Transition to agriculture and first state presence: A global analysis

Oana Borcan^a, Ola Olsson^b, Louis Putterman^c

 ^aSchool of Economics, University of East Anglia, Norwich NR4 7TJ, United Kingdom. Corresponding author: O.Borcan@uea.ac.uk
 ^bDepartment of Economics, University of Gothenburg, Gothenburg 40530, Sweden

^cDepartment of Economics, Brown University, Providence RI 02912, USA

Abstract

It has often been observed that the emergence of states in a region is typically preceded by an earlier transition to agricultural production. Using new data on the date of first state emergence within contemporary countries, we present a global scale analysis of the chronological relationship between the transition to agriculture and the subsequent emergence of states. We find statistically significant relationships between early reliance on agriculture and state age in all sub-samples and also when we use alternative sources of data at different levels of geographical aggregation. A one millennium earlier transition to agriculture among non-pristine states predicts a 317-430 year earlier state emergence. We uncover differences in cases where states were imposed from outside or when they emerged through internal origination. The agriculture-state lag is on average 3.1 millennia in internally originated (including pristine) states, and 2.7 millennia in externally originated states. We also explore some of the mechanisms through which agriculture is believed to have influenced the emergence of states. Our results indicate that the rise of social classes was often an intermediate step towards the presence of early states.

Keywords: Agricultural transition, States JEL classification: N50,O43 It has been generally acknowledged in recent years that the transition from a reliance on foraging as the main source of calories to a reliance on crop cultivation and animal husbandry, had a far-reaching impact on social organization and the subsequent pattern of economic development.¹ Building on Jared Diamond's (1997) famous contribution, Hibbs and Olsson (2004) and Olsson and Hibbs (2005) show that regions that were endowed with a multitude of suitable wild plants and animals for early domestication, are still more developed today than regions that had few or no suitable species. Putterman (2008) introduced new measures of the date of transition to agriculture for all contemporary countries in the world and demonstrated that an earlier reliance on agriculture was associated with higher incomes in 1500 CE.² More recently, Scott (2017) even proposed that the rise of densely populated farming villages led to the self-domestication of humans which, in turn, was a prerequisite for more complex social organization.

A key intermediate mechanism between the emergence of agriculture and long-run economic development, is clearly the evolution of early states. A common observation of anthropology and archaeology is that emergence of the macro polities we call states typically followed by a few millennia the transition to agriculture (Service and Sahlins, 1960; Service, 1971; Diamond, 1997; Johnson and Earle, 2000). Typically, the pattern has been remarked on with reference to a small number of cases, limiting tests for statistical regularity. We partially address this omission by using our recently compiled data (Borcan et al, 2018) that permit the contours of the agriculture-to-state passage to be studied statistically on a global scale. Our main analysis takes as observational

¹We would like to thank the editor Carola Frydman as well as two anonymous reviewers for very helpful comments. Louis Putterman wishes to thank Jared Diamond and Alan Farahani for aiding his understanding of the literatures on early agricultures and states in the course of ongoing research. We also thank James Fenske for a useful discussion, Christopher Paik for his help in sourcing sub-national data, as well participants at the Quantitative History Webinar at Hong Kong Business School. The authors have no relevant or material interests relating to the research described in this paper.

²See Spolaore and Wacziarg (2013) for an overview of this literature.

units the territories of 159 countries of the year 2000 CE, accounting for 96 percent of all countries having populations above 0.5 million in that year. The countries included account for over 90 percent of the world's population and for almost 99 percent of its land surface, excluding Antarctica. We code for presence or absence of states beginning 3500 BCE, the estimated date of transition to centralized political organization above tribal level in southern Mesopotamia. We acknowledge that these code for two key thresholds only—the paramount chiefdom or proto-state, and full state (see below)—so finer gradations in the long-run emergence of political centralization are left partly to ancillary analysis (Section 3) and mainly for other research, as more nuanced and comprehensive data become available.

All countries covered have achieved their first state presence by 2000 CE, with considerable variation in timing and nature of state emergence. We identify as "pristine" those states which emerged in the absence of nearby models of macro polity. Such states arose in eight countries of today (hereafter countries). We designate as "externally originated" the states of 72 countries where initial state emergence is attributed to annexation or colonization from outside. We identify as "internally originated" states 79 intermediate cases in which states emerged earliest as the result of internal political developments but in a world region in which large scale polities were gradually appearing in likely diffusion from an originally pristine core.

In line with the extant literature, we propose a basic framework whereby a transition to agriculture in most cases ushered in socio-economic changes including a higher level of fertility, greater sedentism, and new forms of social organization in larger social units. Across the globe, these social changes then led to similar processes of greater political integration that eventually paved the way for the rise of states. The main contribution of this paper is the empirical characterization of this process. More specifically, we study the dynamics in the development sequence from agriculture to states. Our general prior is that an earlier transition to agriculture also implied an earlier transition to states, and vice versa. We study the length of the temporal lag from agriculture to states among pristine versus non-pristine states and among internally versus externally originated states. Among pristine cases, there was no earlier state experience to build upon. In externally originated states, the state was imposed through an intervention by foreigners and the timing was less dependent on internal conditions. We thus expect the temporal lag to be relatively long in pristine states and have a weaker association with domestic time since agriculture in externally originated states.

The slope in the estimated relationship between time since agriculture (*TimeAgri*) and time since the emergence of states (*StateAge*), further provides interesting information about whether the length of the temporal lag varies with time since agriculture. If, for instance, the estimated parameter β in a regression $StateAge_i = \alpha + \beta \cdot TimeAgri_i + \epsilon_i$ is positive but below unity, then this implies that the temporal lag between agriculture and state emergence decreases as we approach the contemporary era.³ This would be consistent with a hypothesis suggesting that learning from previous transitions would speed up internally originated state formation processes in regions that had a late transition to farming. As mentioned above, it further seems likely that the relationship between time since agriculture and state emergence should be weaker in externally than in internally originated states, implying a relatively low and less precisely measured β in the former category.

In our empirical investigation, we find as expected a strong positive association between time of transition to agriculture and time of state emergence even when controlling for geographic and climatic factors, distance from the relevant diffusion zone's pristine state, and time of first human settlement, as well as when modelling spatial spillovers and using an instrumental variables strategy. Our estimated average time from primary reliance on agriculture to full state emergence is 3406 years for pristine, 3100 years for internally originated and 2731 years for externally originated states, and our estimates imply that a one millennium earlier transition to agriculture among non-pristine states predicts a 317-430 year earlier state emergence depending on the

³Based on the reasoning above, we also hypothesize that the intercept should most often be nonpositive: $\alpha \leq 0$.

exact specification. The β -coefficient is consistently positive but far below unity, implying that the temporal lag between the transitions to agriculture and states typically decreased closer to the contemporary era. We also find that the β -coefficient is indeed lower in externally originated states.

In an attempt to reach a deeper understanding of the mechanisms through which agricultural production led to state formations, we also include an analysis of cultural complexity before the Common Era. We use cultural data from Peregrine (2003) coded for 65-75 polygons (different from modern-day country territories) at different periods in prehistory. We also obtained sub-national polygons by overlapping maps of archaeological tradition regions with modern-day country boundaries, in order to aggregate the cultural complexity data at the country level in two periods in prehistory: 2000 and 1000 BCE. We then coded a binary variable for the presence of supra-tribal rule, based on the Peregrine (2003) political integration variable, which captures the number of levels of political integration above the local community level. We use this data as our dependent variable in regressions at archaeological tradition polygon level. Our polygon-level results for the periods 2000 BCE and 1000 BCE indicate that a reliance on agriculture strongly predicts political integration but that social stratification (the presence of 3 or more social classes) also displays a strong positive correlation with political integration. Although we are not able to pinpoint the exact mechanism through which agricultural communities transformed into states, we argue that our evidence suggests that a reliance on agriculture led to a social division of societies, which in turn led to the emergence of an elite exercising centralized power over a substantial population.

Our study relies on data developed by us in Borcan, Olsson, and Putterman (hereafter BOP, 2014) and BOP 2018, and detailed in the latter's Appendix. Whereas BOP (2018) presented the extended state history data and provided an in-depth analysis of how an accumulated state history affected long-run economic development up to the present day, the current paper focuses on the earlier process through which states emerged with a lag after transitions to agriculture. The key contribution of the current paper is its characterization of the relationship between the transition to agriculture and the rise of states among nearly all modern countries in the world. The analysis of the timing of transition from agriculture to states is considerably more developed here than in the corresponding section of our working paper. For example, we now separately discuss agriculture and state emergence in the six pristine states occupying parts of eight present day countries, we control for distance to origin of relevant state spread zone, spatial autocorrelation, and we add a complementary analysis using grid-cell data for Europe, Middle East and N. Africa, and another analysis using ethno-archaeological zone data for the period 5000-1000 BCE. We also undertake an exploratory analysis of mechanisms discussed in the conventional theories on state emergence (population density, urbanization and social stratification).

Our paper is related to a few existing works on the transition to agriculture and state origins. Petersen and Skaaning (2010) and Boix (2015) estimate correlations between agricultural transition and state emergence, with the former adding supplemental estimates to a previous compilation of state age data by Putterman (2007) that extended to 1 CE, and the latter using dates said to be based on books published in the 1970s through 1990s.⁴ Schönholzer (2019) tests the environmental circumscription hypothesis of Carneiro (1970), proposing that states mainly arose early in environments characterized by a highly productive agricultural core that was surrounded by an inhospitable periphery. Mayshar et al (2019) argue that a necessary condition for the rise of powerful early states was the transparency and storability of the agricultural surplus by a societal elite. While our focus is on the temporal relationship between appearance of agriculture and emergence of states, we briefly review their ideas along with earlier contributions on the channels linking agriculture to states.

⁴The year of first state presence is not shown in the book, nor is any public repository of the data or country by country detailing of sources indicated.

1. Empirical Specification and Data

1.1. Empirical Specification

Our main empirical specification is

$$StateAge_{ij} = \alpha + \beta \cdot TimeAgri_i + \gamma \cdot D_j + \delta' \cdot X_i + \epsilon_{ij}$$

where the dependent variable $StateAge_{ij}$ is the time in millennia (ky) in 2000 CE since the emergence of the first state in a territory defined by the borders of current country i in diffusion region j, $TimeAgri_i$ is the time elapsed since the transition to agriculture in i (ky), D_j is a dummy for state diffusion region j, and X_i is a set of control variables defined below.

Country level units defined over year 2000 borders are used because our research on state age has focused on how early history influences differences in economic and institutional outcomes today, and because comparable estimates have not been assembled for the entire world at grid-cell or other finer levels. *StateAge* is extracted from the *State History Index* developed originally by Bockstette, Chanda, and Putterman (2002) and extended by BOP (2018) to account for states that emerged before the Common Era, as identified to be the case in 58 of 159 modern-day country territories for which data was collected.

Our primary interest concerns the level of the estimated coefficient $\overline{\beta}$, measuring the marginal impact of an additional millennia since transition to agriculture on the time since a state emerged. Furthermore, we will analyze how $\overline{\beta}$ combines with the estimated intercept $\overline{\alpha}$ to determine the expected temporal lag between the transition to agriculture to the emergence of states. If we consider the regression equation above without diffusion region and other controls, we can define the expected temporal lag as:

$$E(TimeAgri_i - StateAge_{ij}) = E(TimeAgri_i) - \overline{\alpha} - \overline{\beta} \cdot E(TimeAgri_i) =$$
$$= (1 - \overline{\beta}) \cdot E(TimeAgri_i) - \overline{\alpha}$$

From this expression, we see that if, for instance, $\overline{\alpha} < 0$ and $\overline{\beta} = 1$, then all countries experience an identical temporal lag between the transition to agriculture and states, given by the absolute value of $\overline{\alpha}$. If, on the other hand, $\overline{\alpha}$ approaches zero and $\overline{\beta} < 1$, then the temporal lag increases with $E(TimeAgri_i)$. Hence, if time since agriculture is small so that the transition happened close to the contemporary era, the transition to states will be shorter than among the countries in the Fertile Crescent with a high $E(TimeAgri_i)$. We are further interested in whether the level of $\overline{\beta}$ varies between internally and externally originated states. In externally originated states, created through an outside intervention, the timing of agricultural transition in the country and the subsequent dynamics of increased domestic social stratification should play less prominent roles, implying a $\overline{\beta}$ that is closer to zero or insignificant.

Although we control for an extensive set of anthropological and geographic variables, we recognize that there might potentially still be omitted factors that influence both the transition to agriculture and the emergence of states. One challenge is related to the changes of state borders over time and possible correlation between the location of these borders and the transition to agriculture. For instance, agriculture and states may have emerged earlier in territories that are naturally circumscribed by geographical features like mountains or deserts, automatically becoming state borders. Moreover, duplicate values for state age and time since agriculture transition in clusters of countries with shared histories could lead to the over-weighting of observations in our OLS estimates relative to the weight given to regions that ended up as parts of larger countries, a problem that could be further compounded if the latter regions' histories were less like one another.

While sub-national data at a global level is unavailable, we go some way towards addressing this concern by using grid-cell level data for the roughly contiguous region that includes Europe, Middle East and North Africa. This data from Harish and Paik (2019) is compiled based on data from the Euratlas for states (covers a restricted period between 1 and 2000 CE) and Pinhasi et al (2005) for agriculture (without restriction on time depth). We describe this data in section 2.3. We also provide a partly parallel analysis using global archaeo-ethnological polygons (rather than units defined by today's country boundaries) and data from 2,000 and 1,000 BCE (from the ACE data of Peregrine, 2003), which we discuss in the Mechanisms section of the paper (Section 3).

Another concern is that we might not be capturing that state formation was likely shaped by a process of geographical diffusion, from the earliest (pristine) states outwards, including via neighbouring territories. All our results therefore include specifications with a measure of distance from the original (pristine) states. However, we go a step further to correct for likely spatial correlation between error terms and between dependent variable values of neighbouring units, to reduce the bias in OLS estimates and standard errors (SARAR models).

1.2. Data on state presence and state age

Data are compiled guided by the conceptions of Service (1960), Tilly (1990), Johnson and Earle (2000), and adopting the convention that political structures from bands to simple chiefdoms fall short of being states, whereas paramount chiefdoms which incorporate multiple individually substantial chiefdoms can be understood as incipient (or proto-) states. A still larger scale including a specialized administration and soldiery is required to qualify as a "full state." The main challenge was to determine the approximate date of state emergence, considering these two thresholds for political development. Along with the scale and specialization requirements mentioned above, BOP (2018) adopted Weber's definition of a state (1978, p.54): an entity which "upholds the claim to the monopoly of the legitimate use of physical force in the enforcement of its order". Decisions on the date of state formation were based on evidence of emerging monopolization of power from archaeological traces of a transition to a scale of organization above the tribal level (e.g., the monumental structures, armies), or from historical records. The main source for data coding was Encyclopaedia Britannica Online, but alternative sources were consulted when information in Britannica was sparse or ambiguous.⁵

For each modern-day country in the sample, BOP (2018) recorded the approximate date of state emergence within its territory and whether or not the manifestation of political centralization can best be characterised as a proto-state or a full state (as well as when it would be reasonable to consider the transition from proto- to a full state complete). The authors found the first presence of a state to have occurred in the form of a paramount chiefdom in present-day Iraq in 3500—3401 BCE, with full state designation beginning there in 3400 BCE.

We employ two dimensions of the state history index compiled by BOP (2018) in our analysis. First, we use time from 2000 CE to first appearance of either a paramount chiefdom or a full state as our main measure of state age, with time to first full state alone as an alternative measure in analyses of robustness. Second, BOP determine whether a country's first state was created by external colonizers versus by internal actors, permitting us to distinguish between internally originated and externally originated states, as mentioned above. To these, we add our identification of pristine states and our assignments of each non-pristine state to the diffusion zone of one or another pristine state, as detailed in Appendix Table A1.

1.3. Data on time of transition to agriculture

Time of transition to agriculture is defined conceptually as approximate year in which a substantial population in some part of a country relied mainly on cultivated crops and domesticated animals for their subsistence. It was compiled by Putterman and Trainor (2006, revised 2018) relying on expert compilations including Smith (1995), MacNeish (1992), and Piperno and Pearsall (1998). We note that first cultivation of individual crops and domestication of animals occurred at considerably earlier dates than we assign for emergence of agriculturally-based society (e.g., South America (Piperno, 2011)), but these items were at first contributing to diets still dominated by foraged

 $^{^{5}}$ All data and replication files for the paper are openly available as Borcan et al (2021)

plants and hunted animals.

1.4. Data on pristine states

Pristine states are identified on the basis of the best archaeological and anthropological evidence to date. We follow, among others, the work of archaeological anthropologist Charles S Spencer (Spencer, 2010, p 7119), which cites locations in Mesopotamia, Egypt, Indus Valley, China, Mesoamerica and Peru as the most likely six cradles of primary state formation ("whereby a first-generation state evolves in pristine fashion, without contact with any preexisting states").

Mesopotamia is indisputably the location of the earliest state. Designation of subsequent pristine states is based on assumptions supported by archaeological evidence. For instance, we presume that the Mesoamerican and Andean civilizations each arose with no direct influence of ideation regarding political structures either from each other or from Mesopotamia. These three fully independent points of state origin are linked by us to the countries most often associated with their initial centers of gravity, i.e. Mexico, Peru and Iraq (our findings would change little were we to substitute, say, Guatemala for Mexico or Bolivia for Peru). In the case of China, although some Fertile Crescent crops had reached it by the time of state emergence there, proto-state building in the East Asian civilizational core around Erlitou arose mainly from local crop and animal packages (Morris, 2010). The Indus Valley cities in what are presently India and Pakistan are also treated as giving rise to states independently of Mesopotamia, despite considerably stronger influence of West Asian agriculture, since signs of direct cultural influence from Mesopotamia are limited. We treat the first state within present-day Iran, on the Susiana Plain, as pristing although not as an independent origin point for state diffusion, because despite influence from contacts with pre-state Uruk, southern Mesopotamia gave rise to states at nearly the same time. Egyptian civilization is viewed as generating macro political structures independently of and only slightly after Mesopotamia, despite the fact that it transitioned to an agriculture based mainly on

West Asian domesticates considerably later than that region had done (Allen, 1997).⁶

1.5. Data on state diffusion from pristine states

To control for potential influence on timing of the gradual spread of (non-pristine) state polities across regions, a process driven not only by conquest and attempts to stave off conquest but possibly also by example, we assign each country to a zone of likely diffusion from a pristine state. Our decisions to assign each country to a particular diffusion zone are guided by a combination of geographical proximity, timing of state emergence, patterns of colonization and conquest, and where archaeological evidence is available, similarity in political organization (collected in the State Antiquity Coding Decisions Appendix that accompanies BOP 2018). We aim to capture primarily the diffusion of institutions (rather than the diffusion of agriculture as a precursor of institutions). For example, the first Mesopotamian states inspired instances of state emergence around the Mediterranean (including North Africa, outside of Egypt, through Phoenician and Greek colonies) and ultimately northward to Scandinavia, Britain and Ireland and southward to Mali. The first (or at least subsequent) Indian states likely influenced the emergence of states in Cambodia, Indonesia, and neighbors (which display similar political structures, known as mandala states). In a similar manner, the first Chinese states arguably inspired the emergence of the early

⁶There are signs of trading contacts between Mesopotamia and the Indus Valley cities, but they followed Persian Gulf/Indian Ocean coastal routes, with no known evidence of gradual dispersal of population eastward to the Indus prior to the Indo-European migrations that largely post-date the early Indus Valley cities. The building plans, pottery styles, as yet undeciphered symbol system, and other artefacts of the Indus cities lead experts to view it as a fundamentally independent culture. Egypt is also regarded as a case of independent political development, which is supported by the evidence that its gradual state formation began with a number of small chiefdoms in the south rather than the north of Egypt (Upper Egypt, ca. 3700-3400 BCE, with Hierakonpolis emerging as the first large urban center ca. 3400-3200 BCE), according to Spencer and other scholars.

states in Korea and Japan.⁷

Distance of each country to the pristine state with which it is identified is given in thousands of km of geodesic distance, from initiation points (as documented by Spencer, 2010) at Uruk (Iraq), Erlitou (China), Mohenjo Daro (Indus Valley), Chavin de Huantar (Chavin, Peru), Monte Alban (Oaxaca Valley, Mexico), and Hierakonpolis (Egypt). Note that state diffusion zones are often similar to, but not fully overlapping with agriculture spread zones. For instance, much of Sub-Saharan Africa (excepting Egypt, Sudan, Eritrea, Djibouti and Somalia) falls into the Western state diffusion region (spreading from Mesopotamia down through Northern Africa), but the adoption of agricultural practices in Central, Southern and Eastern Africa (south of Ethiopia) is tied to the Bantu expansion from West Africa. The substantial though incomplete overlap between agriculture and state diffusion paths makes it unlikely that our measure of distance from pristine states captures only the diffusion of institutional knowledge. Nevertheless, where we could document similarities in political organization between a pristine state and a polity in geographical proximity, we have used this as a key criterion in our assignment decisions.

1.6. Controls

We also include a number of controls X in our estimations to take into account anthropological and geographic characteristics of the territories in our sample, which may influence state emergence. The first is the time (in ky) since the initial uninterrupted settlement by anatomically modern humans (in 2000 CE), which was originally coded by Ahlerup and Olsson (2012) and updated in 2018 following recent developments in

⁷We use the terms "inspired" and "influenced" to convey that the first state in each country of a given diffusion zone was established either by elites familiar with nearby models from within the zone, or by settlers or invaders from a polity within the zone. While need for judgment is sometimes unavoidable, we believe even difficult calls such as assigning Vietnam to the East Asian zone but Cambodia to that of South Asia are defensible on historical and cultural grounds. See Appendix Table A1 for the assignments.

Oppenheimer (2012a, 2012b, 2014). We control for time of first human settlement because both agriculture and states could conceivably have emerged many thousands of years earlier in Africa and the Near East than in (for example) Ireland, Australia, or the Americas, by virtue of later arrival of humans to the latter land masses.

We follow the assumption of Oppenheimer and collaborators according to which anatomically modern humans (AMH) made a single decisive exit from Africa to Eurasia by initially following a southern Asian coastal route, an approach that treats earlier signs of AMH in Fertile Crescent and other sites as largely lacking in longer-term contribution to the AMH gene pool, although the dates we assign to Fertile Crescent countries—52 kya—are earlier than in Oppenheimer (2003, 2012a, 2012b, 2014) and Soares et al. (2009) in recognition of the earlier dates preferred by other experts. See the "Details on Methods" section of the Appendix for further information, including the assumption that earlier AMH appearance in the Fertile Crescent was probably not the decisive long-term exit from Africa. We judge it impossible at present to assign firm dates for individual sub-Saharan African countries, and accordingly use the 135 kya estimate of Oppenheimer (2003) for the entire region, while also confirming the robustness of our qualitative results to adopting a more recent estimate, 90 kya. We also try substituting as an alternate proxy for AMH arrival time the (mainly) land distance from Addis Ababa, used in several studies of long history by economists, on assumption that AMH radiation throughout the world begins somewhere in or near present-day Ethiopia (Ashraf and Galor, 2013).

The geographic controls in X include absolute latitude, an indicator of whether the present-day country is landlocked, distance to coastline and sea-navigable rivers, land suitability for agriculture, mean elevation, temperature, precipitation, and percentage population at risk of contracting malaria. These are conventional controls in the long-run development literature, which account for the main influences for the adoption of agriculture and a sedentary life. We also calculate and control for distances (i.e. the length of the shortest curve) to the relevant pristine state - the nucleus of the diffusion region. All the variables' construction is described in the "Details on Methods" section

of the Appendix. Furthermore, Table A1 in the Appendix displays the data on state age, transition to agriculture, time since first human settlement and state diffusion region for each country in the sample.

In the next section we first present the standard OLS results for pristine and nonpristine states and alternative specifications with sub-country level data.

2. Results

2.1. Agricultural Transition and State Emergence in Pristine States

The six clusters of pristine states widely accepted by anthropologists emerged on the territories of present-day Iran and Iraq, Egypt, India and Pakistan, China, Mexico and Peru. The time before 2000 CE since the transition to agriculture on the territories of these eight countries is strongly predictive of the timing of autochthonous and independent state emergence (Figure 1A), with a correlation coefficient of 0.85. A fitted line emerges very close to the cases of Mexico and Peru, India and Pakistan, Iran and Iraq, indicating that state formation would have occurred around 400 years earlier for each millennium earlier that reliance on agriculture emerged. Given the very small number of observations, it is clear that the significance of the positive slope hinges on the inclusion of Mexico and Peru. In Egypt, the lag between the transitions to agriculture and the presence of state is shorter than predicted by the slope, and vice versa in China. Note that these estimates are based on our definition of state age including the early phase of proto-states. Political institutions come considerably later in India and Pakistan than in China if we go by full state rather than proto-state.⁸

⁸The difference is due to the fact that BOP assign proto-state status to the Indus Valley cities in keeping with the lack of evidence of full state presence in them. The matter appears to remain unsettled, however, with some viewing the large scale and highly organized appearance of these cities as prima facia evidence for state presence.



Figure 1: State age and time since transition to agriculture in pristine (A) and nonpristine states (B).

State age is calculated in BOP (2018). The time since agricultural transition is compiled by Putterman and Trainor (2006 [revised 2018]).

Notes: State age is plotted against the time since the transition to agriculture in millennia (ky) before 2000 CE in both figures. (B) includes three-letter isocodes for individual countries and a thin dashed line to the left showing where time since agricultural transition equals state age. Both figures include a fitted OLS regression line with a 95 percent confidence interval for the predicted mean (grey area). In (A), the estimated bivariate regression equation is given by $StateAge_i = 1.104^{***} + 0.414^{***} \times TimeAgri_i + \epsilon_i$. Whether the relationship in the figure represents a causal link due to the opportunity (or demand) for large scale socio-political organization which agriculture would have gradually created (often over millennia), is less controversial in pristine states, where the transition to agriculture took place on average 3.4 ky before state formation.

2.2. State age and time since agricultural transition in non-pristine states

2.2.1. The relationship between state emergence and agricultural transition

From the agricultural cores, the practice of domesticating plants and animals gradually spread to the periphery of five main regions of agriculture diffusion: West Asia -Europe - North Africa (starting from the Fertile Crescent), Southeast Asia and Oceania (spreading from China), Sub-Saharan Africa (through the Bantu expansion out of the territory of modern-day Cameroon), North and Central America (from Mexico), and South America (starting from the Andes).⁹ Soon after the emergence of pristine states, adjacent territories saw the formation of state institutions and large scale political organization sprung up fast across areas of diffusion which largely (but not entirely) overlap with the agriculture diffusion regions. Some of these states emerged autochthonously (internally originated states), but were (in some cases quite evidently) influenced by pristine state development in those regions. A prominent example is the spread of mandala states from India into Southeast Asia. Other states emerged as a result of expansion and conquest by pre-existing states (externally originated states). The Western state diffusion zone, which started in Iraq (the Fertile Crescent), eventually includes many countries of today that were not home to states in our sense (for instance Malawi, Cuba, and New Guinea) before being swept up in the European colonial era. Internally originated states (e.g., those on the territories of what

⁹Smaller scale pristine domestication events occurred at many other sites (Larson et al., 2014), for example Japan, the Mississippi catchment region of North America, and southern India. But in most cases those regions transitioned to more complete reliance on agriculture after the influx of crops from one of the five regions mentioned (i.e. China, Mexico and the Fertile Crescent, respectively, for the cases of Japan, the Mississippi and India).

are today Sweden, Ghana or Indonesia) emerged on average one millennium earlier than externally originated states, which currently count on average a thousand years of existence.

The positive, bivariate relationship between state age and time since agriculture in 151 non-pristine states is shown in figure 1B. The best fit line intersects the vertical axis below 0, consistent with presence of a lag between adoption of agriculture and emergence of a state. With a slope less than unity, this lag tends to be larger in places where the transition to farming occurred earlier. In for instance Turkey (TUR, upper right corner of figure 1B), states emerged 5 ky after agriculture, whereas the lag was only 0.6 ky in Angola (ANG, lower left corner). In all non-pristine states, the transition to agriculture either preceded or (in a few cases like Seychelles and Somalia) coincided with state formation (mean lag is 2.9 ky), indicating that the prospect of reverse causality is of little relevance.

In Table 1 Panel A we display OLS estimates from our main empirical specification for non-pristine states. The results point to 0.430 ky of earlier state emergence for each additional 1 ky of reliance on farming. We also report these estimates after controlling for the distance of states from their diffusion regions' pristine states (column 2); in a third specification we control for unobserved characteristics of the state diffusion regions through region indicators and we additionally control for geographic and climatic characteristics (column 3). These further controls ensure that our main estimate captures the influence of agriculture on state formation, and not other favourable conditions for institutional development along the diffusion paths from pristine states. We note that the unconditional estimate in column 1 is very similar for pristine and non-pristine states. However, when accounting for the distance to pristine states and other characteristics, the estimate of the marginal impact of agricultural transition timing on state emergence in non-pristine states drops to 0.317 ky.¹⁰ Thus, regardless

¹⁰The estimates in the common sample of 137 observations are slightly larger than the counterparts in columns 1 and 2 of Table 1 Panel A (0.448 and 0.380).

Figure 2: State age and time since the transition to agriculture in non-pristine internally versus externally originated states.



Notes: State age is plotted against the time since the transition to agriculture in 2000 CE. The figure includes separate fitted OLS regression lines for internally originated (solid line) and externally originated states (dashed line). Thin line to the left shows where time since agricultural transition equals time since state emergence. The time since state emergence is calculated in BOP (2018). The timing of the transition to agriculture is compiled by Putterman and Trainor (2006 [revised 2018]).

of whether state formation ensued independently or through conquest, closer or further away from the original cradles of civilization, the evidence suggests earlier reliance on agriculture expedited state emergence: states emerged at least 300 years faster for each millennium earlier that reliance on agriculture began.

Even if these results are not affected by reverse causality, there remain concerns that border definitions and factors common to the spread of agricultural practices and large scale political centralization may bias the OLS estimates. We address these concerns below in sections 2.3 and 2.5.

Finally, we compare the estimates in internally originated states with those in externally originated states. We use the term "internally originated" state for those territories where the first ever recorded manifestation of centralized power above the tribal level was a result of political development within the territory. By contrast, externally originated states had their first supra-tribal rule imposed from outside the territory (through conquest or colonization). In coding the BOP (2018) data we have carefully reviewed the evidence for the internal or external origination of states in the BCE, and we have taken a cautious approach by flagging cases where early domestic rule was mixed with external influence (most of these cases are treated as externally originated states). However, despite our best efforts, there remains a risk of misclassification of states, particularly external origination miscoded as internal. We discuss the implications of this type of error below after we show the results.

		State Age	
Panel A: All non-pristine states	(1)	(2)	(3)
Time since agriculture (ky)	0.430***	0.366***	0.317***
	(0.025)	(0.030)	(0.041)
Distance to pristine state		-0.085***	-0.130***
		(0.019)	(0.043)
Time since first human settlement (ky)		0.000	-0.005**
		(0.001)	(0.002)
Constant	-0.403***	0.255	0.526
	(0.108)	(0.230)	(0.429)
Observations	151	151	137
R-squared	0.609	0.644	0.806
Panel B: By internally and externa	lly originat	ed states	
Time since agriculture (ky)	0.328***	0.286***	0.252***
	(0.035)	(0.037)	(0.046)
Internally originated x	0.117^{***}	0.101***	0.077^{***}
time since agriculture	(0.030)	(0.033)	(0.026)
Constant	-0.262**	0.312	0.331
	(0.101)	(0.211)	(0.380)
Observations	151	151	137
R-squared	0.652	0.683	0.826
Geographic Controls	No	No	Yes
Diffusion Region	No	No	Yes

Table 1: State Age and Transition to Agriculture in Non-Pristine States - OLS

Notes: The table presents OLS regression estimates of the relationship between time since state emergence as of 2000 CE, and time since transition to agriculture in 151 countries, excluding 8 countries identified as places of emergence of pristine states. For a description of state age, time since transition to agriculture in 2000 CE and time since first human settlement, see the note to Table 1. Distances to pristine state are calculated as the length of the shortest curve between the centroid of each country and the centroid of its assigned pristine state (the region from which state diffusion into the territory of the country in question is most likely to have originated). The estimates in Panel B columns 2 and 3 include controls for distance to pristine state and time since first human settlement. The geographic and climatic controls and historical variables' construction is detailed in the Additional data subsection of the Appendix. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

We may expect a weaker relationship between agriculture and state emergence in externally originated states, if the expansion of power or conquest of new territories was driven by factors orthogonal to the time since agriculture became the dominant food source (e.g., opportunities to expand in ungoverned territories, mineral resources or difference in military technologies). On the other hand, territories not yet politically organized, but where agriculture was the main mode of food production may have been more attractive, thus becoming earlier targets for expanding states. In either case, there can be vastly different (and potentially arbitrary) processes of state formation in externally-originated states, so it is ex ante unclear what the time lag between agriculture and states should be. We believe it is reasonable to think that if agriculture is one of potentially many more factors leading to political integration in these states, its relationship to state formation is weaker. Below we undertake an exploratory analysis to document any difference across the two types of states.

The simple association between state age and time since the transition to agriculture is positive and significant for both types of states, but slightly weaker in the externally originated states, where the line fitted through the scatter of cases has a flatter slope and a lower R-squared value (Figure 2). The OLS estimates in Table 1 Panel B confirm that there is a significantly higher magnitude of the influence of agriculture on state transition in internally originated states, as seen from the coefficient of the interaction term between internal status and timing since agriculture transition. The magnitude of the slope of agriculture is 0.252 - 0.328 in externally originated states compared to 0.329-0.445 in internally originated states (implying that a state emerged about 290 years later for each millennium of delay in agricultural transition, for the former group, versus about 390 years later for each millennium in the latter group). The agriculturestate lag is on average 3.1 millennia in internally originated (including pristine) states, and 2.7 millennia in externally originated states. The correlation coefficients are both significant, but they also differ, having the values 0.66 and 0.83 for externally and internally originated states, respectively. Most of the correlation is probably driven by the early conquests (well before the European colonization era commenced around

1450 CE), for example the annexations of the Roman Empire. It is worth emphasizing that internally originated states have on average emerged before the Common Era (mean state age is 2.19 millennia), while externally originated states are much younger (mean state age 0.97 millennia). With over half of the states established autonomously early on, the arrival of political rule in the remaining territory through diffusion or imposition seems to have diluted the role of early agriculture. In fact, for countries that experienced their first state presence during the era of European colonization, there is no correlation between the time of first agriculture and that of first state.¹¹ Given the weaker relationship in externally originated states, misclassification of states as initiating within rather than being imposed from outside the territory, would lead to a downward bias in the beta estimate and a weaker correlation between agriculture and state timing in internally originated states.

Nevertheless, these results suggest that while earlier mechanisms linking agriculture to state formation became weaker over time, they endured for much of history, and it was only in the last 550 years (since the radical changes in overseas transport and communication, the most recent 10 percent of the overall history of states) that these channels were muted by novel processes of state diffusion (as the post-1450 CE results in Appendix tables A11 and A12 suggest).

¹¹See appendix table A12, where we estimated regressions separately for 18 countries having states externally originated in 1-1450 CE and 40 states having ones externally originated after 1450 CE, paralleling those of Table 1, panel A. While the slope estimate is large and significant for the countries with states externally formed before 1450, it is insignificant for those formed in 1450 and later (and the correlation coefficient in the latter case is close to zero). The latter result appears related to Ertan et al. (2016)'s finding that in the era of European colonization, places having earlier established agriculture and longer pre-existing states were less likely to be colonized at all and if colonized, tended to be colonized later (after improvements in European military capabilities and abilities to withstand tropical diseases).

2.3. Results from grid-cell data

For a partial sub-national analysis, we rely on data from Harish and Paik (2019), who coded state presence between 1 and 2000 CE and date since the transition to agriculture for a fine grid of rectangular cells within Europe, the Middle East and North Africa. The authors divided the region into about 2,400 arbitrary rectangular cells (of approximately 77 km \times 62 km) with land mass present. They used the Euratlas century-by-century geocoded maps of political history of Europe (and parts of the Middle East and North Africa) to determine whether each grid-cell had a polity ruling over at least 0.1 square kilometers of its area for each century in the Common Era. Thus, for each grid-cell they compile a variable that retains the number of centuries of state presence until 2000 CE. This is the measure we use as our dependent variable (rescaled to millennia BP), as the closest proxy to our country-level state age measure. We note that it is an imperfect proxy because states in existence before 1 CE (such as those in Italy, Greece and the Levant) have their presence capped at 20 centuries. Also, grid-cells may have a polity present in the early centuries, followed by no rule for some time, followed by the return of political organization. With these caveats in mind, we compare the grid-cell level results with the results with our country-data for the subsample of countries of this region.

The grid-cell data on agriculture transition dates is compiled by Harish and Paik (2019), based on a sample of 765 Neolithic sites (C14-dated) in the Near East and Europe from Pinhasi, Fort, and Ammerman (2005), also mapped by Olsson and Paik (2020). This is the average number of millennia (across sites) from the transition to agriculture to 2000 CE within each cell.

In the grid-cell level regressions we control for geographical factors (agriculture suitability, distance to the nearest coasts or rivers, mean elevation) and our own calculated distance to Mesopotamia (the pristine state for this region), to keep the specifications as comparable as possible to our cross-country regressions.¹²

¹²Note that we do not include latitude as a control in the grid-cell level analysis, because of near

The results are displayed in Table 2 (while appendix table A3 displays analogous results for our cross-country sample covering the same region as in Harish and Paik, 2019). Specifications 2 and 3 include modern-day country fixed effects. The grid-cell level estimates of the effect of time since agriculture are positive and significant at 1 percent significance level. They suggest that one millennium of agriculture is associated with one century of earlier state presence. These estimates are likely underestimating the true association, because the available state presence data is capped at maximum 20 centuries.¹³ The country-level estimates for this region are larger, and comparable with the estimates in Table 1 for the entire sample of countries. Overall, the results provide additional support for the hypothesis that early transition to agriculture expedited state emergence.

2.4. Heterogeneity of effects across time

One question that merits further exploration is whether the association between time since the Neolithic transition and state emergence is stable, or whether it changes over time. Figure A4 in the appendix displays the time lag since the transition to agriculture until the first date of state emergence, plotted against state age in 2000 CE. While there is a large variation in lag values at all levels of the state age, there is a large concentration of countries with young states (below two millennia) and smaller lags (below three millennia) and also a cluster of old states with lags of over four millennia (generating the appearance of a slight positive association, which is confirmed by a correlation coefficient of 0.34). The line of best fit in the figure has a slope of 0.38 significant at 1 percent significance level, suggesting that every additional millennium of state presence is associated with nearly four centuries longer lag between agriculture and state onset, on average.

perfect correlation between latitude and time since the emergence of agriculture.

¹³Up to 38.6 percent of the grid cell observations are upper-censored. Indeed, Tobit regressions accounting for these observations yield larger coefficients in specifications including the same controls as those in columns 2 and 3 in Table 2, around 0.150 (results available on request).

	(1)	(2)	(3)
Number of centuries	state pres	ent in grid	-cells
Time since Agriculture (ky)	0.292^{***}	0.095^{***}	0.103^{***}
	(0.010)	(0.015)	(0.015)
Distance to Mesopotamia		-0.000***	0.000
		(0.000)	(0.000)
Observations	2,277	$2,\!277$	2,223
R-squared	0.232	0.841	0.856
Controls	No	No	Yes
Country FE	No	Yes	Yes

Table 2: Sub-country analysis

Notes: The table presents grid-cell level OLS regression estimates of the relationship between the number of centuries of state presence 1-2000 CE, and time since transition to agriculture in Europe, Near East and North Africa, using data from Harish and Paik (2019). Distances to pristine state are calculated as the length of the shortest curve between the centroid of each cell (country) and the centroid of its assigned pristine state. The geographic and climatic controls and historical variables' construction is detailed in the Additional data subsection of the Appendix. All regressions include intercepts. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The millennium-by-millennium state samples are too small to permit a precise estimation of the association between duration of reliance on agriculture and timing of state emergence for every thousand years of state history. Therefore we propose to examine the relationship separately for the BCE and CE periods. The majority of states (101 in our sample) emerged in the CE, when state formation occurred at a much faster rate than in antiquity (the lag between transition to agriculture and the emergence of states is 2.5 millennia on average for countries in which states emerged after 1 CE, and 3.5 millennia on average for ones in which states emerged BCE).

Table A4 in the appendix displays specifications analogous to those in Table 1, separately for countries in which states emerged before 1 CE (columns 1-3) and those in which they emerged after 1 CE (columns 4-6). The estimates of the time since agriculture coefficient are significant and positive in both sub-samples, but differ starkly in magnitude. Agriculture is associated with earlier state emergence in the BCE period (up to 400 years per millennium earlier agriculture in the most complete specification, compared to 56 years earlier in the sample in which states emerged after 1 CE). Alternative specifications with agriculture interacted with the CE era indicator show that this difference is significant at 1 percent level (the coefficient on the interaction is -0.287, while that of agriculture itself is 0.358). The results suggest that the time elapsed since the emergence of agriculture continued to influence the process of state formation in the more recent history, but that this influence diminished over time.

As polities emerged and expanded around the globe, new dynamics (such as increased competition over land and longer-distance maritime trade and colonization) spawned large scale political organization in the remaining territories. This is in line with the results on externally originated states, which also display a weaker association with agriculture, and which are on average younger than internally originated states (table 1 panel B).¹⁴

 $^{^{14}}$ As an added check, Table A11 divides the 101 countries in which a non-pristine state appeared after 1 CE into those with first state dated 1 CE to 1450 CE and after 1450, respectively; the correlation

2.5. Robustness checks

We now turn our attention to the issue of diffusion of state institutions across space and how this might affect our estimates. Political and economic transformation certainly travels across neighbouring geographic regions. State institutions may emerge indigenously in some regions and prompt further autonomous development in proximate regions (for example as a defense response), or they may be imposed by external powers expanding their conquest into adjacent territories. Whatever the case, in many regions of the world we see geographic clusters of territories with similar times of state emergence. For instance, North African regions saw their earliest large scale polities in the 8th century BCE when seafarers from Phoenicia established Carthage and semi-independent colonies which gradually expanded on the shores of the Mediterranean in what are today Libya, Tunisia, Algeria and Morocco. These spillovers violate the standard OLS assumptions, something which merely controlling for distance from pristine states cannot mitigate, because the latter does not capture the influence of all neighbouring territories. We therefore estimate spatial dependence models for all states where we allow for spatial autocorrelation in the dependent variable and in the disturbances (SARAR). In these models, spatial lags of the dependent variable enter the model explicitly, through a matrix of weights on the time since state emergence in other countries. The weights are inversely-proportional to the distance from the centroid of each country to other countries, to indicate that developments in further away places are less likely to affect political change in any given region.¹⁵

In the SARAR estimates in Appendix Table A6 the coefficients on years of agriculture are positive and significant at 1 percent level (we include the entire sample of

of time of agricultural transition and time of state emergence is insignificant in the latter group.

¹⁵Specifically, SARAR models are $y = \lambda \mathbf{Wy} + \mathbf{X\beta} + \mathbf{u}$, with $\mathbf{u} = \rho \mathbf{Mu} + \epsilon$, where \mathbf{W} and \mathbf{M} are $n \times n$ spatial weighting matrices, n is the sample size, λ and ρ are scalars, and \mathbf{Wy} and \mathbf{Mu} are $n \times 1$ vectors representing spatial lags. We estimate a model where $\mathbf{W}=\mathbf{M}$ is a matrix with diagonal elements equal to zero, and off-diagonal elements representing the inverse great-circle distances between geodesic centroids.

countries). Modelling spatial spillovers yields estimates around 0.320-0.417 ky earlier states for each millennium of earlier agriculture.

Apart from the issue of spatial spillovers, we also performed other sensitivity checks (results in the appendix), where we include: redefining State age to be based only on the date when a full state was established, rather than a proto-state (Tables A2 and A7); an interaction term that captures how state diffusion speed (proxied by distance to pristine state) may vary depending on how early the pristine state made the transition to agriculture (Table A10); using the original, unrevised, version of the data on time since first human settlement of Ahlerup and Olsson (2012) (Table A9). As an additional strategy to mitigate the issue of omitted variable bias, we also ran two-stage least squares regressions, where in the first stage, the *geographic and biogeographic endowments* (domesticable plants and animals) compiled by Hibbs and Olsson (2004) are used as instruments to predict the time since transition to agriculture, and in the second stage, the resulting values are themselves used as predictors of time since state emergence. The time since agriculture coefficients displayed in appendix Table A5 are slightly smaller than in Table 1 Panel A, falling in the range 0.28-0.35.

We also report regressions where we proxy patterns of initial human settlement by the migration distance from East Africa (appendix table A8). We find consistently significant and similar estimates of the impact of agriculture timing on state formation.

3. Mechanisms

In this section we discuss the main arguments in the literature for the nexus between early reliance on agriculture and state formation. Data availability limits how much we can tests these mechanisms. Nevertheless, we undertake an exploratory analysis to understand whether some popular narratives mediate fully or partially the effect of agriculture on state emergence.

3.1. Literature

Scholars from across the disciplines studying the evolution of human societies have established that the prehistoric transition to agricultural technology as a main source of food production was associated with a sedentary lifestyle, surges in population density, and the emergence of social classes (Diamond, 1997; Johnson and Earle, 2000; Peregrine et al, 2007).

In some stratified societies, a dominant class was capable of controlling a surplus from food production based on agriculture and establishing a monopoly over violence for the purpose of defense and rule enforcement (Weber, 1919). Different (not necessarily mutually exclusive) theories of state formation highlight the varying roles of this elite including: defense against "roving bandits" (Olson, 1993), domination and conflict (Gennaioli and Voth, 2012), taxation and the provision of public goods such as defense forces, irrigation systems, city walls and temples (Dal Bo, Hernandez and Mazzuca, 2015; Wittfogel, 1957; Besley and Persson 2010), and proliferation of trade (Bates, 1983; Fenske, 2014).

All these theories fundamentally link the emergence of a political elite back to the shift to reliance on agriculture (which almost always occurred first), but they highlight different pathways. For instance, some scholars argued that population pressures gave rise to conflict, in turn spawning institutions capable of containing it (e.g., Johnson and Earle, 2000). An alternative explanation for the emergence of social stratification lies in the gradual enclosure (tantamount to claiming property rights) of more and less productive lands, where better-off residents of more productive lands hire labour from worse-off outsiders, setting the premise for social inequality (Dow and Reed, 2013). Other scholars singled out specific environmental or technological constraints that created favourable conditions for social stratification and political centralization. Most theories have been probed through casual observation or selected case studies. To date, very few studies have provided large-sample causal evidence of these mechanisms.

Robert Carneiro proposed the environmental circumscription argument that early autonomous agricultural communities had no outside options to escape exploitation or conflict in productive territories confined within natural barriers (mountains, desserts, sea). The entrapped were therefore easily dominated by powerful elites from within their territory or by neighbouring competing communities. Carneiro (1970) offers the example of the contrast between the Amazonian basin, where vast expanses of sparsely populated productive land allowed population movements in response to conflict over territory, and the Peruvian coastal valleys flanked by mountains and the sea, where mounting population pressures gave rise to early political integration. Schönholzer (2020) undertook the first large scale empirical test for this hypothesis, using archaeological sites across the world to show that circumscribed agricultural territories saw earlier transition to statehood. Moreover, pre-state population density was shown to have little impact on the effect of agricultural circumscription, and urbanization was shown to follow state emergence. Schönholzer suggests that the temporal sequence consistent with his results is one where conditions favorable to agriculture in an environmentally circumscribed or constrained setting led to migratory pressures, conflict and state formation, and subsequent population growth and urbanization.¹⁶

Other studies have focused on the facilitation of taxation as a catalyst in the emergence of social hierarchy with a dominating elite. Mayshar et al (2019) propose and test a model where the emergence of productive, grain based agriculture and of elites adept at appropriating a substantial share of the harvest gave rise to storable food surplus. It was thus the storability of the crops, rather than overall land productivity, that kick-started the process of social stratification and the emergence of the political elites with incentives to invest in fiscal capacity and the growth of their polities.

Mayshar et al challenge the circumscription hypothesis based on examples of states that developed in non-circumscribed territories where cereals were the main crop (e.g.,

¹⁶See also Mayoral and Olsson (2019) who present a dynamic version of the environmental circumscription model and show how weather shocks in the core and periphery of ancient Egypt affected political stability between 2686-1150 BCE. Dow and Reed (2018) present a related view of developments in southern Mesopotamia.

maize in the Mayan states), as opposed to the Amazon basin, where tubers were dominant. However, while they make a compelling point in comparing state emergence across different bio-geographies, examples can be found to support the hypothesis that environmental circumscription made a difference in the onset of states in territories with comparable crops. Both the Mayan states and the pristine Mesopotamian states derived their wealth from cereal crops, but the circumscribed Mesopotamian polities emerged millennia before the Mayan states, a fact potentially consistent with circumscription theory although perhaps alternatively explained by later emergence of agriculture itself.

The main contribution of our paper is to adopt a fully global view in examining the timing of state formation in relation to the variation in time since the transition to agriculture. In so doing, we account for the conventional channels that state formation has been attributed to: population density, urbanization and social stratification. Our aim is not to test or validate the recent arguments in the literature on circumscription and appropriability of agricultural surplus, but the results presented below are compatible with both these mechanisms.

3.2. Data

For this analysis we use data on prehistoric societal changes from the Index of Cultural Complexity (described in the Atlas of Cultural Evolution, Peregrine, 2003, henceforth ACE, based on the methodology of Murdock and Provost, 1973). The index includes ten 3-point scales (on writing, fixity of residence, agriculture, urbanization, population density, technological specialization, land transportation, money, political integration and social stratification).¹⁷

We select political integration, agriculture, urbanization, density of population and

¹⁷This dataset has been explored by economists mainly for compiling data on technology in prehistory. See e.g., Comin et al's (2010) technology adoption index. To our knowledge, this data has not been explored for constructing regional or country measures of urbanisation, population density, social stratification and political integration in antiquity.

social stratification, because these are the main variables and conventional channels hypothesized in the literature on the link between reliance on agriculture and state emergence.¹⁸ These are coded as follows in ACE: Political integration is 1 for autonomous local communities, 2 for integration one or two levels above community and 3 for three or more levels above community. Agriculture is coded 1 if there is no agriculture, 2 if agriculture was a secondary source of subsistence and 3 if it was a primary source. Urbanization is coded 1 if communities had fewer than 100 people, 2 for 100-399 people and 3 for 400 or more people. Population density is 1 for less than 1 person/square mile, 2 for 1-25 persons/square mile and 3 for 26+ persons/square mile. Social stratification is 1 for egalitarian societies, 2 for societies with two social classes, and 3 for societies with three or more social classes.

The advantage of this data is that it is coded for "archaeological traditions" rather than modern-day countries. These "archaeological traditions" were designed to serve as basic units of analysis for archaeoethnology, and are defined "as a group of populations sharing similar subsistence practices, technology, and forms of socio-political organization, which are spatially contiguous over a relatively large area and which endure temporally for a relatively long period" (Peregrine 2003). These traditions are described in the *Encyclopedia of Prehistory* (Peregrine and Ember, eds. 2001). The ACE archaeological traditions dataset lists all traditions (distinct geographical polygons that shift over time), with their approximate start and end date. The ACE presents maps of archaeological traditions (approximate geographical coverage of each tradition), for every thousand years between 10000 BCE and 1 CE. For each archeological tradition's unique polygon on the map, the associated dataset contains the values of the underlying cultural complexity scales.

¹⁸Moreover, the other ACE scales are associated with some of our selected variables: the fixity of residence is usually implied by reliance on agriculture; money and writing are widely viewed as correlates of state emergence; technological specialization is highly correlated with the other variables but not suggested to be a channel from agriculture to states in the literature.

To get a sense of the data, consider the archaeological tradition identified as number 3075 ("European Megalithic"). This tradition started around 4000 BCE and ended around 2500 BCE and had the following values for the cultural complexity scales: agriculture was the primary source of subsistence (value 3), largest settlements less than 100 people (urbanization value 1), population density classified as 1-25 per square mile (value 2), social stratification based on two social classes (value 2), and political integration one or two levels above the community (value 2). For each case (including this example), we geocoded the archaeological tradition polygons for 2000 BCE and 1000 BCE. For each prehistoric period we may have different polygons, depending on the geographic coverage of successive archaeological traditions. Note that when a tradition enters the historical record (written records), the cultural complexity scales are no longer recorded. Thus, from 1000 BCE, parts of South Europe and North Africa linked to the Roman expansion are omitted from the dataset. For these polygons in 1000 BCE, we imputed maximum values for the cultural complexity scales.

We then adopted a set of coding conventions to make the analysis and interpretations more intuitive:

1) For individual archaeological tradition polygons (between 65 and 75 in each period) we recoded the scale values to take values 0 (instead of 1), 1 (instead of 2) and 2 (instead of 3). Then, for each polygon and scale, we created a dummy for value 1 and a dummy for value 2. The only exception is a new dummy for political integration, which is 1 if there is political integration above the community level (values 2 and 3 in the original political integration scale) and 0 otherwise. We code political integration in this way to account for gradually emerging polities. We use these dummy variables in our polygon-level regressions. This approach avoids the shortcomings of working with modern-day country borders.

2) For analysis at the level of countries of today (i.e., circa 2000 CE), we aggregated the 0-1-2 recoded scales of all polygons accounting for some of the current country's territory, for each period. By overlapping the ACE polygons with modern-day country borders, we obtained around 350 sub-country polygons. When more than one such polygon occupies space within the borders of the modern country, we took a simple average of the values of each scale across the relevant archeological traditions within each country, and then divided the result by 2, to obtain an average score between 0 and 1. We use these average scores in country-level regressions.

Appendix table A13 presents the polygon proportions for the binary cultural complexity indicators for 2000 BCE and 1000 BCE. Of the sample, 52.7 percent of polygons had some political integration above the community level in 2000 BCE, rising to 54.5 percent in 1000 BCE. Agriculture as a primary source for subsistence was present in 44.6 percent of polygons in 2000 BCE, rising to 47 percent in 1000 BCE. Large scale urbanisation (over 400 people) and high population density (over 26 people per square mile) were present in about 16.7 percent and 7.6 percent of polygons respectively in 1000 BCE, whereas social stratification (2 classes and above) had reached over 45 percent of territories by 1000 BCE. This is suggestive of a closer link of political integration on one side to agriculture and social stratification on the other, than to the other factors. This is further confirmed in appendix table A14, which displays the pairwise correlations between political integration and all scales, and between agriculture as a primary source and all scales, for 2000 and 1000 BCE. Political integration is most strongly correlated with the presence of agriculture as a primary source (up to 0.85correlation in 2000 BCE), and with social stratification (two classes, correlation coefficient 0.63), followed by an intermediate level of correlation with population density (nearly 0.56). Agriculture is in turn strongly correlated with social stratification (up to 0.64 in 1000 BCE), and also with population density and large scale urbanisation. The patterns do not suggest specific causal links between the different cultural evolution indicators and political integration, but agriculture emerges as the strongest correlate, which also links tightly to social stratification. Below we include these indicators in polygon- and country-level regressions to see which correlations persist when other factors are accounted for.

3.3. Results

3.3.1. Timeline: agriculture, state and ACE cultural complexity indicators

Our analysis focuses on the data on archaeological traditions for 5000 BCE, 3000 BCE, 2000 BCE and 1000 BCE (while the ACE data maps archaeological traditions for 10000 BCE and each millennium between 8000 BCE - 0 CE). We followed a parsimonious approach in extracting the data, selecting 3000-1000 BCE as the most relevant period to our analysis, because the first states according to our preferred definition emerged between 3500 and 3000 BCE. According to ACE's less strict definition of state, political integration at three or more levels of hierarchy above community level was only present in eight archaeological traditions in 3000 BCE, and only in one - the Ubaid tradition in Mesopotamia - in 5000 BCE. Accordingly, we also examine the data for 5000 BCE to capture early pre-state political and other societal developments in their temporal sequence at and shortly after the adoption of agriculture.

To shed light on the evolution of agriculture and ACE cultural complexity indicators (urbanization, population density, social stratification and political integration), we display the shares of territories of countries-of-today hosting traditions which had achieved a certain level of technology or complexity at each time period.

Figure 3 displays the country-level mean values for the cultural complexity indicators across ACE tradition polygons within the boundaries of countries-of-today, at each of four points in time. For each of the indicators, the closer the mean score is to 0, the closer is a society to its initial form (e.g., no stratification, integration or urbanization). Values around 0.5 reflect a society where agricultural technology or cultural complexity is emerging (e.g. agriculture is a secondary food source, stratification only displays two classes, etc.). Values close to 1 indicate complex organization (e.g., society is stratified in two classes or more, etc.) and a mean of 1 for agriculture would mean it has become a primary food source everywhere in the world. In 5000 BCE, the mean agriculture score across countries is 0.2, more than trebling by 3000 BCE, and going up to 0.79 in 1000 BCE, meaning that most countries-of-today had agriculture as a primary food source on most of their territory by 1000 BCE. By contrast, all other indicators start off at very low levels (means at or below half of the agriculture mean) in 5000 BCE, and increase, but with averages remaining at or below the intermediate value 0.5 in 1000 BCE (with the highest being political integration and stratification, mean 0.48). Urbanisation and population density seem to lag behind social stratification and political integration, although the two sets of measures are not fully comparable in scale. We add the share of countries with state presence as per our own definition (BOP) between 3000 and 1000 BCE, for comparison. The share rises from 4 percent to 15 percent of the sample, reflecting that our threshold for political integration is much higher compared to the ACE one.

The average values are not fully indicative of the temporal sequence of transition to full-fledged agriculture and high levels in other complexity scales. To capture the early marked developments within countries-of-today, we focus on identifying countries where at least one ACE tradition polygon had achieved the highest level of development by a specific point in time. Figure 4 displays the share of countries-of-today with top-level values, separately for each cultural complexity indicator and period. By 5000 BCE, 21 percent of territories of countries-of-today had at least one region with agriculture as a primary source. This share rises to 67 percent, 70 percent and 77 percent in 3000, 2000 and 1000 BCE. By contrast, the share with high level of stratification (three classes and above) rises from 3 percent to only around 23 percent by 1000 BCE. Urbanization (communities above 400 people) has a similar evolution, with a higher rise (reaching 31 percent of country territories) by 1000 BCE. The shares with top level political integration are similar to those using our definition of state (16 percent by 1000 BCE).

The evidence supports a sequence in which agriculture developed and spread predominantly before the other societal developments being tracked took place. Stratification had on average a broader early expansion than urbanization, but the levels are comparable by 1000 BCE. Political integration closely mirrors the evolution of stratification.



Figure 3: Mean scores for agriculture and cultural complexity indicators (ACE) within countries-of-today

Notes: The bars represent the mean of the agriculture indicator and several cultural complexity indicators across ACE polygons within boundaries of countries-of-today. The indicators take values 0, 0.5 and 1 for no, intermediate level and high level of each indicator (e.g., no social stratification, two classes and three or more social classes)



Figure 4: Maximum level agriculture and cultural complexity indicators (ACE) within countries-of-today

Notes: The bars represent the share of countries-of-today with top-value complexity in each indicator in at least in one ACE polygon. E.g. in 5000 BCE, 21 percent of all territories of countries-of-today had agriculture as a primary food source (at least in one ACE polygon within that territory).

3.3.2. ACE cultural complexity polygon data

Table 3 displays OLS regressions (linear probability models) at ACE tradition polygons level for 2000 and 1000 BCE, with the ACE-based political integration indicator (value 1 if there was some political rule above the community level) as the dependent variable. The presence of agriculture as a primary source is the main explanatory variable, and successive sets of cultural complexity factors are included (urbanization in column 3, population density in column 4, social stratification in column 5, and all 3 alongside agriculture in column 6). In columns 5 and 6 we include an interaction term between primary agriculture and stratification (three and above classes) in order to account for a potential moderation of the agriculture effects on state development by social stratification.¹⁹ Panel A shows the results for 2000 BCE, and panel B for 1000 BCE. All except the first specification include controls for the polygon centroid latitude and distance from the centroid to the pristine state closest to the polygon, as well as continent fixed effects. From 1 CE onward, much of the world entered the historical era and is omitted from the ACE dataset.

The most striking result is that, regardless of the cultural complexity indices included, as well as geographic and country controls, the significant correlation between the presence of agriculture and political integration persists. It is very large in 2000 BCE, particularly in the specifications with successive sets of ACE complexity indicators (up to 80 percent more likely for a territory to display political integration when agriculture is dominant, in specifications including controls). The effect drops to 51.3 percent in the specification including urbanization, population density and social stratification, but remains significant at 5 percent level. These individual factors in turn also correlate positively and significantly with political integration, but only social strati-

¹⁹Similar interaction terms between the primary agriculture indicator and indicators for the highest levels of urbanization and population density are omitted because all polygons achieving the highest levels of urbanization for the historical period under consideration have agriculture as the primary source of subsistence.

fication persists in a horse-race against the others in 2000 BCE, with strikingly high coefficients (nearly 87.6 percent higher incidence of political integration where society was divided into at least 3 classes already in 2000 BCE). , the agriculture-stratification interaction term is significant at 5 percent level, negative and almost the same magnitude as the coefficient of agriculture, indicating a substitution effect between the presence of agriculture and that of advanced social stratification. In polygons with three and above social classes, the correlation between political development and agriculture as a primary source is reduced to zero. This is indicative of the moderating role of social stratification in the historical process of state formation.

In 1000 BCE, the association with agriculture as dominant food source remains high and persistent (72 percent higher incidence of political institutions above community level). The coefficient of advanced social stratification exceeds that of agriculture in column 6, while rising population density and urbanization displays no significant link to political integration. Here again we note the large negative and significant coefficient of the interaction term between agriculture and social stratification. It suggests that agriculture is no longer the decisive factor in regions having achieved a high level social stratification by 1000 BCE, although it is likely to have been an underpinning of stratification, which now seems to play a more decisive role; it also indicates that agriculture continues to be decisive in those regions without (or with barely emerging) social stratification. Note that if we redo the estimations with a dependant variable redefined to capture exclusively the presence of political integration at 3 or more levels above community (around 15 percent of the sample of polygons in both periods), population density emerges as a significant predictor even in full specifications, while the coefficient of agriculture markedly drops (see appendix table A15). This suggests that population growth (ultimately made possible by extension, intensification, and improvements in agriculture) is associated with political centralization at later stages in the state formation process.

The results suggest that the rise of social classes, entailing economic and social specialisation, was a likely proximate channel from the Neolithic food production revolution to the emergence of political elites. At these points in the ancient history of political institutions, population density and urbanization appear secondary to social stratification. The results are consistent with the theories that propose that a food surplus generated by agriculture, leading to population growth and further resource pressures, did not single handedly trigger political centralization. A fundamental transformation of society based on some initial social differentiation was probably critical for the emergence of institutional structures for organizing ever-larger scale communities. This social transformation does not fully mediate the impact of agriculture, which points to alternative channels not explicitly modelled here. The extant literature suggests other channels could include the type of food surplus (appropriability) and taxation (fiscal capacity), as argued by Mayshar et al (2019) and Schönholzer (2020) among others, which may explain the persistent effects of agriculture found here.

Panel A Political integration above community 2000 BCE Agriculture primary 0.886*** 0.798*** 0.625*** 0.590*** 0.513 2000 BCE (Agri3) (0.055) (0.105) (0.125) (0.205) (0.21)	**
Agriculture primary 0.886^{***} 0.798^{***} 0.678^{***} 0.625^{***} 0.590^{***} 0.513 2000 BCE (Agri3) (0.055) (0.105) (0.125) (0.127) (0.205) (0.21)	** ~\
$2000 \text{ BCE (Agri3)} \qquad (0.055) (0.105) (0.125) (0.127) (0.205) (0.21)$	۲)
	5)
Agriculture secondary 0.219 0.172 0.115 0.101 0.075 0.04	1
$2000 \text{ BCE} \qquad (0.204) (0.171) (0.150) (0.130) (0.098) (0.098)$	1)
Urbanization > 400 0.160* -0.02	21
people 2000 BCE (Urb3) (0.081) (0.05)	6)
Urbanization > $100-399$ 0.179^* 0.00	8
people 2000 BCE (0.093) (0.06)	4)
Pop density $26+2000$ BCE 0.158 0.08	1
ppl/sqm 2000 BCE (Dens3) (0.098) (0.08	4)
Pop density 1-25 0.214* 0.09	9
ppl/sqm 2000 BCE (0.111) (0.08)	5)
Social stratification $3+$ 0.901^{***} 0.876°	***
classes 2000 BCE (Stra3) (0.055) (0.07)	9)
Social stratification 2 0.392^* 0.38	5*
classes 2000 BCE (0.204) (0.204)	2)
Agri3 x Stra3 -0.529** -0.501	**
(0.217) (0.23)	1)
Obs 74 74 74 74 74 74 74 74	/
R-squared 0.736 0.782 0.795 0.795 0.896 0.89	9
Panel BPolitical integration above community 1000 BCE	
Agriculture primary 0.879^{***} 0.833^{***} 0.796^{***} 0.761^{***} 0.804^{***} 0.725^{**}	***
1000 BCE (Agri3) (0.058) (0.101) (0.141) (0.167) (0.176) (0.21)	2)
Agriculture secondary 0.379 0.338 0.354 0.349 0.457 0.48	5
-1000 BCE (0.367) (0.387) (0.394) (0.398) (0.342) (0.356)	5)
$Urbanisation > 400 \qquad 0.003 \qquad 0.02$	5
people 1000 BCE (Urb3) (0.144) (0.04)	6)
Urbanisation > 100-399 0.099 0.11	1
people 1000 BCE (0.160) (0.09)	2)
Pop density $26+$ 0.010 0.10	1
ppl/sqm 1000 BCE (Dens3) (0.174) (0.114)	9)
Pop density $1-25$ 0.107 0.04	
ppl/sqm 1000 BCE (0.171) (0.09)	6)
Social stratification $3+$ 1.046 ⁺⁺⁺⁺ 1.046 ⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺⁺	4)
classes 1000 BCE (Stra3) (0.041) (0.041) (0.041)	4)
Social stratification 2 0.230 0.23	1
$\begin{array}{c} \text{classes} (0.186) (0.186$	9) ***
Agris x Stras $-0.806^{+++} -0.839$	
(0.168) (0.12)	()
UDS 00	G
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	0
Agri3 x Urb3 / Agri3 x Dens3 0 0	,

Table 3: Political integration and agriculture by ACE ethno-archaeological zones

Notes: The table displays OLS regressions of political integration on agriculture presence, urbanization, population density and social stratification indicators based on the cultural complexity indices from the Atlas of Cultural Evolution - ACE (Peregrine, 2003), at 1000 BCE and 2000 BCE. These indices are compiled for homogeneous archaeological traditions. Using the ACE maps, our sample includes all polygons for archaeological tradition regions. For each polygon, we generated dummy variables for each factor of cultural complexity, based on the original ACE indices. Controls: polygon centroid latitude and distance to pristine state. All regressions include intercepts. Agriculture-urbanization and agriculture-population density interactions are omitted due to multicollinearity (no polygons have the highest levels of urbanization and population density before agriculture as a primary source). Robust standard errors in parentheses. * p < 0.05, ** p < 0.01, *** p < 0.001

3.3.3. Cross-country data with aggregated ACE cultural complexity indices

Relying exclusively on the ACE data permits us to only account for the presence of agriculture as a dominant food source, and not for the timing since the agriculture transition. The latter is important, as it encompasses additional potential channels of societal transformation. For instance, a longer period since the switch to agriculture may have allowed for the development of superior storage technologies, trade and taxation infrastructures, all of which require the creation of institutions to carry out and safeguard these activities. The time dimension of the passage from agriculture to states was the unifying element of our previous analysis.

We thus return to our original country-level data (based on modern-day borders) on the time since the Neolithic revolution, and state presence (based on the BOP 2018 state age), coded for 2000 BCE and 1000 BCE.²⁰ For each period, the subsample of countries used is restricted by which territories were already reliant on agriculture at the time. We then regress a binary state presence indicator (1 if centralized power manifested at period t on the territory of country i, according to the BOP 2018 classification) on the time since agriculture adoption. To explore the potential mechanisms discussed above, we include the period-specific country-average indicators of cultural complexity (urbanization, population density, social stratification).

The results are displayed in Table 4, with specifications similar to those in Table 3, but instead including diffusion region fixed effects and geographic controls. The most important result is that the proximate channel variables from the ACE dataset display larger coefficients in 1000 BCE than in 2000 BCE. Of the three ACE average indices, only urbanisation has significant coefficients in 1000 BCE. The time since agriculture adoption is significant and consistent in magnitude in 2000 BCE. One additional millennium of agriculture is associated with about 10-17 percent increase in the likelihood of developing state institutions within the territories defined by modern-country

²⁰Reassuringly, the correlation between the country-aggregate measure of political integration based on ACE and the BOP state presence indicator is 0.68 in 1000 BCE.

borders, in 2000 BCE. The coefficients of the interaction terms between time since agriculture adoption and urbanization, population density and social stratification are noisily estimated and insignificant in most specifications. However, in 2000 BCE, the negative estimate of the interaction between agriculture timing and social stratification is broadly consistent with the results in Table 3. This suggests that having a higher level of stratification became more important than the marginal effect of an additional millennium of agriculture for state formation by 2000 BCE. By 1000 BCE, urbanization and stratification appear to predict state formation, while the marginal effect of timing since agriculture adoption becomes insignificant. These differences relative to the results in table 3 may be attributable to several factors. First, it is possible that our state presence indicator captures states at a more advanced level of political integration than the ACE-based political development indicator. Second, the time since the emergence of agriculture captures a different feature than the presence or absence of agriculture as a primary source of sustenance. Third, our measures for the proximate channels are based on simple averages of the scales across ACE archaeological polygons within modern-country territories and the specifications reported include three different interaction terms using these averaged scales.²¹

It is worth emphasizing that these results are also limited by the unavailability of more detailed data on how other channels developed over the millennia BCE, such as warfare, food storability, taxation. However, to the extent that a longer time elapsed since the transition to agriculture captures developments in capacity to manage, distribute and tax food resources, these are accounted for in the model, and feed into the main significant effects in these regressions.

To conclude, our analysis explores the traditional channels put forth in the literature as correlates (if not drivers) of state emergence: population pressures, urbanization and the emergence of social classes. We find some evidence that the latter (and

²¹Weighting each polygon by its share of current territory cannot be readily automated and population weights, which would be preferable, are unavailable.

some limited support for urbanisation) was associated with increased likelihood of state emergence in 2000 and 1000 BCE. Moreover, where developed, social stratification appears to assume greater marginal importance in explaining state presence or formation than does reliance on agriculture, although the latter may have underpinned the former's emergence. The evidence seems to challenge the often cited temporal sequence agriculture-population pressure-state emergence in the process of state formation. This is consistent with recent evidence from archaeology, which shows a non-linear pattern in population growth following the transition to agriculture in Europe, where initial population booms were followed by (very likely) endogenous population declines, due to disease or soil depletion (Shennan et al, 2013).

4. Concluding Remarks

The association between a population's transition to reliance on domesticated crops and animals for its subsistence, and changes in its political structure culminating in the emergence of states, is strongly evident in our data. To be sure, only the pristine cases might be accepted as fully independent, with the strictest level of independence being limited to four to six cases only. Transition to primary reliance on agriculture is highly correlated with independent state emergence, despite the small sample. Emergence of states through internal political developments in countries that we classify as being in the spread zones of pristine states, must be viewed as providing less fully independent evidence.

Nevertheless, the similar way in which time transpires from adoption of agriculture to emergence of states in these cases offers further support for the idea of a process whereby, by engendering population growth, social stratification and density of settlement, opportunity for new forms of political organization to take root were likely fostered in similar ways across a large number of localities. Even those cases in which the first macro polity was directly attributable to an external group or empire display a similar pattern at least on average, perhaps because until recent centuries, the conquest and rule of territory was usually focused on areas more populous than those occupied by foragers alone. No countries in our sample display simultaneous arrival of both agriculture and the state from without before that phenomenon became common in the post-1450 colonial era. Our analysis based on the territories of most of the world's countries today thus supports, with expanded coverage and statistical precision, the long held belief that transition to agriculture was in the large majority of cases a prologue to the emergence of states throughout the world.

Table 4:	BOP	State	presence,	time	since	a griculture	and	cultural	complexity	across
countries										

	$\begin{array}{c} (1) \\ \textbf{Panel A:} \end{array}$	$\begin{array}{c} (2)\\ \mathbf{BOP} \ \mathbf{St} \end{array}$	(3) ate Pres	$\stackrel{(4)}{\mathbf{ence in } 2}$	000 BCE
	0 1 4 0 4 4 4	0.000**	0.000*	0.100*	
Agyears: Time since transition	0.146^{+++}	0.092^{**}	0.092^{*}	(0.106^{*})	0.172^{**}
Agreears x	(0.021)	(0.037) 0.076	(0.050)	(0.059)	0.186
ACE urbanization in 2000 BCE		(0.064)			(0.140)
ACE urbanization in 2000 BCE		-0.043			0.042
A		(0.217)	0.000		(0.398)
ACE population density in 2000 BCE			(0.082)		(0.095)
ACE population density in 2000 BCE			-0.105		(0.211) -0.365
			(0.306)		(0.756)
Agyears x			()	0.060	-0.278^{*}
ACE stratification in 2000 BCE				(0.071)	(0.150)
ACE stratification in 2000 BCE				-0.154	(0.194)
				(0.203)	(0.392)
Obs	82	82	82	82	82
R-squared	0.708	0.727	0.718	0.711	0.755
	Panel B: BOP State Presence in 1000 BCE				
	Panel B:	BOP St	ate Pres	ence in 1	000 BCE
Agreers: Time since transition	Panel B:	BOP St	ate Pres	ence in 1	000 BCE
Agyears: Time since transition to Agriculture in 1000 BCE (ky)	Panel B: 0.115*** (0.018)	BOP St 0.081^{**} (0.035)	ate Pres 0.061^{*} (0.033)		$\begin{array}{r} 000 \text{ BCE} \\ 0.009 \\ (0.060) \end{array}$
Agyears: Time since transition to Agriculture in 1000 BCE (ky) Agyearse	Panel B: 0.115*** (0.018)	$\begin{array}{c} \textbf{BOP St} \\ 0.081^{**} \\ (0.035) \\ 0.045 \end{array}$	ate Pres 0.061* (0.033)	$\frac{\text{ence in 1}}{\begin{array}{c} 0.012\\ (0.055) \end{array}}$	$\begin{array}{r} 000 \text{ BCE} \\ 0.009 \\ (0.060) \\ -0.125^* \end{array}$
Agyears: Time since transition to Agriculture in 1000 BCE (ky) Agyearse x ACE Urbanisation in 1000 BCE	Panel B: 0.115*** (0.018)	$\begin{array}{c} \textbf{BOP St} \\ 0.081^{**} \\ (0.035) \\ 0.045 \\ (0.039) \end{array}$	ate Pres 0.061* (0.033)	$\frac{\text{ence in 1}}{\begin{array}{c} 0.012\\ (0.055) \end{array}}$	$\begin{array}{c} \hline 000 \ \text{BCE} \\ \hline 0.009 \\ (0.060) \\ -0.125^* \\ (0.073) \end{array}$
Agyears: Time since transition to Agriculture in 1000 BCE (ky) Agyearse x ACE Urbanisation in 1000 BCE ACE Urbanization in 1000 BCE	Panel B: 0.115*** (0.018)	BOP St 0.081** (0.035) 0.045 (0.039) 0.239**	$\frac{\text{ate Pres}}{0.061^*}_{(0.033)}$	$\frac{\text{ence in 1}}{0.012} \\ (0.055)$	000 BCE 0.009 (0.060) -0.125* (0.073) 0.278* (0.278*
Agyears: Time since transition to Agriculture in 1000 BCE (ky) Agyearse x ACE Urbanisation in 1000 BCE ACE Urbanization in 1000 BCE	Panel B: 0.115*** (0.018)	$\begin{array}{c} \text{BOP St} \\ 0.081^{**} \\ (0.035) \\ 0.045 \\ (0.039) \\ 0.239^{**} \\ (0.112) \end{array}$	ate Pres 0.061* (0.033)	$\frac{\text{ence in 1}}{0.012} \\ (0.055)$	$\begin{array}{c} \hline 0.000 \ \text{BCE} \\ \hline 0.009 \\ (0.060) \\ -0.125^* \\ (0.073) \\ 0.278^* \\ (0.150) \\ 0.111 \end{array}$
Agyears: Time since transition to Agriculture in 1000 BCE (ky) Agyearse x ACE Urbanisation in 1000 BCE ACE Urbanization in 1000 BCE Agyears x ACE Population Density in 1000 BCE	Panel B: 0.115*** (0.018)	BOP St 0.081^{**} (0.035) 0.045 (0.039) 0.239^{**} (0.112)	$ \underbrace{ \text{ate Pres}}_{0.061*}_{0.033)} \\ \underbrace{ 0.049}_{0.047}_{0.047} $	$\frac{\text{ence in 1}}{0.012} \\ (0.055)$	$\begin{array}{c} \hline 0.00 \ \text{BCE} \\ \hline 0.009 \\ (0.060) \\ -0.125^* \\ (0.073) \\ 0.278^* \\ (0.150) \\ 0.111 \\ (0.089) \end{array}$
Agyears: Time since transition to Agriculture in 1000 BCE (ky) Agyearse x ACE Urbanisation in 1000 BCE ACE Urbanization in 1000 BCE Agyears x ACE Population Density in 1000 BCE ACE Population Density in 1000 BCE	Panel B: 0.115*** (0.018)	BOP St 0.081^{**} (0.035) 0.045 (0.039) 0.239^{**} (0.112)	$\underbrace{\begin{array}{c} \text{ate Pres} \\ 0.061^{*} \\ (0.033) \end{array}}_{0.049} \\ (0.047) \\ 0.330 \end{array}$	$\frac{\text{ence in 1}}{(0.012)}$	$\begin{array}{c} \hline 0.000 \ \text{BCE} \\ \hline 0.009 \\ (0.060) \\ -0.125^* \\ (0.073) \\ 0.278^* \\ (0.150) \\ 0.111 \\ (0.089) \\ 0.184 \end{array}$
Agyears: Time since transition to Agriculture in 1000 BCE (ky) Agyearse x ACE Urbanisation in 1000 BCE ACE Urbanization in 1000 BCE Agyears x ACE Population Density in 1000 BCE ACE Population Density in 1000 BCE	Panel B: 0.115*** (0.018)	BOP St 0.081^{**} (0.035) 0.045 (0.039) 0.239^{**} (0.112)	$\begin{array}{c} \underline{\text{ate Pres}}\\ 0.061^{*}\\ (0.033)\\ \end{array}$	<u>ence in 1</u> 0.012 (0.055)	$\begin{array}{c} \hline 0.000 \ \text{BCE} \\ \hline 0.009 \\ (0.060) \\ -0.125^* \\ (0.073) \\ 0.278^* \\ (0.150) \\ 0.111 \\ (0.089) \\ 0.184 \\ (0.399) \\ \end{array}$
Agyears: Time since transition to Agriculture in 1000 BCE (ky) Agyearse x ACE Urbanisation in 1000 BCE ACE Urbanization in 1000 BCE Agyears x ACE Population Density in 1000 BCE ACE Population Density in 1000 BCE	Panel B: 0.115*** (0.018)	BOP St 0.081^{**} (0.035) 0.045 (0.039) 0.239^{**} (0.112)	$\begin{array}{c} \underline{\text{ate Pres}}\\ 0.061^{*}\\ (0.033) \\ \end{array}$ $\begin{array}{c} 0.049\\ (0.047)\\ 0.330\\ (0.244) \end{array}$	<u>ence in 1</u> 0.012 (0.055) 0.135**	$\begin{array}{c} \hline 0.000 \ \text{BCE} \\ \hline 0.009 \\ (0.060) \\ -0.125^* \\ (0.073) \\ 0.278^* \\ (0.150) \\ 0.111 \\ (0.089) \\ 0.184 \\ (0.399) \\ 0.135 \\ 0.25 \end{array}$
Agyears: Time since transition to Agriculture in 1000 BCE (ky) Agyearse x ACE Urbanisation in 1000 BCE ACE Urbanization in 1000 BCE Agyears x ACE Population Density in 1000 BCE ACE Population Density in 1000 BCE ACE Stratification in 1000 BCE	Panel B: 0.115*** (0.018)	BOP St 0.081** (0.035) 0.045 (0.039) 0.239** (0.112)	$\begin{array}{c} \underline{\text{ate Pres}}\\ 0.061^{*}\\ (0.033) \\ \end{array}$ $\begin{array}{c} 0.049\\ (0.047)\\ 0.330\\ (0.244) \end{array}$	$ \begin{array}{r} ence in 1 \\ 0.012 \\ (0.055) \end{array} $ $ 0.135^{**} \\ (0.064) \\ 0.004 \end{array} $	$\begin{array}{c} \hline 0.000 \ \text{BCE} \\ \hline 0.009 \\ (0.060) \\ -0.125^* \\ (0.073) \\ 0.278^* \\ (0.150) \\ 0.111 \\ (0.089) \\ 0.184 \\ (0.399) \\ 0.135 \\ (0.087) \\ 0.220 \end{array}$
Agyears: Time since transition to Agriculture in 1000 BCE (ky) Agyearse x ACE Urbanisation in 1000 BCE ACE Urbanization in 1000 BCE Agyears x ACE Population Density in 1000 BCE ACE Population Density in 1000 BCE Agyears x ACE Stratification in 1000 BCE ACE Stratification in 1000 BCE	Panel B: 0.115*** (0.018)	BOP St 0.081** (0.035) 0.045 (0.039) 0.239** (0.112)	$\begin{array}{c} \underline{\text{ate Pres}}\\ 0.061^{*}\\ (0.033) \\ \end{array}$	$ \begin{array}{r} ence in 1 \\ 0.012 \\ (0.055) \end{array} $ $ 0.135^{**} \\ (0.064) \\ -0.094 \\ (0.188) \end{array} $	$\begin{array}{c} \hline 0.009 & 0.009 \\ (0.060) \\ -0.125^{*} \\ (0.073) \\ 0.278^{*} \\ (0.150) \\ 0.111 \\ (0.089) \\ 0.184 \\ (0.399) \\ 0.135 \\ (0.087) \\ -0.220 \\ (0.215) \end{array}$
Agyears: Time since transition to Agriculture in 1000 BCE (ky) Agyearse x ACE Urbanisation in 1000 BCE ACE Urbanization in 1000 BCE Agyears x ACE Population Density in 1000 BCE ACE Population Density in 1000 BCE ACE Stratification in 1000 BCE ACE Stratification in 1000 BCE	Panel B: 0.115*** (0.018)	BOP St 0.081^{**} (0.035) 0.045 (0.039) 0.239^{**} (0.112)	$\begin{array}{c} \underline{\text{ate Pres}}\\ 0.061^{*}\\ (0.033) \\ \end{array}$	$\frac{\text{ence in 1}}{0.012}$ (0.055) 0.135^{**} (0.064) -0.094 (0.188)	$\begin{array}{c} 0.009 \\ (0.060) \\ -0.125^{*} \\ (0.073) \\ 0.278^{*} \\ (0.150) \\ 0.111 \\ (0.089) \\ 0.184 \\ (0.399) \\ 0.135 \\ (0.087) \\ -0.220 \\ (0.215) \end{array}$
Agyears: Time since transition to Agriculture in 1000 BCE (ky) Agyearse x ACE Urbanisation in 1000 BCE ACE Urbanization in 1000 BCE Agyears x ACE Population Density in 1000 BCE ACE Population Density in 1000 BCE ACE Population Density in 1000 BCE ACE Stratification in 1000 BCE ACE Stratification in 1000 BCE	Panel B: 0.115*** (0.018)	BOP St 0.081^{**} (0.035) 0.045 (0.039) 0.239^{**} (0.112)	$\begin{array}{c} \text{ate Pres} \\ 0.061^{*} \\ (0.033) \\ \end{array}$ $\begin{array}{c} 0.049 \\ (0.047) \\ 0.330 \\ (0.244) \end{array}$ $\begin{array}{c} 114 \\ 0.524 \end{array}$	$\begin{array}{c} \underline{\text{ence in 1}}\\ 0.012\\ (0.055) \end{array}$ $\begin{array}{c} 0.135^{**}\\ (0.064)\\ -0.094\\ (0.188) \end{array}$ $\begin{array}{c} 114\\ 0 \end{array}$	$\begin{array}{c} \hline 0.009 & 0.009 \\ (0.060) \\ -0.125^{*} \\ (0.073) \\ 0.278^{*} \\ (0.150) \\ 0.111 \\ (0.089) \\ 0.184 \\ (0.399) \\ 0.135 \\ (0.087) \\ -0.220 \\ (0.215) \\ \hline 114 \\ 0.024 \end{array}$
Agyears: Time since transition to Agriculture in 1000 BCE (ky) Agyearse x ACE Urbanisation in 1000 BCE ACE Urbanization in 1000 BCE Agyears x ACE Population Density in 1000 BCE ACE Population Density in 1000 BCE ACE Population Density in 1000 BCE ACE Stratification in 1000 BCE ACE Stratification in 1000 BCE Obs R-squared Diffusion region FF	Panel B: 0.115*** (0.018) 114 0.751 Vos	BOP St 0.081^{**} (0.035) 0.045 (0.039) 0.239^{**} (0.112) 114 0.786	$ \begin{array}{r} ate Pres \\ 0.061^{*} \\ (0.033) \\ 0.049 \\ (0.047) \\ 0.330 \\ (0.244) \\ 114 \\ 0.784 \\ Vos $	$ \begin{array}{r} ence in 1 \\ 0.012 \\ (0.055) \\ \hline 0.135^{**} \\ (0.064) \\ -0.094 \\ (0.188) \\ 114 \\ 0.778 \\ \hline Voc \\ \hline Voc \\ \hline \end{array} $	$\begin{array}{c} \hline 000 \ \text{BCE} \\ \hline 0.009 \\ (0.060) \\ -0.125^* \\ (0.073) \\ 0.278^* \\ (0.150) \\ 0.111 \\ (0.089) \\ 0.184 \\ (0.399) \\ 0.135 \\ (0.087) \\ -0.220 \\ (0.215) \\ \hline 114 \\ 0.804 \\ \hline \text{Vec} \end{array}$

Notes: The table displays panel data OLS regressions of state presence (indicator =1 if a proto-state or a full state had emerged at time t in country i, based on BOP, 2018) on agriculture presence and the Atlas of Cultural Evolution - ACE urbanization, population density and social stratification indicators (Peregrine, 2003), years 2000 and 1000 BCE. Using the ACE maps, we coded the territories of countries-of-today by averaging the ACE values for those polygons corresponding to distinct archaeological traditions in ACE that overlap each country of today. The country-level ACE indicators are simple averages of the cultural complexity indicators. Controls: latitude, distance to coast and rivers, agriculture suitability, mean elevation, precipitations, temperature, landlocked indicator, percentage population at risk of malaria. All regressions include intercepts. Robust standard errors in parentheses. * p < 0.05, ** p < 0.01, *** p < 0.001

References

- Ahlerup, Pelle and Ola Olsson. 2012. "The roots of ethnic diversity." Journal of Economic Growth 17 (2): 71-102.
- Allen, Robert C. 1997. "Agriculture and the origins of the state in ancient Egypt." Explorations in Economic History 34 (2): 135-54.
- Ang, James B. 2015. "What drives the historical formation and persistent development of territorial states?" The Scandinavian Journal of Economics 117 (4): 1134-75.
- Ashraf, Quamrul and Oded Galor. 2011. "Dynamics and stagnation in the Malthusian epoch." American Economic Review 101 (5): 2003-41.
- Ashraf, Quamrul and Oded Galor. 2013. "The Out of Africa Hypothesis, Human Genetic Diversity, and Comparative Economic Development." American Economic Review, 103(1):1–46.
- Bates, Robert. 1983. Essays on the political economy of rural Africa. University of California Press.
- Besley, Timothy and Torsten Persson. 2010. "State capacity, conflict, and development." *Econometrica*, 78(1):1–34.
- Bleaney, Michael and Arcangelo Dimico. 2011 "Biogeographical conditions, the transition to agriculture and long-run growth." *European Economic Review* 55 (7): 943-54.
- Bockstette, Valerie, Areendam Chanda, and Louis Putterman. 2002. "States and markets: The advantage of an early start." Journal of Economic Growth 7 (4): 347-69.
- Boix, Carles. 2015. Political Order and Inequality: Their Foundations and their Consequences for Human Welfare. New York: Cambridge University Press.
- Borcan, Oana, Ola Olsson, and Louis Putterman. 2014. "State history and economic development: evidence from six millennia." Brown University Department of Economics Working Paper 2014–6.
- Borcan, Oana, Ola Olsson, and Louis Putterman. 2018. "State history and economic

development: evidence from six millennia." Journal of Economic Growth 23 (1): 1-40.

Borcan, Oana, Ola Olsson, and Louis Putterman. 2021. "Transition to agriculture and first state presence: A global analysis" Open-ICPSR, https://doi.org/10.3886/E140722V1.

Carneiro, Robert .1970. "A Theory of the Origin of the State," Science, 169:733-738.

- Comin, Diego, Easterly, William, Gong, Erick. 2010. "Was the Wealth of Nations Determined in 1000 BC?" American Economic Journal: Macroconomices 2: 65-97.
- Dal Bo, Ernesto, Pablo Hernandez and Sebastian Mazzuca, 2015, "The Paradox of Civilization: Pre-Institutional Sources of Security and Prosperity," NBER working paper 21829.
- Diamond, Jared. 1997. Guns, Germs, and Steel: The Fates of Human Societies. 1st ed. New York: W.W. Norton & Co.
- Dow, Gregory and Clyde Reed. 2013. "The origins of inequality: insider, outsider, elites, and commoners." *Journal of Political Economy* 121(3):609–641.
- Dow, Gregory and Clyde Reed. 2018. "The Origins of City States in Southern Mesopotamia," unpublished paper, Simon Fraser University.
- Ertan, Arhan, Martin Fiszbein, and Louis Putterman. 2016. "Who was colonized and when? A cross-country analysis of determinants." *European Economic Review* 83: 165-84.
- Fenske, James. 2014. "Ecology, Trade, and States in Pre-colonial Africa," Journal of the European Economic Association, 12(3): 612—640.
- Gennaioli, Nicola and Hans-Joachim Voth. 2015. "State Capacity and Military Conflict," *The Review of Economic Studies*, 82(4): 1409–1448.
- Harish, S.P., and Christopher Paik. 2019. "Historical State Stability and Economic Development in Europe," *Political Science Research and Methods*, 1-19.
- Hibbs, Douglas A. and Ola Olsson. 2004. "Geography, biogeography, and why some countries are rich and others are poor." *Proceedings of the National Academy of Sciences* 101 (10): 3715-20.
- Johnson, Allen W. and Timothy K. Earle. 2000. The Evolution of Human Societies:

from Foraging Group to Agrarian State. 2nd ed. Stanford: Stanford University Press.

- Larson, Gregor, Dolores Piperno, Robin Allabyd, Michael Purugganan, Leif Andersson, Manuel Arroyo-Kalin, Loukas Barton, Cynthia Climer Vigueira, Tim Denham, Keith Dobney, Andrew N. Doust, Paul Gepts, M. Thomas Gilbert, Kristen Gremillion, Leilani Lucas, Lewis Lukens, Fiona B. Marshall, Kenneth Olsen, J. Chris Piress, Peter Richersont, Rafael Rubio de Casas, Oris Sanjur, Mark Thomas, and Dorian Fuller. 2014. "Current Perspectives and the Future of Domestication Studies," Proceedings of the National Academy of Sciences 111 (17): 6139-6146.
- MacNeish, Richard. 1992. The Origins of Agriculture and Settled Life. Norman: University of Oklahoma Press.
- Mayoral, Laura and Ola Olsson (2019) "Pharaoh's Cage: Environmental Circumscription and Appropriability in Early State Development" SSRN Working Paper http://dx.doi.org/10.2139/ssrn.3496225.
- Mayshar, Joram, Omer Moav, Zvika Neeman and Luigi Pascali. 2019. "The Origin of the State: Land Productivity or Appropriability?" unpublished, University of Warwick.
- Morris, Ian. 2010. Why the West Rules for Now: The Patterns of History and What They Reveal About the Future, London: Profile Books.
- Murdock, George P. and Caterina Provost (1973) "Measurement of Cultural Complexity" In: A. Tuden ed. *Ethnology*, 12(4):379-392.
- Olsson, Ola and Douglas A. Hibbs. 2005. "Biogeography and long-run economic development." *European Economic Review* 49 (4): 909-38.
- Oppenheimer, Stephen. 2003. *Out of Eden: The Peopling of the World*. London: Constable and Robinson.
- Oppenheimer, Stephen. 2012a. "A single southern exit of modern humans from Africa: Before or after Toba?" *Quaternary International* 258: 88-99.
- Oppenheimer, Stephen. 2012b. "Out-of-Africa, the peopling of continents and islands: Tracing uniparental gene trees across the map." *Philosophical Transactions of The*

Royal Society B 367 (1590): 770-84.

- Oppenheimer, Stephen. 2014. "Modern humans spread from Aden to the antipodes with passengers and when?" In Southern Asia, Australia and the Search for Human Origins, edited by Robin Dennell and Martin Porr, 228-42. New York: Cambridge University Press.
- Peregrine, Peter N.. 2003. "Atlas of Cultural Evolution" In:J.P. Gray ed. World Cultures, 14(3): 1-75.
- Peregrine, Peter N. and Melvin Ember eds.. 2001. Encyclopedia of Prehistory. Vols1-8. New York: Kluwer Academic/Plenum Publishers.
- Peregrine, Peter. N., Carol R. Ember, and Melvin Ember. 2007. "Modeling State Origins Using CrossCultural Data," Cross-Cultural Research 41(1), 75-86.
- Petersen, Michael B. and Svend-Erik Skaaning. 2010. "Ultimate causes of state formation: The significance of biogeography, diffusion, and neolithic revolutions." *Historical Social Research* 35 (3): 200-26.
- Piperno, Dolores R. 2011. "The origins of plant cultivation and domestication in the new world tropics: Patterns, process, and new developments." *Current Anthropol*ogy 52 (S4): S453-70.
- Piperno, Dolores R. and Deborah M. Pearsall. 1998. The Origins of Agriculture in the Lowland Neotropics. Cambridge, MA: Academic Press.
- Putterman, Louis. 2007. "State antiquity index: Introduction and appendix." Brown University Department of Economics. http://www.brown.edu/Departments/Economics/ Faculty/Louis_Putterman/State_Antiquity_Index_V3%201%20update%20Nov.%202017. pdf.
- Putterman, Louis. 2008. "Agriculture, Diffusion and Development: Ripple Effects of the Neolithic Revolution" *Economica* 75: 729-748.
- Putterman, Louis and Cary A. Trainor. 2006 [Revised 2018]. "Agricultural transition year - country data set." Brown University Department of Economics. http://www. brown.edu/Departments/Economics/Faculty/Louis_Putterman/agricultural%20data% 20page.htm.

- Schönzolzer, David (2020) "The Origin of the Incentive Compatible State: Environmental Circumscription," unpublished paper, Institute for International Economic Studies.
- Scott, James. 2017. Against the Grain: A Deep History of the Earliest States Yale University Press.
- Service, Elman. 1971. Primitive Social Organization: An Evolutionary Perspective.2nd ed. New York: Random House.
- Service, Elman and Marshall Sahlins. 1960. Evolution and Culture. Ann Arbor: University of Michigan Press.
- Smith, Bruce D. 1995. The Emergence of Agriculture. New York: Scientific American Library.
- Soares, Pedro, Luca Ermini, Noel Thomson, Maru Mormina, Teresa Rito, Arne Röhl, Antonio Salas, Stephen Oppenheimer, Vincent Macaulay, and Martin B. Richards. 2009. "Correction for purifying selection: An improved human mitochondrial molecular clock." American Journal of Human Genetics 84 (6): 740-59.
- Spencer, Charles S. 2010. "Territorial expansion and primary state formation." Proceedings of the National Academy of Sciences 107 (16): 7119-26.
- Spolaore, Enrico and Romain Wacziarg. 2013. "How Deep Are the Roots of Economic Development?" Journal of Economic Literature 51 (2): 325-369.
- Tilly, Charles. 1990. Coercion, Capital, and European States, AD 990-1990. Cambridge, MA: Basil Blackwell.
- Weber, Max. 1919. "Politics as vocation," published by Duncker Humblodt, Munich.
- Wittfogel, Karl A.. 1957. Oriental Despotism: A Comparative Study of Total Power, Yale University Press, New Haven CT.