of our empirical strategy, the empirical specifications and estimation techniques employed and a brief preliminary analysis of the data.

6.1 Sample

Our empirical analysis examines the evolution of political instability and government capacity in ancient Egypt over the period 2685 BCE – 760 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.[19] The unit of analysis in our baseline regressions is five year-periods. The reason why we aggregate the data in such a way is twofold. Firstly, the type of mechanisms we have in mind (i.e., state weakening and increasing migrating pressures) are unlikely to follow after a single year of low agricultural productivity. Instead, they are more likely to arise in *periods* with repeated low harvests. We capture this by considering average Nile behaviour in a given time period (5 years in our baseline specification, but other frequencies have also been considered, see Table 2). Secondly, given the antiquity of our data the precision in the dating of our variables is an important concern. By considering 5 (or, in some regressions, 10) year periods we can reduce, at least partly, measurement error due to inexact dating. This, of course, would not eliminate entirely measurement error which most likely will lead to attenuation bias.[20]

6.2 Empirical strategy, Estimation and Model Selection

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.[21] 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 exogeneous 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

[19]The 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.

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

[21]As 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.

outcomes can be attributed to the effect of the treatment variables. Admittedly, this can be a strong assumption as well, as we cannot discard that unobserved variables might create differences across periods. Nevertheless, the fact that we are using a very different research design and data, complements very well the existing literature.

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 (sequence of) bad harvests in the core has a direct effect on the tax collection of that period but also an indirect one on future collection, that operates through the decision of farmers to migrate to/from the core in the following period. This decision, in turn, is made based on beliefs about the evolution of climate. These beliefs might need some time to form 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.

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. Focusing on the low-frequency component of the climate data is important for our purposes as the type of effects we hypothesize (i.e., severe weakening of the state and credible migration threats) are unlikely to follow a single year of aberrant floods but rather, they would be the outcome a sequence of negative weather shocks. 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, some of our specifications contain time trends as well as period dummies.

Our main dependent variables measuring political instability are discrete. For this reason, we employ non-linear specifications (ordered logit and logit) estimated by maximum likelihood and, for robustness, also linear ones. Both types of specifications have pros and cons, so by using both we hope to be able to overcome some. A key characteristic of our data is a high degree of autocorrelation. Linear models provide a more robust and flexible framework to test and correct for autocorrelation in the residual term than non-linear ones. This is an important advantage, as residual autocorrelation can provoke biases in the estimators when dynamic time series models are considered. Also, since our models contain squared terms and sometimes interactions as well, interpretation of results is simpler in linear models (Ai and Norton, 2003). Nevertheless, since our dependent variables are often ordinal indices, OLS is also problematic, as it interprets literally the assigned (arbitrary) values defining the different categories. For this reason, in our baseline specifications we use non-linear models but we also report estimates obtained in linear specifications, which will be helpful in interpreting the results and in testing for residual autocorrelation.

To capture dynamic effects, we allow for lags of the dependent and the independent variables in the model. As is customary in time series analysis lags of the dependent variable are chosen using information criteria. In particular, we use the Bayesian information criterion (BIC). [22] For each model, we report the Cumby-Huizinga test for (first order) autocorrelation. Given the difficulties of testing for residual autocorrelation in non-linear specifications, when the latter models are employed, autocorrelations tests are computed on linear specifications containing identical variables as the non-linear ones.

We also allow for a number of controls. In the first table we use the BIC to choose a model specification for them and, to facilitate comparability across equations, we maintain that specification in all subsequent tables. More specifically, for each control we consider the variable in levels, its square and 1 lag of these variables, see Table 1 for details.

Table D.4 in Appendix D presents the results of unit root tests applied on our dependent variables. We are able to reject the null hypothesis of a unit root in all of them, except for the case of AREA, the log of the area under state control. For this reason, in the empirical analysis we model this variable in first differences, see Section 7.2.

7 Results

This section describes our 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, and the area state control) and weather shocks. Variable definitions and summary statistics are provided in Appendix D.

7.1 Political Instability

Tables 1 and 2 present our main results relating political instability and the proxies for the weather conditions in the Nile and surrounding areas. In all cases, p-values based on robust standard errors are reported in brackets.

Table 1 focuses on POL INSTABILITY. Ordered logit models are estimated in all columns except in Column 8, where a linear specification is employed. In addition to estimated coefficients and their p-values, Table 1 reports the model BIC and the p-value of the Cumby-Huizinga test of residual autocorrelation (computed on a linear specification of the model). The null hypothesis of this test is the lack of first order autocorrelation in the residuals, so large p-values are associated to uncorrelated residuals.

Column 1 regresses POL INSTABILITY on NILE FLOODS and lags of the dependent variable. [23]

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

[23] Two lags of the dependent variable were considered in all columns, except in Columns 3 and 9, where three lags

Table 1: POLITICAL INSTABILITY AND CLIMATE SHOCKS

	Dependent Variable: POL INSTABILITY									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NILE FLOODS$_t$	-0.021	-1.320*	-1.171	-1.223*	-1.247*	-1.274*	-1.435**	-0.176*	-1.424*	-1.148
	(0.863)	(0.052)	(0.102)	(0.092)	(0.086)	(0.079)	(0.049)	(0.086)	(0.061)	(0.104)
NILE FLOODS$_t^2$		0.137*	0.128*	0.137*	0.139*	0.142*	0.154*	0.020	0.158*	0.143*
		(0.061)	(0.100)	(0.082)	(0.077)	(0.073)	(0.052)	(0.101)	(0.050)	(0.065)
RAIN HINTERLAND$_t$				0.188						
				(0.119)						
RAIN HINTERLAND$_{t-1}$					0.207*		-2.358			
					(0.086)		(0.106)			
RAIN HINTERLAND$_{t-2}$						0.224*	2.565*	0.038*	0.322**	0.385**
						(0.064)	(0.078)	(0.068)	(0.026)	(0.031)
TENURE$_t$			0.359***	0.319***	0.320***	0.321***	0.320***	0.067***	0.338***	0.323***
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
TENURE$_t^2$			-0.004***	-0.004***	-0.004***	-0.004***	-0.004***	-0.001***	-0.004***	-0.004***
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pseudo R^2	0.409	0.414	0.493	0.489	0.490	0.491	0.496	.752	0.520	0.494
Obs	308	308	307	308	308	308	308	308	308	308
Lags Dep. variable	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lags Tenure	–	–	✓	✓	✓	✓	✓	✓	✓	✓
Period dummies	–	–	–	–	–	–	–	–	✓	–
Time trends	–	–	–	–	–	–	–	–	–	✓
Infor. Crit (BIC)	382.5	385.2	366.3	368.9	368.4	367.9	370.8	383.5	379.2	377.5
Cumby-Huizinga test	0.257	0.291	0.535	0.130	0.136	0.142	0.126	0.142	0.635	0.149
Estimation	OLOGIT	OLOGIT	OLOGIT	OLOGIT	OLOGIT	OLOGIT	OLOGIT	OLS	OLOGIT	OLOGIT

Note: Dependent variable is POL INSTABILITY. Estimation is by maximum likelihood in an ordered logit specification, except in Column 8 where a linear specification and OLS has been employed. p-values (based on heterokedasticity robust standard errors) are reported in parentheses. The period covered is 2685 BCE –1140 BCE and the unit of analysis is 5-year periods. Two lags of the dependent variable were included in all columns except in columns 3 and 9, where 3 lags were needed to avoid serial correlation in the residual term. 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. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

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.[24] 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 10% level). While the former keeps the negative sign, the square term is positive, suggesting that ruler instability is lowest whenever Nile floods were not extreme, either because they were too low, in which case periods of droughts would follow, or too high, as excessive floods could cause destruction of the irrigation infrastructure. [25]

were needed to avoid serial correlation in the residuals. Since these coefficients are not the focus of our analysis, their estimates are omitted to save space.

[24]The inclusion of lags of these variables was also considered but they were never significant and since the BIC increased significantly they were not introduced in the model.

[25]This non-linear relationship between Nile floods and conflict bears resemblance to that reported in Chaney (2013) for a much later period (1169–1425 CE). Chaney (2013) finds that in periods of aberrant Nile floods the highest-

In Column 3, we introduce TENURE$_t$, measuring the number of years from the ruler's first year in office to the first year of the current period. We also add its square, to account for the possibility that ruler replacement becomes more likely as the ruler gets old, and 1 lag of these variables, as dictated by the BIC. These control variables are highly significant, while the coefficients of the Nile variables are similar as in the previous column, although the estimation is a bit noisier, resulting in slightly larger p-values.

Column 4 introduces the contemporaneous value of RAIN HINTERLAND$_t$, our proxy for rainfall in the areas surrounding the Nile valley. Its associated coefficient is positive, indicating that more rain in the surrounding areas tends to increase political instability, although it's not significant (p-value is 0.12). Our theory suggests that the main impact of RAIN HINTERLAND on political instability is through (future) exit pressures, which implies that the main effect will be driven by lagged values of RAIN HINTERLAND. To examine this prediction, Column 5 replaces RAIN HINTERLAND$_t$ by RAIN HINTERLAND$_{t-1}$. This leads to a decrease in the BIC, meaning that under this criterion this model is preferred. The coefficient of RAIN HINTERLAND$_{t-1}$ is positive and significant which suggests that an improvement of weather conditions in the neighboring areas increased ruler instability although the effect needed some time to become apparent. Column 6 considers RAIN HINTERLAND$_{t-2}$ in place of RAIN HINTERLAND$_{t-1}$. The coefficient associated to rainfall in the hinterland increases and its significance improves, while the BIC is smaller, which again suggests that using RAIN HINTERLAND$_{t-2}$ improves the fit of the model. Column 7 introduces in the regression both RAIN HINTERLAND$_{t-1}$ and RAIN HINTERLAND$_{t-2}$ simultaneously. Only the latter is significant and, since introducing both variables in the regression increases the BIC considerably, from now on we capture the effect of weather conditions in the hinterland by using RAIN HINTERLAND$_{t-2}$. Column 8 re-estimates Column 6 by OLS, showing very similar results.

Column 9 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).[26] 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. Introducing period dummies does not change our conclusions. Finally, Column 10 introduces deterministic time trends (a quadratic polynomial of time) to capture systematic trends in the data, obtaining similar results

Summarizing, the models in Table 1 explain a significant fraction of the variance of POL INSTABILITY (that reaches 75% when linear models are considered), and support our hypothesis that political stability in Ancient Egypt was higher in times of relatively high circumscription, that is, whenever Nile floods were favourable and when the weather conditions in the neighboring areas were rough.

---

ranking religious authority was less likely to be replaced. He attributes this finding to the fact that in those periods it was easier for the religious authority to coordinate a revolt to overthrow the political leader. See Section 2 for a discussion of Chaney's paper.

[26] See Appendix B.3 for details about the Egyptian chronology.

Table 2: POLITICAL INSTABILITY AND CLIMATE SHOCKS: VARIATIONS

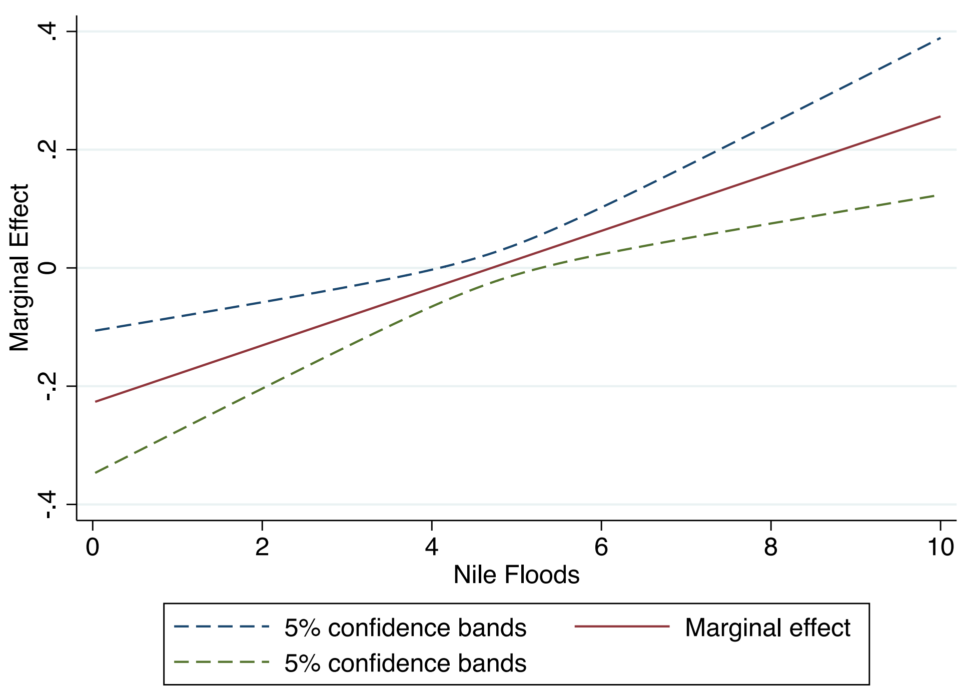
	Dependent Variables: POL INSTABILITY and POL INSTABILITY(DUMMY)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NILE EXTREME FLOOD$_t$	-0.385								
	(0.585)								
NILE EXTREME DROUGHT$_t$	2.290***								
	(0.000)								
NILE EXTREME$_t$		2.306**	0.261**						
		(0.038)	(0.024)						
NILE EXTREME $_t\times$		-0.905**	-0.108**						
RAIN HINTERLAND$_{t-2}$		(0.047)	(0.036)						
NILE FLOODS$_t$				-1.368*	-1.071*	-1.588**	-0.226**	-2.932*	-0.903
				(0.062)	(0.080)	(0.021)	(0.026)	(0.056)	(0.151)
NILE FLOODS$_t^2$				0.134*	0.125*	0.160**	0.023*	0.301*	0.104
				(0.059)	(0.071)	(0.036)	(0.052)	(0.087)	(0.118)
RAIN HINTERLAND$_{t-2}$	0.214*	0.265**	0.045**	0.195*	0.189*	0.119	0.017	0.262*	0.197**
	(0.076)	(0.032)	(0.039)	(0.081)	(0.099)	(0.351)	(0.400)	(0.099)	(0.032)
(Pseudo) R^2	0.493	0.492	.752	0.497	0.443	0.295	.362	0.341	0.737
Obs	308	308	308	296	384	308	308	153	1536
Lags Dep. variable	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Inf Criteria (BIC)	366.644	367.2885	383.2916	353.5956	466.0557	353.0494	358.225	266.0714	690.3597
Cumby-Huizinga test (p-val)	0.116	0.165	0.165	0.128	0.996	0.674	0.674	0.133	0.142
Estimation	OLOGIT	OLOGIT	OLS	OLOGIT	OLOGIT	LOGIT	OLS	OLOGIT	OLOGIT

Note: Dependent variable is POL INSTABILITY except in columns 8 and 9 where POL INSTABILITY(DUMMY) was employed instead. Ordered logit models are estimated except in Columns 3 and 7, where linear specifications are employed, and Column 6, where a logit one is considered. The controls are TENURE, its square and 1 lag of these variables. The unit of analysis is 5 year periods, except in columns 8 and 9 where the unit of analysis is 10-year and 1-year periods, respectively. The Baxter-King filter has been applied on the climate variables in Column 4. The sample period is 2686 BCE – 1140 BCE except in Column 5, where it is 2686 BCE – 760 BCE. Lags of the dependent value are included in all models. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Table 2 considers additional variations to probe into the robustness of the results in Table 1. Ordered logit specifications are employed unless otherwise noted, and the same controls as in our baseline specification (Column 6 in Table 1) are considered in all columns, but the value of the associated coefficients is not reported to save space.

Column 1 is identical to Column 6 in Table 1 but captures the non-linear relationship of Nile floods and ruler stability by defining two dummy variables, NILE EXTREME FLOOD and NILE 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 political instability than extreme floods. Column 2 considers whether there is an interaction effect between the two weather shocks. To examine whether the effect of RAIN 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 RAIN HINTERLAND$_{t-2}$ are positive and significant, as expected, but the interaction of the two is negative and also significant. Since interpreting interaction terms in non-linear models is not as direct than in linear ones, Column 3 re-estimates Column 2 by OLS, obtaining very similar results. Using an F-test we cannot reject that the sum of the coefficients of RAIN HINTERLAND$_{t-2}$ and the interaction term is equal to zero. [27] 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.

Figure 4: MARGINAL EFFECT OF NILE FLOODS ON POL INSTABILITY



NOTES: This graph plots the marginal effect of NILE FLOODS on POL INSTABILITY(DUMMY). Estimates correspond to Table 2, Column 7.

The remaining columns in Table 2 provide additional variations: Column 4 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 6, Table 1. Column 5 considers a longer time period than in our baseline regressions, more specifically from 2686 BCE to 760 BCE.[28] Column 6

[27] Using a similar F-test, we do reject that the sum of the coefficients of NILE EXTREME$_t$ and the interaction term is zero.

[28] We don't consider this interval in Table 1 because the level of resolution of RAIN HINTERLAND becomes very crude after 1140 BCE.

considers a binary dependent variable, POL INSTABILITY(DUMMY), and a logit specification is employed. Column 7 re-estimates Column 6 using a linear specification estimated by OLS. Finally, columns 8 and 9 replicate our baseline specification (Column 6 in Table 1) using different units of analysis. In particular, Columns 8 and 9 consider 10-year and 1-year periods, respectively, as unit of analysis. In general, results in Table 2 provide support to the circumscription hypotheses, i.e., favorable weather conditions in the core together with rough ones in the hinterland contributed to higher political instability in Ancient Egypt.

Figure 4 reports the marginal effect of NILE FLOODS on POL INSTABILITY, which has been computed using the coefficients in Column 7 from Table 2 (corresponding, for simplicity, to the case where the dependent variable is binary and the equation has been estimated by OLS). It shows that increases in NILE FLOODS have a negative and statistical significant effect on political instability for moderate values of NILE FLOODS but the marginal effect becomes positive for large values of NILE FLOODS.

Table 3: POLITICAL INSTABILITY(DYNASTIES) AND CLIMATE SHOCKS

	Dependent Variables: POL INSTABILITY(DYNASTIES) and POL INSTABILITY(DYNASTIES, DUMMY)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NILE FLOODS$_t$	0.134	-2.565**	-4.272**	-8.030***		-0.224**	-8.953**	-8.746	-6.910***	-0.145**
	(0.606)	(0.033)	(0.012)	(0.007)		(0.033)	(0.045)	(0.111)	(0.000)	(0.036)
NILE FLOODS^2_t		0.287**	0.530***	0.938***		0.026**	1.073**	1.050*	0.794***	0.017**
		(0.029)	(0.004)	(0.004)		(0.026)	(0.033)	(0.052)	(0.000)	(0.035)
NILE EXTREME FLOOD					5.048**					
					(0.019)					
NILE EXTREME DROUGHT					2.919*					
					(0.070)					
RAIN HINTERLAND$_{t-2}$				1.046**	0.709**	0.008	1.246**	1.371*	0.769**	0.005
				(0.017)	(0.014)	(0.449)	(0.042)	(0.097)	(0.023)	(0.533)
(Pseudo) R^2	0.645	0.659	0.737	0.755	0.736	.872	0.785	0.816	0.741	0.752
Obs	308	308	308	308	308	308	308	308	308	308
Lags Dep. variable	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls (TENURE DYNASTY)	–	–	✓	✓	✓	✓	✓	✓	✓	✓
Add. controls (TENURE RULER)	–	–	–	–	–	–	✓	–	–	–
Period dummies	–	–	–	–	–	–	–	✓	–	–
Inf. Crit (BIC)	147.4	148.5	145.2	144.9	151.4	99.3	157.9	147.4	132.2	-89.4
Cumby-Huizinga test (p-val)	0.480	0.425	0.245	0.627	0.627	0.724	0.679	0.821	0.622	0.548
Estimation	OLOGIT	OLOGIT	OLOGIT	OLOGIT	OLOGIT	OLS	OLOGIT	OLOGIT	LOGIT	OLS

Note: Dependent variable is POL INSTABILITY(DYNASTIES) POL INSTABILITY(DYNASTIES, DUMMY). P-values based on HC standard errors are reported in brackets. Ordered logit models have been estimated except in Columns 9 and 10 where a logit and a linear specification has been considered, respectively. Controls are TENURE(DYNASTY), its square and 1 lag of these variables. Additional controls are similar but referred to ruler's tenure. The sample period is 2686 BCE – 1140 BCE *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

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. 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 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. More specifically, POL INSTABILITY(DYNASTIES) is equal to 1 if there is a dynasty replacement in period t and equals 2 in periods where there are two or more dynasties in power, as those periods reflect lack of central rule and a high degree of political instability. Otherwise, it takes a value equal to zero.

This variable presents some advantages with respect to POL INSTABILITY, 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. In particular, ordered logit models are estimated unless otherwise stated. Column 1 in Table 3 regresses POL INSTABILITY(DYNASTIES) on NILE FLOODS and lags of the dependent variable. As in previous columns, two lags of the dependent variable were enough to avoid residual autocorrelation. This column shows that there is a not a statistically significant (linear) relationship between the former variables. However, the results change if the square of NILE LEVEL is added to the former specification. Column 2 shows that too low or too high Nile floods are associated with higher dynastic instability.[29] 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.[30] As in the case of TENURE, we also allow for a squared term of TENURE DYNASTY to capture potential nonlinear effects, as well as 1 lag of these variables (whose coefficients are not reported to save space). The BIC decreases after introducing these controls suggesting an improvement in model fit, but otherwise results are unaffected by the introduction of these controls.

Column 4 introduces RAIN HINTERLAND$_{t-2}$ in the model.[31] The coefficient of this variable is positive and statistically significant. Column 5 captures aberrant Nile floods by employing 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 behavior. Column 6 re-estimates Column 4 by OLS in a linear specification. Similar results are found but the coefficient associated with RAINFALL HINTERLAND$_{t-2}$ is now smaller and not significantly different from zero. Column 7 introduces additional controls, more specifically it considers TENURE and its square, as well as 1 lag of these variables, as in Tables 1 and 2, which doesn't alter the conclusions from previous columns. Column 8 adds period dummies, which increase a bit the p-values but have little impact on the estimated coefficients. Finally,

[29]We also 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.

[30]If several dynasties are in power, this variable measures the time spanned from the last dynastic change.

[31]As in previous tables, other lags of RAIN HINTERLAND were considered as well and RAIN HINTERLAND$_{t-2}$ provided the best fit according to the BIC.

Columns 9 and 10 employ a bivariate dependent variable that equals 1 whenever there's a dynasty replacement or whenever there were two or more dynasties in power. Column 9 considers a logit specification and obtains very similar results. Column 10 re-estimates the same model using a linear model. Similar results are also found, except for the coefficient of RAIN HINTERLAND$_{t-2}$, which now is smaller and becomes insignificant.

Summarizing, the results in Table 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. If anything, weather shocks seem to be more strongly linked to POL INSTABILITY (DYNASTIES) than to the previous variables, suggesting that they are associated with more severe forms of political instability. The following section presents additional evidence supporting our claim that periods of high circumscription helped the consolidation of the state.

7.2 Additional evidence: State capacity

The theory outlined in Section 3 (and Appendix A.1) posits that periods of high circumscription are associated with a higher tax collection and a greater state capacity. This section explores the impact of changes in circumscription on our proxy for the latter.

Our 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 km^2).[32] 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-run, leading therefore to a reduction in the area under the state's control.

Table 4 examines these predictions. As mentioned in subsection D.3, we could not reject the null hypothesis of a unit root in AREA (see Table D.4 in Appendix D). Then, Table 4 mainly focuses on the first difference of the log of the area under state control, denoted by AREA(GROWTH), which can be interpreted as the area's growth rate.[33] Column 1 regresses AREA(GROWTH) on NILE FLOODS and its square as well as on similar controls as in previous tables.[34] No lags of the dependent variable have been introduced as residuals do not exhibit autocorrelation and the BIC increases after introducing them. The coefficient of NILE FLOODS is positive while that of NILE FLOODS^2 is negative, suggesting that AREA(GROWTH) is positively associated to intermediate levels of Nile floods.

[32] The data come from Geacron, `http://geacron.com/home-es/?lang=es`.

[33] We've also tested the non-stationarity of AREA(GROWTH), which was rejected at at the 1% level.

[34] This includes TENURE and its square, TENURE(DYNASTY) and its square and 1 lag of these variables.

Table 4: AREA CONTROLLED BY THE STATE AND CLIMATE SHOCKS

	Dependent Variable: AREA(GROWTH) (Columns 1–6) and AREA (Column 7)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
NILE FLOODS$_t$	0.056*	0.063*	0.065*	0.067*		0.091*	0.074*
	(0.092)	(0.079)	(0.076)	(0.073)		(0.060)	(0.065)
NILE FLOODS$_t^2$	-0.007*	-0.007*	-0.008*	-0.008*		-0.010*	-0.009*
	(0.068)	(0.062)	(0.060)	(0.058)		(0.055)	(0.057)
NILE EXTREME FLOOD$_t$					-0.010		
					(0.728)		
NILE EXTREME DROUGHT$_t$					-0.033		
					(0.209)		
RAIN HINTERLAND$_t$		-0.010**					
		(0.047)					
RAIN HINTERLAND$_{t-1}$			-0.011**				
			(0.041)				
RAIN HINTERLAND$_{t-2}$				-0.012**	-0.013**	-0.018	-0.011*
				(0.038)	(0.035)	(0.122)	(0.083)
R^2	0.085	0.089	0.090	0.091	0.087	0.101	0.962
Obs	308	308	308	308	308	308	308
Lags Dep. Variable	–	–	–	–	–	–	✓
Controls	✓	✓	✓	✓	✓	✓	✓
Period dummies	–	–	–	–	–	✓	–
Inf. Criteria (BIC)	-67.0	-62.5	-62.8	-63.1	-61.7	-43.8	-57.9
Cumby-Huizinga test (p-val)	0.539	0.503	0.494	0.485	0.507	0.362	0.584

Note: The dependent variable is AREA(GROWTH), the first difference of the area under state control except in Column 7, where it's AREA. Estimation is by OLS. P-values based on HC standard errors are reported in brackets. The control variables are: TENURE and its square, TENURE(DYNASTY) and its square, and one lag of all these variables. The table also reports the value of the Bayesian information criterion (BIC) as well as the p-value of the Cumby-Huizinga test of first order residual autocorrelation. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

Column 2 introduces the contemporaneous value of rainfall in the hinterland. It has a negative and significant coefficient, implying that more rainfall in the areas surrounding the Nile is associated to a decrease in AREA(GROWTH). Columns 3 and 4 lag 1 and 2 periods the value of RAIN HINTERLAND. The size of the coefficient tends to increase (moderately) in absolute value with the lag and the value of the BIC decreases, suggesting a similar behavior as in previous tables. Column 5 captures the Nile behavior by considering dummies for extreme floods and extreme droughts, which have a negative sign but are insignificant. Column 6 considers period dummies and the results are similar but rainfall in the hinterland is now measured less precisely (p-value is .12). Finally, Column 7 regresses the variable in levels, i.e, AREA on the same variables as Column 4 and a lag of the dependent variable (needed to capture autocorrelation in the residual term), obtaining very similar results.

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 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 outlined 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 or switch to non-transparent food subsistence activities, 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 sudden disappearance of the Maya culture on the Yucatan peninsula around 900 CE. Is it even possible that lessons from such 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

Allen, R.C. (1997) ”Agriculture and the Origins of the State in Ancient Egypt” *Explorations in Economic History* 34, 135-154.

Allen, R.C., M.C. Bertazzini, and L. Heldring (2020) ”The Economic Origins of Government” working paper.

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 planktonicforaminifera 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.and C. Guerriero (2021) “Climate Change and State Evolution ” *PNAS* 118(14), doi.org/10.1073/pnas.2020893118.

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.

Blaydes, L. and E. Chaney (2013) ”The Feudal Revolution and Europe’s Rise: Political Divergence of the Christian West and the Muslim World before 1500 CE” *American Political Science Review* 107(1), 16-34.

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.

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.

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.

Hurst, H.E. (1951) “Long-term storage capacity of reservoirs”, *Trans. Amer. Soc. Civil Eng.* 116, 770-808.

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 (2022) “Cereals, Appropriability and Hierarchy” *Journal of Political Economy*, forthcoming.

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.

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

Moreno Garcia, J.C. (2020) *The State in Ancient Egypt* London: Bloomsbury.

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. (2020) ”The Origin of the Extractive Compatible State: Environmental Circumscription” UB Berkeley, *mimeo*.

Scott, J.C. (2009) *The Art of Not Being Governed: An Anarchist History of Upland Southeast Asia*. Yale University Press, New Haven.

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.

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

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

A Theory

A.1 A Formal Model

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

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

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

A.1.1 Individuals

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

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

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

$A^j_{t+1} = \rho^j A^j_t + r^j_{t+1}$, with $\rho^j \in (0,1), j = \{c,h\}$.[37] We also assume that the shocks r^c_t and r^h_t are mutually independent, i.i.d. random variables with positive support and means $\{\mu^c, \mu^h\}$, respectively.

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

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

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

$$E_t((1-\tau)A^c_{t+1}|A^c_t) > E_t(\psi_i A^h_{t+1}|A^h_t),$$

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

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

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

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

It follows that the larger the productivity gap between the core and the hinterland at time $t = 1$, the higher the share of population in the core at time $t+1$. From the characterization of the threshold $\overline{\psi}_t$, it is straightforward to see what the determinants of L^c_{t+1} 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 $t = 1$; Specifically, weather shocks that increase A^c_t or that decrease A^h_t will increase population in c at time $t+1$.*

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

A.1.2 The state

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

So far we have considered that d_t is simply a realization from a random variable D_t. However, it is reasonable to expect that D_t and TR_t are related. The sign of this association is likely to be related to the ultimate goal of the attack. There is ample support in the literature that economic

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

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

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

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

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

downturns (that in our framework translate into a reduction in TR_t) increase the probability of political instability (Acemoglu and Robinson 2006, Burke and Leigh 2010, Chaney 2013, etc.). Periods of economic stress facilitate the organization of internal revolts, as factions of the elite can take advantage of the weakness of the state (and thus, the smaller size of mobilized defensive resources) and the discontent in the population (due to scarcity and famine) to organize an insurrection and replace the leader. This is the so-called "opportunity cost" effect. An alternative scenario would be one in which the aggressor is primarily interested in conquering TR_t, in which case D_t are TR_t will be positively related (the "rapacity" effect, see Dube and Vargas, 2013).

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

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

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

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

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

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

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

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

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

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

A.2 Proofs.

This section presents the proof to Propositions 1 and 2.

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

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

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

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

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

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

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

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

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

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

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

Proof of Proposition 2

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-1}^h). We consider here the case where λ=0 and discuss at the end the general case. Under the assumptions of Proposition 2 this derivative is given by

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

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

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

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

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

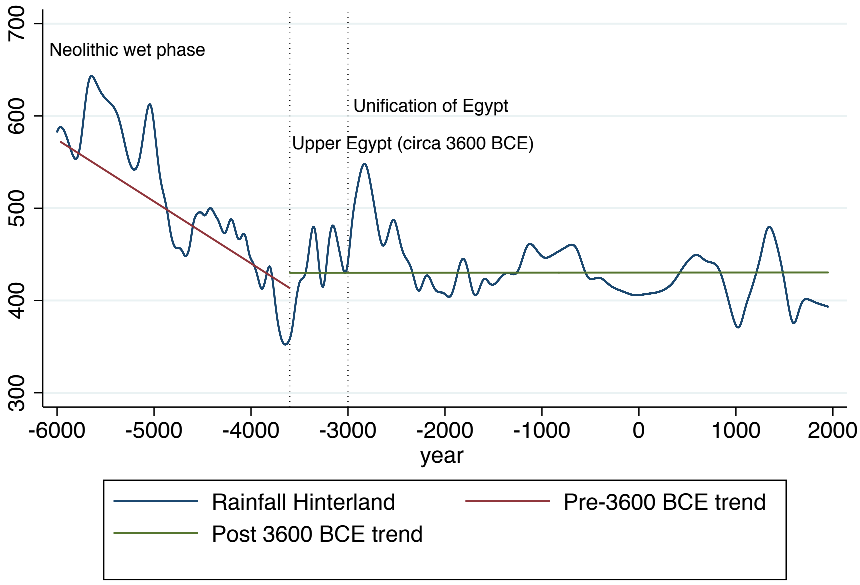
In the case of Egypt, however, we argue that the case where $\lambda \leq 0$ could be the most relevant one. The pharaoh was believed to be responsible for the Nile floods. Then, in case of famines derived from extreme Nile floods he was often directly blamed for them (Bell, 1971). This could have led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

B Historical Background: Additional documentation.

B.1 The Neolithic wet phase and the creation of the Egyptian state

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

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



B.2 Exogenous shocks and political instability in historical Egypt

Two recent studies have examined the relationship between exogenous geophysical shocks and political instability in historical Egypt. Using the Roda Island (Cairo) Nilometer data from the Muslim period in Egypt, Chaney (2013) studies the impact of extreme Nile floods on the probability of replacement of the Head Judge, i.e. the main religious authority in the country during the Abbusid dynasty. The empirical study covers mainly the years 1169-1425 CE. Using text data from an annual chronicle of the period, Chaney (2013) finds that extreme Nile floods are associated with more references to high food prices and a more frequent use of words like “combat, riots and looting”. However, the main result of the paper is that extreme floods implied a lower risk of replacement of the Head Judge and a greater construction of religious monuments.[45] 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.

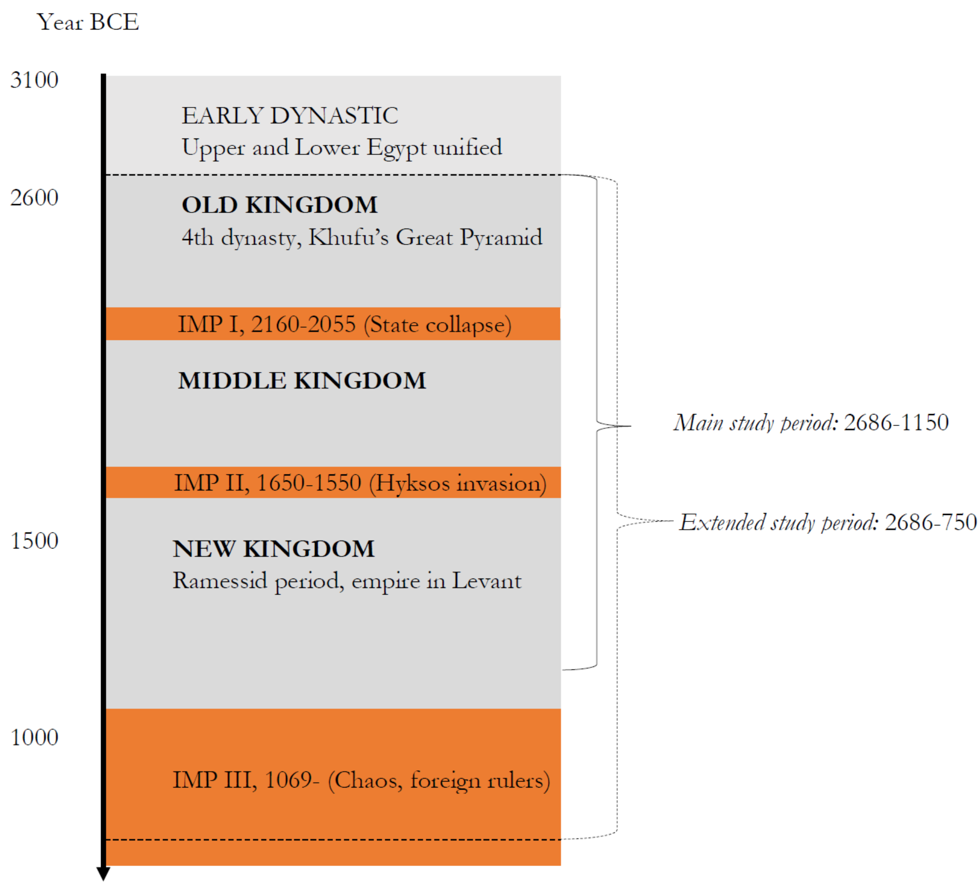
B.3 Rulers and Dynasties: Chronologies and Dating issues

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 3nd century BC historian Manetho, as well as inscriptions on monuments and other archeological artefacts such as the *Palermo stone* that offer chronologies of events during certain rulers. Modern scholars have further used astronomical observations, referred to in ancient texts, of lunar phenomena with known regular cycles in order to anchor ruler reigns to specific years. Despite these efforts, there are several remaining measurement issues with the exact dating of ruler tenures, including the recurring periods of fundamental uncertainty during intermediate periods when many state functions collapsed.

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

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

Figure B.2: ANCIENT EGYPT CHRONOLOGY.



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

ternative 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.[46] 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

[46]See `http://www.metmuseum.org/toah/hd/phar/hd_phar.html` (October 2002).

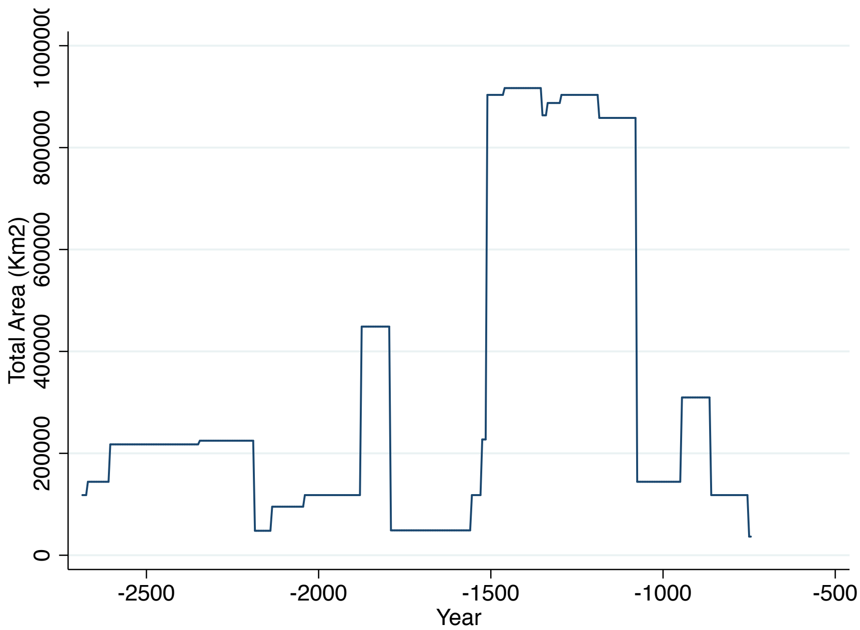
study (2686-1150 BCE) covers dynasties 3-20 in the Egyptian chronology, all of which were native to Egypt. Later dynasties, such as the 22-23rd dynasties, were often of foreign descent.

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.

Figure B.3: AREA (IN KM^2) OF ANCIENT EGYPT.



NOTES: The figure shows the total area controlled by the Egyptian state during antiquity. Source: Geacron

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 (δ18O isotope-levels) from a stalagmite in Qunf Cave in Southern Oman (Fleitmann et al, 2003). We argue that it should provide a reliable indicator for the intensity of African Monsoon precipitation in the Ethiopian highlands and thus reflect the annual movements of the Intertropical Convergence Zone (ITCZ). Figure 2 shows the location of Qunf cave in relation to the movements of the ITCZ. Furthermore, we believe that there are several advantages of using this data including: i) The high time resolution with an average of 3.85 years during the relevant period; ii) Unlike data from historical Nilometers, the Qunf data have been uncontaminated by human manipulation (the data from the Roda Nilometer in Cairo was partly collected for tax collection purposes by different ruling elites during almost 1300 years); and iii) It is one of the most frequently cited works on the paleoclimate in the area from this period.[47]

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.

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

We would expect to find a positive relationship between our Nile flood (Qunf) observations and the Lake Tana Ti concentration. Figure C.4a 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.

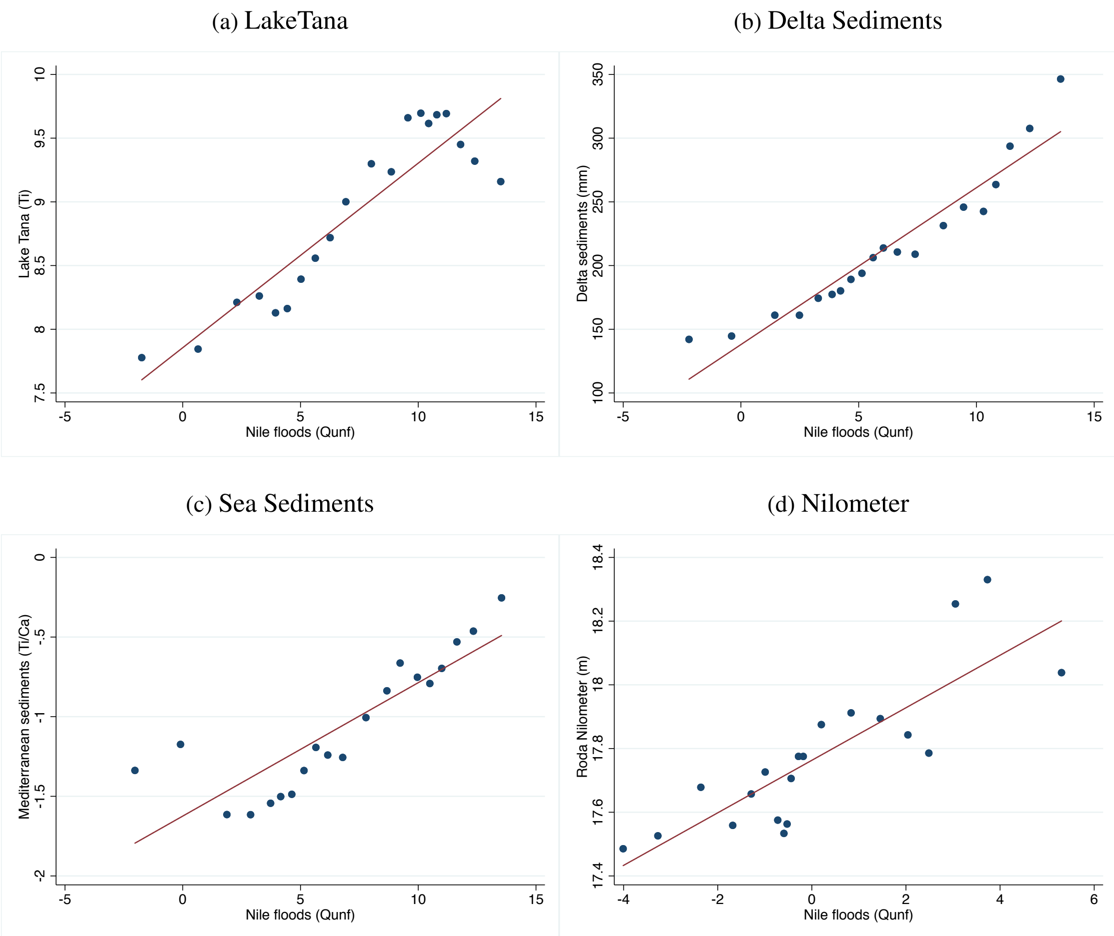
A second method for assessing historical Nile floods exploits the thickness of sedimentation layers in the Nile delta and also out in the Mediterranean. Marriner et al (2012) collected sediment levels from several locations in the Nile Delta for successive centennial observations. In Figure C.4b, we show the relationship between interpolated, annual sediment levels and Nile

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

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

Figure C.4: BINSCATTER PLOTS BETWEEN NILE FLOODS AND ALTERNATIVE PROXIES.



NOTES:

The figure shows the fitted binned scatter plot relationship between our Nile floods (Qunf) index and various other climate indicators in the same order as specified in Table C.1.

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.4c shows the binscatter for 6,805 interpolated, annual observations and Table C.1 shows that the correlation coefficient is almost 0.68.

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

As discussed above, the Roda Island Nilometer in Cairo has data for a much later time period than the one studied in the paper (622-1921 CE). In order to investigate the association between the Qunf data and the Roda Nilometer, we correlate the filtered, annual observations of the maximum Nile levels with our Nile floods-index. In total, we can use 1,005 yearly observations during 638-1642 CE. Figure C.4d 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 $\delta 18$O-levels from Soreq cave, 18 km west of Jerusalem. Figure C.5 shows a map of the region, which also includes the location of the Roda Island Nilometer (in Cairo).

A first and critical issue is how well our rainfall index, created from isotope observations from a stalagmite inside Soreq cave, captures actual observed rainfall outside the cave. Figure C.6a 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.

A second question is how well rainfall in Israel also captures rainfall in Egypt's hinterland? Our

Figure C.5: SOREQ CAVE LOCATION

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.6b, we show the binned scatterplot of 1,344 monthly precipitation observations in Israel and Egypt over the time period 1901-2012 from World Bank (2019). 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 is 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.6c, we show that there is, as predicted, a negative and significant relationship for the time period 5000 BCE to 1150 BCE ($\rho = -0.63$).

Fourth, it is well known from ancient texts that a major threat to the Egyptian civilization during periods of drought was that the drying up of the hinterland led to the growth of massive sand dunes that sometimes invaded or completely swallowed cultivated fields near the Nile (Bell, 1971). The sediments layer collected by Revel et al (2015) out in the Mediterranean also includes data on the elemental ratio Si/Al, which derives from aeolian dust or sand, blown into the Nile from the nearby deserts. An important component of desert sand is quartz (SiO_2) and the Si/Al-ratio might thus serve as a proxy for the historical sandstorms and drought conditions. In Table C.2, column (4), and in Figure C.6d, we show that there is a moderately strong negative

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

Table 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.

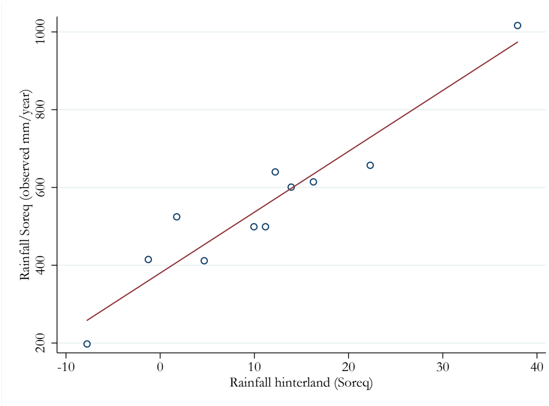
relationship between Rainfall hinterland and dust intensity in the sediment layers over the time period 6686-1150 BCE($\rho = -0.243$). The binscatter in Figure C.6d shows that the correlation is weak in lower levels of our Rainfall hinterland index, suggesting that in periods of very low rainfall levels, the desert climate becomes volatile and might either bring high or low dust storms.

Figure C.7 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.7 shows that both variables display quite similar patterns.

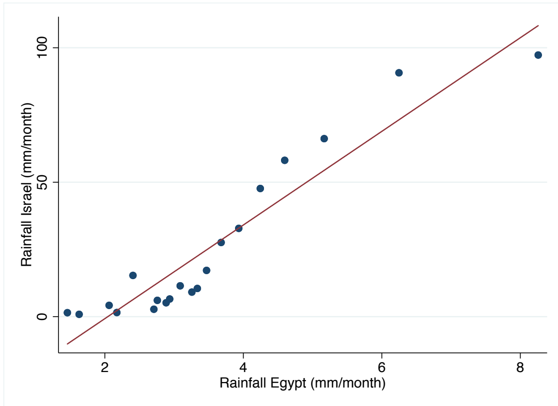
Lastly, we made it an explicit assumption in our theoretical model that shocks to the Mediterranean and Monsoon weather systems are uncorrelated. Table C.2, column (5) and Figure C.8 shows the binscatter plot for the filtered 5-year average values of our Nile floods and Rainfall hinterland-variables 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.6: RAINFALL IN THE HINTERLAND VERSUS ALTERNATIVE PROXIES.

(a) Observed rainfall versus Rainfall Hinterland

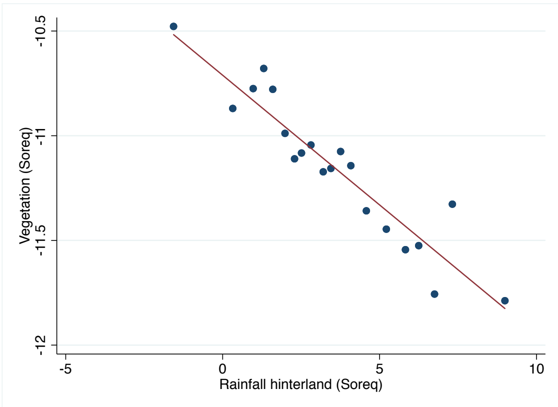


(b) Observed Rainfall: Israel vs Egypt

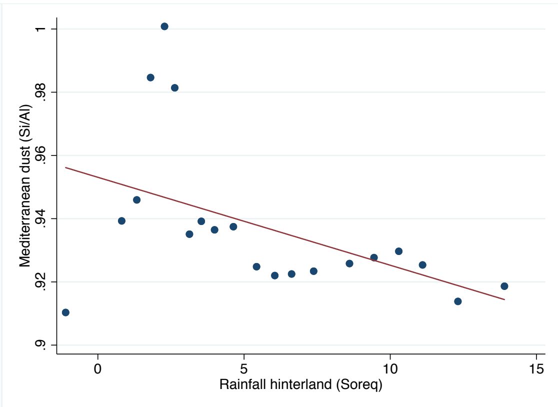


%

(c) Vegetation vs Rainfall Hinterland



(d) Mediterranean Dust versus Rainfall Hinterland



NOTES: Panel (a) shows the scatterplot between our Rainfall hinterland (Soreq) index and observed rainfall outside Soreq cave for available annual observations 2000-2010, whereas Figures in panels (b) to (d) show the binned scatterplot between our Rainfall in the hinterland and variables in the same order as specified in Table C.2.

D Variable definition, Summary Statistics and Preliminary Analysis

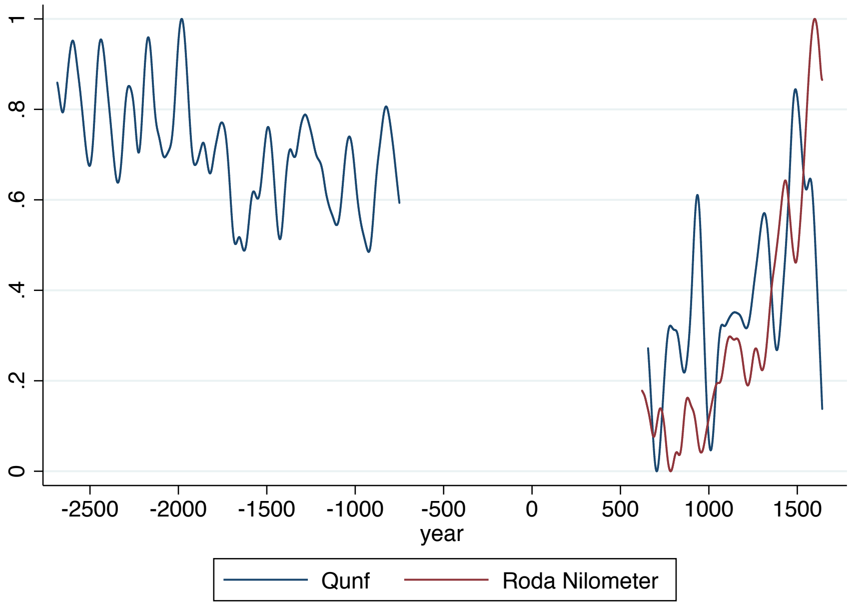
This section provides definitions of the variables employed in the empirical analysis (Section D.1) and a table of summary statistics (see Section D.2). Section D.3 describes a preliminary analysis that examines the time series properties of the variables employed in the empirical analysis.

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

Figure C.7: NILE FLOODS (QUNF) VERSUS RODA NILOMETER DATA.



NOTES: The figure shows the evolution of the Nile floods data obtained from the Qunf cave as well as direct measurements of the Nile river obtained from the Roda Nilometer. Both series have been filtered using a Butterworth index and to facilitate the comparison, the resulting variables have been normalized to be in the interval [0,1].

contain a ruler change];[48] 2 if any of the years in the five-year period $(t, t+1)$ corresponds to periods of high political instability for which exact chronology is not available [around 17% of the periods];[49] and 0 otherwise.[50]

- POL INSTABILITY (DYNASTY)$_t$: This variable is defined in five-year periods and takes three values: 1 if there is a change of dynasty in the 5-year period $(t, t+1)$, 2 if two or more dynasties coexist in that period according to Shaw's chronology, as the fact that two or more dynasties coexist in the same period is a symptom of lack of centralization and political instability, and 0 otherwise. See Appendix B1 for definitions and main historical sources of Egyptian dynasties.[51]

- AREA$_t$: Area under State Control. This variable captures the log of the area (in km^2) under

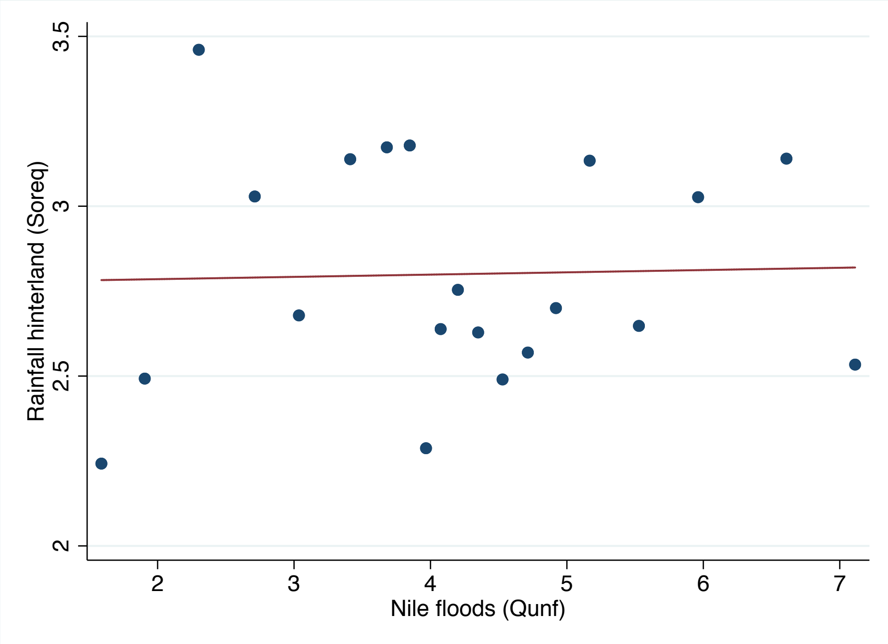
[48] See Appendix B for a discussion about the chronologies and dating issues for rulers and dynasties. Thanks to the information provided by a Ramessid king-list written on a papyrus on the Museo Egizio in Turin (Turin cannon), there are few weak links in the order and dating of Old Kingdom rulers (exceptions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

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

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

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

Figure C.8: NILE FLOODS VERSUS RAINFALL IN HINTERLAND.



NOTES: The figure shows the binned scatterplot relating our Nile floods (Qunf) index and rainfall hinterland (Soreq) index as specified in Table C.2.

State control at the end of the five-year period starting in t. Source: Geacron.

- AREA$_t$(GROWTH): Growth rate of the area under State control. This variable is the first difference of AREA$_t$. Source: Geacron.

- 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 3 (upper pannl) 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 3 (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.

- 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
AREA	390	327365.2	316089.9	36561	916795
TENURE	390	14.24	15.43	1.00	93.00
TENURE DYNASTY	390	76.89	57.17	1.00	255.00

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 dependent variables employed in the empirical analysis. Table D.4 presents the results. The modified GLS Augmented Dickey-Fuller test (ADF-GLS, Elliot et al., 1996) was employed and lags were chosen using the Modified BIC. A constant and a trend were included in all cases.

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

Table D.4: UNIT ROOT TESTS. This table presents the value of the ADF-GLS test for unit roots (Elliot et al, 1996) for each of the dependent variales employed in the empirical analysis. The null hypothesis is the existence of a unit root in the data. Large negative values indicate rejection of the null hypothesis. The number of lags was chosen using the Modified BIC. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

E Additional references

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

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

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

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

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

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

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/.