Deluge and the Rise of Civilizations:
From the Neolithic to Early States

Shuo Chen, Xinyu Fan, and Yanfei Yin*

Abstract: The legends of almost all ancient civilizations involved catastrophic floods and collective human efforts to recover from the disasters. Using the great deluge of the Yellow River around 1920 BCE as a natural experiment, we show that the deluge accelerated early civilization development from the Neolithic to early states, including class differentiation, public buildings, the use of bronze, writing, and early cities. In a horse race of possible channels, both the cooperation and productivity hypotheses contributed to civilizational development, with no single hypothesis dominating the other. Therefore, the paper highlights the multivariate evolution of civilizational development and state formation.

Keywords: deluge, early civilization, multivariate development, cooperation, productivity

JEL: N45, O10, H70, Q10
1. Introduction

All four ancient human civilizations originated from river basins, nourished by fertile soil and easy access to irrigation (Wittfogel, 1957), but were constantly suffering from the threats of floods (Chaney, 2013). Accounts of world-ending floods are found from Genesis to the Epic of Gilgamesh, while legends are told in Native Americans on the rise of new civilizations after great floods (Frazer, 1918). Throughout human history, one thing is clear: floods are instrumental shapers of early civilizations. On the other hand, existing studies on early civilizations focus more on the formation of states and government (Sanchez de la Sierra, 2020; Mayshar et al., 2022; Allen et al., 2023), while little is known about how human civilizations evolved from hunter-gatherer societies with stone-made tools in the Neolithic to well-functioning cities and states – the transition of which paved the way for long-term economic and social prosperity (Giuliano and Nunn, 2021; Galor et al., 2023). To address this gap, we examine the dynamics of civilization development along the Yellow River from 5000 BCE to 500 BCE: the period that witnessed the transformation of early Chinese civilization from a Neolithic way of living to established states with frequent inter-state conflict, using the great deluge in 1920 BCE as a source of variation that profoundly reshaped civilization development.

The Yellow River basin, covering an area of 795,000 square kilometers (twice the size of modern Germany), is considered as the birthplace of Chinese civilization. In about 1920 BCE, the region experienced a gigantic flood caused by the collapse of an earthquake-induced weir in the Yellow River that had blocked its path for over six months. This event caused 12-17 billion cubic meters of water flooding the middle and lower reaches of the river, forming a massive deluge. The great deluge resulted in a substantial increase in soil productivity due to the deposition of rich sediments. However, it also posed survival challenges which required collective efforts to secure resources through peaceful cooperation or violent conflicts. We thus expect the deluge accelerated the civilization development in the middle and lower
reaches of the Yellow River, from the Neolithic to early states and other complex social organizations.

To do this, we collect information on 1,983 archaeological sites, construct 10×10 kilometer grid-level panel data for 5000-500 BCE, and systematically examine the impact of this deluge on civilizational development. To reflect the development phase from the Neolithic to early states, we follow the archaeological literature and track a set of five early civilization characteristics: the presence of class differentiation, the presence of public buildings, the use of bronze, evidence for writing, and traces of early cities (Child, 1950; Trigger, 1978; Xia, 1985). Class differentiation and public buildings are often prerequisites for complex social hierarchies, while the presence of bronze, writing, and early cities are generally considered to be the three essential elements in defining the birth of civilization. Specifically, we examine both the extensive margin, which determines whether one of these characteristics is present, and the intensive margin, which determines whether the grid has achieved multiple characteristics. Our treatment consists of the flood-affected grids whose distance from the Yellow River is less than 50 kilometers. Our control group consists of grids further away from the river and therefore less affected by the deluge.

Using a difference-in-differences (DID) strategy, our main results show that, after the deluge, areas closer to the Yellow River, i.e., areas more affected by the flood, were more likely to form early civilization characteristics in both the extensive and intensive margins. A parallel trend test confirms the validity of our DID strategy that the treatment and control groups feature no systematic difference in civilization development prior to the deluge. The results are robust after addressing the potential selection bias in the presence of archaeological data, the incomplete coverage of civilization characteristics, the measurement error from mapping the sites to grids, and the clustering standard errors in grid sizes. The results are also robust to considering the confounding low-temperature climate shocks around 2000 BCE and the political system changes around 1000 BCE.
Regarding the mechanism, we highlight the multivariate origin of early civilizations that no single causal factor explains the evolution of civilizations (Peregrine et al., 2004; Peregrine et al., 2007), and test three possible channels: the cooperation hypothesis, the productivity hypothesis, and the conflict hypothesis. First, we follow Waldinger (2022) to construct the potential cooperation opportunities for each grid and show that grids with higher cooperation opportunities are more prone to early civilization characteristics, thus validating the cooperation hypothesis (Wittfogel, 1957; Vorlaufer and Steimanis, 2023). Second, we obtain slope gradient data and show that gentler slopes – which allows for more sediment deposition and thus increased soil productivity – predict higher levels of early civilization development, which confirms the productivity hypothesis (Engels, 1884; Childe, 1936; Johnson and Earle, 2000). Third, the conflict hypothesis suggests that the outbreak of floods may accelerate civilization development through wars over limited living space and resources. However, we did not observe a decrease in the number of early civilizations in each grid following conflict-related conquests, thus finding limited support for the conflict hypothesis (Carneiro, 1970; Turchin et al., 2013). Moreover, we conduct a horse race between two validated channels and show that both productivity and cooperation hypotheses are at work, with no single hypothesis dominating the other.

This paper makes three contributions to the literature. First, this paper advances our understanding of the origin of civilizations by filling the gap in civilization development determinants from the Neolithic era to the early states. The existing literature on the origin of the state primarily builds on established infrastructures such as the existence of cities and public good provisions, and explores multiple perspectives of early state formation, such as the need for irrigation (Wittfogel, 1957; Allen et al., 2023); the need for war (Tilly, 1985); the need for resource extraction (Sanchez de la Sierra, 2020); and the natural consequence of

---

1 It is worth noting that the nullification of the conflict hypothesis complements to – rather than contradicts with – Tilly (1985)’s well-known proposition that “War made the state, and state made the war”, in that the interaction between conflicts and state building implies the existence of the state in the first place. Some recent studies (e.g., Abramson, 2017) also challenge the validity of the statement by showing the war did not explain the formation of sovereign states in Europe. Nevertheless, the discussion of established states and their dynamics is not the focus of this paper. Instead, our focus is precisely the development of civilizations when no prior institutions of such ever existed.
productivity improvement (Engels, 1884).\(^2\) Meanwhile, research on the Neolithic revolution mainly focuses on their long-term impact on the modern world (Galor and Moav, 2007; Putterman, 2008; Olsson and Paik, 2020).\(^3\) Taken together, there is a gap in the literature regarding the evolution of civilizations from the Neolithic era to early states with established cities and primitive hierarchies. This paper fills the gap by discussing how external shocks such as the deluge in the Yellow River contributed to the development of early Chinese civilization, from primitive characteristics such as class differentiation and public buildings to more advanced features such as the use of bronze, writing, and early cities.

Second, we contribute to the multivariate origin of civilizations. Echoing Peregrine et al. (2007)’s theory of cultural evolution that “cultural evolution is multivariate (i.e., that several causal factors are involved)”, and “simple univariate models and analyses do not capture the multivariate processes at work”, we explore the logic empirically to early human civilizations, and show that both the productivity hypothesis and the cooperation hypothesis played important and non-substitutable roles in early civilization development. In this vein, the closest paper to ours is Allen et al. (2023), which shows that the demand for irrigation after river shifts in Southern Iraq in 2850 BCE contributed to the formation of government. We improve upon the contributions of Allen et al. (2023) on two fronts: first, we provide a richer and multi-layered measure of early civilization characteristics from class differentiation to the existence of early cities, thus alleviating the concern in the existing literature that artificial irrigation can be introduced independently from the state (Leach, 1959; Woodbury, 1961; Kang, 2006); second, we highlight the multivariate nature in the origin of civilizations that both the cooperation and the productivity hypothesis contribute to civilization development.

---

\(^2\) We differ from Sanchez de la Sierra (2020) in focusing on the development of civilizations in prehistory rather than modern contexts, with no prior forms of states or complex hierarchies can be learned or observed. In a similar vein of prehistoric state formation, Mayshar et al. (2022) utilizes cross-sectional data to demonstrate a causal effect of the cultivation of cereals on the formation hierarchy. We differ from Mayshar et al. (2022) in using panel data and a DID approach with the deluge as a natural experiment, thus our findings are more robust to omitted variable concerns.

\(^3\) Another strand in the literature on early human civilization aims to construct indicators of prehistoric development, such as household consumption (Gupta and Halket, 2023), inequality (Boix and Rosenbluth, 2014; Kohler, et al., 2017), economic development (Bakker et al., 2018), state capacity (Fenske, 2014), and city sizes (Barjamovic et al., 2019). Within this strand, Murdock and White (1969) select 186 societies globally to construct the Standard Cross-Cultural Sample (SCCS) that includes civilization characteristics such as political power, taxation and hierarchical complexity. The construction of indicators is not the focus of this paper. However, as shown in Section 3, the indicators in Borcan et al. (2018) lend support to the validation of our data.
Third, we contribute to the literature on endogenous institution formation. Existing studies mainly conduct theoretical discussions on the endogenous formation of political institutions or regard them as exogenous setups (Banerjee and Iyer, 2005; Dell et al., 2018). This paper, on the other hand, empirically discusses how environmental shock affects the endogenous formation of complex social organizations such as early cities, joining in the existing discussions of the impact on religious authority and institutional changes (e.g., Brückner and Ciccone, 2011; Chaney, 2013; Benati et al., 2022). In this regard, a paper close to ours is Chaney (2013), who studies the impact of floods on the Nile on the interactions of political powers in established states. We differ from Chaney (2013) in discussing the role of great floods on the development of civilization, especially how natural disasters accelerated the endogenous formation of early civilization characteristics from zero to one.

The remainder of the paper is organized as follows. Section 2 presents the historical background. Section 3 describes the data and the construction of variables. Section 4 presents the empirical strategy of the paper and the results. Section 5 examines the possible mechanisms. And Section 6 concludes.

2. Background and Historical Context

This section provides the historical background to put our research question in context. First, we introduce the key concepts and indicators in understanding early civilization development. Next, we describe the details of the great deluge in 1920 BCE and its impact.

2.1 Early Civilization Development

Timeline of Civilization Development. The development of early human civilization can be broadly divided into four stages: the Paleolithic, Neolithic, Bronze, and Iron Ages (Lubbock,
In the Paleolithic Age, human utilized tools made of stone, wood, and bone and lived in a hunter-gatherer society without complex social division of labor or social hierarchies. Around 7000 BCE, early human civilizations entered the Neolithic Age, which marked the transition from hunter-gatherer societies to sedentary agriculture. During this period, social classes began to differentiate, with evidence of wealth inequality found in tombs and the emergence of a social hierarchy. At the same time, early cities appeared, accompanied by public buildings such as warehouses, temples, and irrigation facilities.

Around 3300 BCE, the extensive use of bronze marked the beginning of the Bronze Age and the emergence of large cities, including Lagash in Mesopotamia and Erlitou in the Yellow River Basin. The Bronze Age also witnessed the emergence of writing. While the order and timing differed, archaeologists widely agree that the origin of civilizations can be attributed to the interplay of bronze, writing, and cities (Daniel, 1968; Xia, 1985). The state, as the most complex form of political organization, was also born during this period (Service, 1975). After 1200 BCE, human civilization entered the Iron Age, when cities continued to expand, and trade and conflicts between countries became very frequent. Figure 1 displays the timeline of civilization development for four major ancient civilizations.

[Insert Figure 1 here]

**Five Indicators Measuring Early Civilization Development.** Given our focus on civilization development, we follow Sanchez de la Sierra (2020) to track the major indicators reflecting the phases of civilizations. Based on the timeline of civilization development and

---

4 According to Fagan and Durrani (2020), a city should meet the following criteria: 1. a concentration of population, usually above 5,000 individuals in size; 2. the presence of division of labor; and 3. a complex social organization, with the presence of public buildings such as shrines and walls. Archaeologists can usually determine whether a site is a prehistoric city based on the size of the site or the scale of its relics.

5 This paper presents a novel approach to studying the three elements by extending the period of analysis from their origins to their continuous expansion. This differs from the existing literature where empirical studies often focus on the period after the formation of these elements. As an illustration, Allen et al. (2023) analyzed the impact of a river shift in Mesopotamia on the origins of government during the timeframe of 3900-2700 BCE. However, it should be noted that the formation of Uruk City dated back to around 4000 BCE. The current paper, however, focuses on the time frame between 5000-500 BCE, examining the development of the three elements from zero to one. For example, the earliest city in the Yellow River Basin, the Xishan site, was formed around 3300-2800 BCE (Qian, 1999), and the earliest writing appeared after 1200 BCE. Figure 1 also visualizes and compare the timeframe of Allen, et al. (2023) and this paper.
conventions in archaeology, we select five major indicators for this study, namely, class differentiation, public buildings, use of bronze, written records, and cities (Child, 1950; Trigger, 1978; Xia, 1985). Among them, class differentiation – typically evidenced by the disparity in burial objects – and the emergence of large-scale buildings indicate social stratification, which forms the basis of civilization development (Engels, 1884; Liu and Chen, 2012; Bulliet et al., 2010). Meanwhile, the three key elements of civilization origins, i.e., bronze, writing, and cities, indicate a higher level of civilization development (Zou, 1987; Chang, 2004). In what follows, we introduce the critical periods in Chinese civilizations through the lens of these key indicators.

Timeline of Chinese Civilization. The earliest Chinese civilization originated in the Yellow River Basin, which covers an area of 795,000 square kilometers, twice as large as modern Germany. Known as the mother river, the Yellow River witnessed Chinese civilization entering the Neolithic Age around 7000 BCE. Until 5000 BCE, several Neolithic civilizations surfaced in the Yellow River Basin, with evidence of crop cultivation, livestock breeding, and wide usage of ground stone tools and pottery. Chinese civilization continued to develop between 5000 BCE and 2000 BCE, with the appearance of large-scale public buildings such as altars and irrigation facilities. Class differentiation also emerged with richer burial objects and much evidence of human sacrifice (Ren, 2000; Yan, 2009).

Around 2000 BCE, the Erlitou civilization was formed in the middle and lower reaches of the Yellow River, with advanced pottery, jade, and bronze crafting skills. A high degree of social differentiation was observed, and traces of early cities emerged. Around 1600 BCE, the Shang civilization established the first formal state in Chinese history. The earliest writing was also discovered in this period. Around 1000 BCE, the Zhou civilization replaced the Shang and established several small vassal states based on blood relations, where the number

---

6 Features such as the development of trade (Bulliet et al., 2010) and population growth (Murdock, 1957) are also considered important indicators of early civilizations. However, quantifying these indicators in prehistoric contexts is challenging due to data availability.

7 Social stratification refers to a social hierarchy created by the classification of individuals and groups based on power, wealth, or occupation.
of cities expanded rapidly (Tan, 1996). In the later Zhou civilization, especially after 770 BCE, China entered an extended period of war and turmoil, with many vassal states involved in constant conflicts. Eventually, the Qin dynasty, the first unified state, was established in 221 BCE.

In summary, from 5000 BCE to 500 BCE, Chinese civilizations prospered in the Yellow River Basin, from differentiating social classes, building large public structures, to the extensive usage of bronze, writing, cities, and eventually early states. During this evolution, a great deluge struck the Yellow River Basin around 1920 BCE, which fundamentally shaped the natural conditions of the middle and the low reaches of the Yellow River Basin. In the following, we introduce the details of the deluge.

2.2 The Deluge in 1920 BCE

The Scale of Deluge. Around 1920 BCE, a major deluge broke out in the Jishi Gorge area of the upper Yellow River (Wu et al., 2016). The deluge was triggered by an earthquake that created a 240-meter-high dam blocking the Yellow River for over six months. 12-17 billion cubic meters of water were impounded, forming a vast weir whose capacity was about 54%-77% of that of the Three Gorges Dam – the largest water conservancy project in contemporary China. Six to nine months after the earthquake, the dam collapsed, and 11-16 billion cubic meters of water poured down the Yellow River, creating a great deluge. The deluge flow peaked at 360,000 to 480,000 cubic meters per second, equivalent to more than 500 times the present-day flow of the Yellow River in similar regions and more than four times the peak of the Great Mississippi Flood in 1927, one of the greatest floods in the American history.

The Impact of Deluge. The deluge wreaked havoc on Chinese civilization throughout the Yellow River Basin in both the short and long term. In the short term, the unprecedented scale of the deluge spread downstream to the middle and lower reaches, directly destroying
civilizations along the Yellow River, such as the Lajia site.\(^8\) In the long term, the deluge had a lasting impact on the civilizations along the Yellow River Basin, which destroyed the natural levees in the lower reaches, causing constant floods more likely. The constant floods consequently influenced civilization development, in which three avenues are most prominently discussed in the literature. First, the sediment deposition from floods increased soil productivity, thus contributing to more production surplus, a prerequisite for the advancement of social structures (Engels, 1884). Second, the floods encouraged cooperation across regions, such as building dams, dredging rivers, sharing information, and sharing risks, which requires fundamental organizational changes – such as the state – to achieve (Wittfogel, 1957; Vorlauffer and Ivo Steimanis, 2023). Third, floods may trigger wars over limited living space and resources, demanding complicated organizational structures (Carneiro, 1970; Turchin et al., 2013). Moreover, these three avenues may co-exist and interact, as illustrated by the multivariate theory of civilization origins by Peregrine et al. (2007).

**Testing Civilization Development.** According to the *Records of the Grand Historian* (*Shiji*), the first general chronicle of Chinese history,\(^9\) the deluge directly led to the establishment of the Xia civilization, where the founding ruler of the dynasty spent more than a decade managing the deluge and subsequently gained his legitimacy (e.g., Wu and Ge, 2005a, 2005b; Wu and Ge, 2014). Moreover, the Shang civilization was forced to relocate its capital multiple times due to rampant flooding. Historical anecdotes aside, did the great deluge accelerate the development of early civilizations along the Yellow River? If so, through what channels did this acceleration occur? In the following sections, this paper systematically examines the impact of the deluge on the early civilization development in China, using archaeological site data from 5000 BCE to 500 BCE.

---

\(^8\) The Lajia site is located in Qinghai Province, China, about 25 kilometers from the Jishi Gorge area where the deluge originated (Figure C1). The large number of artifacts such as stone tools, pottery, jade, and bone tools unearthed at the Lajia site, and various types of houses, trenches, and burials indicate that the Lajia site already had a very high level of civilization when the deluge struck (Ren, et al., 2002; Ye, 2002). However, the deluge completed destroyed the Lajia site (Xia, et al., 2003), and severely struck several civilizations, such as the Qijia, Longshan, and Taosi cultures (Figure C1).

\(^9\) Some scholars believe that the Shang dynasty – rather than the Xia dynasty – was the first formal state in Chinese history (Gu and Tong, 1982; Loewe and Shaughnessy, 1999), because the large number of excavated archaeological findings directly corroborate the existence of the Shang dynasty, while no direct evidence has been found for the existence of the Xia dynasty. Nevertheless, based on indirect evidence such as archaeological remains, it is widely recognized that the Xia dynasty was the first state of China (Chang, 1986; Sun, 2018).
3. Data

We construct an original panel dataset based on archaeological sites from 5000 BCE to 500 BCE, covering the upstream and midstream Yellow River Basin in China. Section 3.1 describes the construction of our dataset. Section 3.2 introduces the dependent and independent variables for our empirical analysis. Suggestive evidence is presented in Section 3.3.

3.1 Sample Area and Archaeological Sites

Sample Area in Grids. Our dataset spans from 5000 BCE to 500 BCE, covering the upstream and midstream Yellow River basin. The cross-sectional unit of observation is a 10×10 kilometers grid cell, which reduces the impact of shifting boundaries due to dynastic changes. The observation period is chosen to reflect the development from the birth of Chinese civilization to the formation of complex feudal states. The sample area, we exclude the downstream part of the Yellow River basin because the river frequently shifts due to sediment accumulation (Shelach-Lavi, 2015), leading to the instability of the downstream basin. Figure 2 depicts the area of the Yellow River Basin relevant to our study. Moreover, we exclude the upstream area of the Jishi Gorge since the great deluge broke out from the Jishi Gorge. We also exclude the grid whose distance to the Yellow River is greater than 100 kilometers to make the grids more comparable. Figure C2 shows the corresponding sample area. For clarity, the remaining graphs of the Yellow River basin follow Figure 2.

---

10 Some may be concerned that after 2000 BCE, the Chinese civilization underwent rapid development, thus the five early civilization indicators cannot capture the levels of development accurately. However, as shown in Figures 6 and 7 in the following sections, there existed substantial variation among the levels of development among the Yellow River basin that several indicators did not emerge in many regions. This means that our indicators remained considerably relevant.

11 According to the definition of the Yellow River Conservancy Commission, the demarcation point of the middle and lower reaches of the Yellow River is Taohuayu, Zhengzhou City, Henan Province, which is located between 113 and 114°E. Therefore, we have excluded areas in the Yellow River basin beyond 114°E.
Archaeological Sites. We manually collect information on archaeological sites and match them to the grids in both spatial and temporal dimensions. Specifically, we collect archaeological sites from *The Dictionary of Chinese Archaeology*, which documents major and up-to-date archaeological sites excavated in China. In total, we collect information on 1,983 archaeological sites. The detailed coding procedure is discussed in Appendix A.

Matching Sites with Grids. We match the archaeological sites to the grids in spatial and temporal dimensions. The detailed procedure is discussed in Appendix B. Figure 3 depicts the location of the archaeological sites in our sample area. The larger the gray circle, the more archaeological sites there are in the grid. In the temporal dimension, we match the archaeological sites to the grids in terms of the existence time of the sites. Specifically, we divide our observation period into nine sectors, each spanning 500 years, such as 5000 BCE - 4500 BCE and 4500 BCE - 4000 BCE, and map the sites to the sectors if their existence time intersected with the sectors – a common practice used in the literature to soften the measurement error in archaeological estimates (e.g., Anderson et al., 2017).

---

12 An alternative dataset on the development of prehistoric civilization in China is the *Chinese Archaeological Database* compiled by the HKU-Tsinghua Quantitative History Research Team, based on published archaeological reports and briefs. The database digitizes information on prehistoric sites in China, such as excavated artifacts, and covers the period from the Paleolithic era to 220 CE. The database is currently under construction, thus no public access is available. According to the description of the database, both the database and the *Dictionary of Chinese Archaeology* are based on a large number of archaeological excavation reports. The *Dictionary of Chinese Archaeology* was compiled by the Institute of Archaeology of the Chinese Academy of Social Sciences, making it a solid source in the field. To verify the validity of our data, we also compare it with other databases that describe the development of prehistoric civilization in China, as elaborated in the variable construction section.

13 More than 10,000 Neolithic sites have been discovered in China, but very few have been excavated (Liu, 2005). For example, in Shanxi Province of China, a total of 2,179 Neolithic sites have been discovered by 2020, but only about 80 of them have been excavated ([https://weibo.com/tarticle/p/show?id=230940456582228012370](https://weibo.com/tarticle/p/show?id=230940456582228012370)). The reason is twofold: on the one hand, the excavation process may cause damage to the sites. Therefore, those sites that are difficult to protect after excavation are not excavated; on the other hand, the excavation of sites often takes up a large area of land, which affects local economic development. Since the number of sites actually excavated is only a small fraction, this may lead to selection bias in our data, which will be discussed in our robustness checks.

14 We set the length of the period at 500 years because the existence of some sites may be vague. For example, the existence of the Chengtoushan site has been estimated to be around 4000-2800 BCE. In order to minimize measurement errors, we set the time interval to 500 years. We also do not use a dynamic, stage-dependent time interval because the definition of stages may be endogenous.
3.2 Treatment: The Impact of Deluge

The main explanatory variable of this paper is the intersection term of Deluge and Post. Deluge is a dummy variable indicating whether the grid is affected by the great deluge. Ideally, we need to consider various factors, such as the peak, flow rate, and flood velocity, to determine the affected grids. However, due to data limitations, we determine the treatment status according to the grids’ distance to the Yellow River. Following Hornbeck and Naidu (2014) and Allen et al. (2023), we define the affected regions as those within 50 kilometers of the River. That is, deluge equals to 1 if its distance to the Yellow River is less than 50 kilometers and 0 otherwise. We discuss the choice of distance cutoffs in Section 4. Figure C3 shows the distribution of treatment and control groups in the dataset. Post is a dummy variable indicating whether the section period is larger than 7 (2000 BCE – 1500 BCE), in which the great deluge outbroke. In sum, our dataset contains 7,033 grids from 5000 BCE to 500 BCE. The summary statistics are presented in Table 1.

[Insert Table 1 here]

3.3 Outcome Variables

Early Civilization Characteristics: Extensive Margin. As stated in Section 2.1, we track the five indicators that reflect the development of civilizations, namely, class differentiation, large public buildings, use of bronze, written records, and cities, from The Dictionary of Chinese Archaeology. Specifically, we use the presence of human sacrifice or aristocratic burials in the excavated tombs of the sites to measure whether there was class differentiation in the civilization. We use the presence of temples, warehouses, meeting halls, plazas, water facilities, public roads, and other buildings in the excavated sites to measure whether there were large-scale public buildings in the civilization. We also track the presence of bronzes,
writings, and early city relics in the excavated sites. Table C1 visualizes the above five characteristics in archaeological excavations. Next, we constructed a dummy variable – the existence of civilization characteristics, taking the value of 1 if at least one of the five characteristics, i.e., class differentiation, large public buildings, bronze, writing, or early cities, is present in the grid, and 0 otherwise. The dummy variable thus measures the extensive margin of early civilization development.

**Early Civilization Characteristics: Intensive Margin.** The extensive margin of civilization characteristics alone does not necessarily lead to civilization development. For example, some civilizations have differentiated social classes and built large public buildings, but the “efforts to build state failed” (Tilly, 1985). A more refined measure of civilization characteristics is needed to better reflect the early process of civilization development. Therefore, we construct a second variable, the intensity of civilization characteristics, which refers to the intensive margin, a la Sanchez de la Sierra (2020). Specifically, the variable takes the value of 1 if there was class differentiation or large public buildings in the grid, 2 if there was use of bronze, 3 if writing was found, and 4 if there were traces of early cities.

The reason for assigning a higher value to the presence of bronze, writing, and early cities is that they represent a higher stage in the development of civilizations than just the appearance of class differentiation and large-scale public buildings (Daniel, 1968; Xia, 1985; Zou, 1987; Chang, 2004). Around the globe, many early tribes developed class differentiation and constructed large-scale public buildings, but only a few evolved to manage the use of bronze and to develop writing and cities. Meanwhile, Morgan (1877) and Peregrine et al. (2007) argue that the emergence of writing represents a higher level of civilization than the use of bronze – thus a higher value placed on writing. Finally, traces of cities have often been associated with the presence of large walls for protection, the presence of a non-productive

---

15 Different scholars may have different definitions of early cities. Here, we follow the definition in the Dictionary of Chinese Archaeology: if the Dictionary specifies that the traces of a city in a site, then we define it as an early city accordingly.

16 Tilly (1985) said: “The enormous majority of the political units which were around to bid for autonomy and strength in 1500 disappeared in the next few centuries, smashed or absorbed by other states-in-the-making. The substantial majority of the units which got so far as to acquire a recognizable existence as states during those centuries still disappeared.”
class, and professional bureaucrats for governance – thus the prototype of the early state (Childe, 1950; Mumford, 1961; Cai, 1988). Fagan (2016) points out that “it may be possible to have states without cities, but it is hard to envisage a city that is not embedded within a state.” We therefore assign the highest weight to early cities. Together, the intensity of civilization characteristics reflects the progress toward a more developed civilization. Figure 4 shows the spatial distribution of the intensity of civilization characteristics: a gray triangle indicates a site with no civilization characteristics, a red circle indicates a site with civilization characteristics, and a larger circle indicates greater intensity.

[Insert Figure 4 here]

**Validity of Measure.** To further solidify the validity of our measure, we compare it with the dataset used in Borcan et al. (2018), who construct a state history index of 159 countries worldwide since 3500 BCE based on the existence of states in prehistoric times when discussing the long-term impact of state existence. We do not directly follow Borcan et al. (2018) because we focus on the evolution of civilizational development, while Borcan et al. (2018) focus on the emergence of prehistoric states. However, the two indicators are comparable because prehistoric states implies a new level of civilization development. We extract the information for China from the dataset and compare it with ours. Figure 5 presents the comparison, where the black line is the state history index for China constructed by Borcan et al. (2018), and the red and blue lines indicate the variables constructed in this paper. The horizontal axis of this figure is the period – each period lasts 500 years; the vertical axis is the mean value of the variable within the period. The upper left corner is the $t$-value.

---

17 It should be noted that even cities can vary in their level of civilizational development. For example, some cities were mere population centers, while others had developed complex political and economic systems. In this paper, we focus on the existence of cities rather than the developmental heterogeneity within them. Moreover, the construction of variables here does not imply a monotonous progression from the use of bronze to writing to early cities. For example, the Maya civilization in Mesoamerica produced writing and cities, but no bronze, while the Inca civilization in South America produced bronze and cities, but no use of writing.

18 Borcan et al. (2018) follow the methodology of the original effort by Bockstette et al. (2002). This combines three dimensions of state development: (1) the existence of a state above the tribal level; (2) whether rule was internally or externally based (i.e., whether a country's territory had an autonomous government or was partly or wholly ruled by an authority outside its borders); (3) how much of its territory was under the control of a government (as opposed to multiple competing governments and regions that still lacked state presence).
obtained from the regression of the state history index for China on the two variables we constructed. Our measures are highly correlated with those of Borcan et al. (2018). Specifically, Borcan et al. (2018) indicate that the earliest states were formed in China around 2000 BCE, during which time our early civilization characteristics also increased sharply. This adds to our confidence that our data adequately reflects the early development of civilizations and states.

[Insert Figure 5 here]

3.4 Suggestive Evidence

Before proceeding to the formal analysis, we provide descriptive evidence to help contextualize our findings. As shown in Figure 5, both the existence of civilization characteristics and the intensity of civilization characteristics witnessed a sizeable increase after 2000 BCE. The pattern is shared by the data from Borcan et al. (2018). Figure 6 further shows the time trends of these five characteristics, which confirms the increase after the deluge. To visualize the association more clearly, we show the distribution of sites before and after the deluge in Figure 7. As shown, areas affected by the deluge had, on average, both more sites and better civilization development.

[Insert Figures 6 and 7 here]

4. Empirical Strategies and Results

In this section, we empirically estimate the impact of the great deluge on early civilization development. Section 4.1 characterizes our DID strategy and validates the identification assumptions. In Section 4.2, we present our DID results and extend our analysis to allow for
greater variation in treatment intensity. In Section 4.3, we conduct a series of robustness checks of the data generation process and the regression process.

### 4.1 Empirical Strategy and Identifying Assumptions

**Empirical Strategy.** Our empirical strategy follows a standard DID approach, in which we compare the relative change in our main outcome variables in grids close to the Yellow River (distance less than 50 kilometers) relative to distant grids. The econometric specification takes the following form:

\[
Y_{ct} = \beta F_{c} \times P_{t} + \delta_{c} + \sigma_{t} + \delta_{c} \cdot f(t) + \epsilon_{ct}
\]

where \(c\) denotes the grid and \(t\) denotes the period. \(Y_{ct}\) is an outcome of interest for grid cell \(c\) in period \(t\). \(F_{c}\) is a dummy variable that equals one if a grid is close to Yellow River and zero otherwise. Hence, the treated group comprises deluge grids, while the control group comprises other (non-deluge) grids. \(P_{t}\) is a dummy variable that equals one for the period after the great deluge. The equation also includes controls for grid and period fixed effects, \(\delta_{c}\) and \(\sigma_{t}\), ruling out concerns that our results are driven by time-invariant factors or temporal effects that affect all grids simultaneously. We also include \(\delta_{c} \cdot f(t)\), which is the grid-specific period trend, to control for heterogeneous trends across grids. The regression is clustered at the grid level. The coefficient of interest in equation (1) is \(\beta\). The coefficient is expected to be all positive, indicating a higher level of civilization development in the treatment group after the great deluge.

**Identifying Assumptions.** The validity of our DID approach hinges on the parallel trend assumption that the treatment group and the control group featured no differences in early civilization development prior to the deluge. To test this assumption, we estimate a fully flexible equation that takes the following form:

\[
Y_{ct} = \sum_{j=-6}^{3} \beta_{j} F_{c} \times P_{t}^{j} + \delta_{c} + \sigma_{t} + \epsilon_{ct}
\]
where all variables are defined as in equation (1). The only difference from equation (1) is that in equation (2), rather than interacting $F_{\text{lood}}_{c}$ with a post-deluge indicator variable, we interact the treatment variable with each of the period fixed effects, treating the period 2500 BCE – 2000 BCE as the reference group. The estimated vectors of $\beta_{j}$ reveal the differences between the treated and control prefectures during each period. Figure 8 plots the estimates of equation (2). As shown, the difference between the treated and control groups is constant over time and small in magnitude before the deluge. After the deluge, we observe a sharp increase in the development of the civilization characteristics in the treatment group. This pattern is consistent with the parallel assumption, which confirms the validity of our DID approach.\(^{19}\)

[Insert Figure 8 here]

4.2 Baseline Results

**Extensive Margin.** We present our baseline results in Table 2. As shown, the deluge significantly increases the extensive margin of civilization development, i.e., the presence of civilization characteristics. Specifically, Column 1 controls for grid and period fixed effects to rule out any time-invariant grid features and period shocks that uniformly affect all grids. Column 2 adds the intersection between the deluge and the period to account for the different trends between the treatment and control groups. Column 3 further includes the grid-specific period trend to control for the heterogeneous trend across grids. Regarding magnitudes, the probability of having early civilization characteristics increased by 1.6% in the grids near the Yellow River after the deluge, which is about 33% higher than the sample mean (0.016/0.012).

\(^{19}\) The DID approach also assumes that shocks are relatively exogenous. Since the outbreak of the deluge is due to the earthquake and not to human activities, we consider the shock to be a qualified natural experiment. However, it remains possible that the potential early civilization development, such as building cities, will influence their exposure to the floods, i.e., the output in $t-1$ correlates with treatment status in $t$. We address the issue as follows. First, we assign the grids within 50 kilometers from the Yellow River as the treatment group, independent of the development level of the grids. Second, we estimate Equation (1) using lagged outcomes, and report the results in Table C1. As shown, the estimates are all insignificant after controlling the grid-specific period trend. This suggests that the lagged outcomes have a limited impact on our results.

Electronic copy available at: https://ssrn.com/abstract=4650064
Intensive Margin. After analyzing the extensive margin of civilization characteristics, we examine the intensive margin – the intensity of civilization characteristics. We re-estimate equation (1) using this new independent variable and present the results in Table 3. As shown, the impact of the deluge remains positive and significant, with a larger magnitude: the intensity of civilization characteristics increased by 0.049, about 44% (0.049/0.034) compared with the sample mean.

Treatment Intensity. One limitation of our empirical study is the lack of well-defined treatment and control groups. In the baseline, we define the treatment group using a cutoff distance of 50 kilometers from the Yellow River. However, the impact of the deluge may spill over to nearby grids due to, for example, migration. Furthermore, the assignment of 50 kilometers as the cutoff can be ad hoc. In the following, we allow for a more flexible measure of treatment intensity using the distance to the Yellow River. In particular, we consider the following equation:

\[ Y_{ct} = \sum_{j=1}^{4} \beta_k Post_t \times Distance_c^j + \delta_c + \sigma_t + \epsilon_{ct} \]  

(3)

where all variables are defined as in equation (1). The only difference from equation (1) is that in equation (3), instead of interacting Post_t with a treatment indicator variable, we interact it with a series of distance dummy variables. For example, Distance_c^1 equals to 1 if the distance to the Yellow River is between 0 and 20 kilometers; Distance_c^2 equals 1 if the distance to the Yellow River is between 20 and 40 kilometers. We use the distance set [80, 100] as the reference group. The estimated vectors of \( \beta_k \) reveal the differences between the grids within different distance sets and those far from the Yellow River.

20 We also re-estimate the parallel trends, and report the results in Figure C4.
Figure 9 presents the regression results based on equation (3). As shown, the impact of the deluge decreases with increasing distance: when the distance to the Yellow River exceeds 60 kilometers, the impact is close to 0 and no longer significant. The result rationalizes the cutoff of 50 kilometers as the division between the treatment and control groups in the baseline.

[Insert Figure 9 here]

4.3 Robustness Checks

The baseline results indicate greater early civilization development in regions closer to the Yellow River after the deluge. However, this result may be affected by settings in the data generation and regression processes. In this section, we conduct four robustness checks from the perspectives of both the data-generating and regression processes. For the data-generating process, there are two possible types of bias: sample selection bias and measurement error. For the regression process, our estimation results can be affected by confounding factors and model selection. In the following, we address each of these four challenges.

Are archaeological records accurate? There are three sources of sample selection bias in our archaeological data, concerning the existence, discovery, and recording of archaeological sites. First, some may be concerned that less developed early civilizations may not survive or be excavated as easily, thus, biasing our sample mean upward. However, our results would only be overestimated only if such less developed sites were more likely to be distributed away from the Yellow River area and less likely to be preserved after the deluge. No evidence has been found to support this variation. Second, some may be concerned that our results might be overestimated if archaeologists were more inclined to excavate prehistoric sites with higher levels of development, especially those with early civilization characteristics. However, this does not occur in practice because archaeologists cannot determine whether the site has early civilization characteristics prior to excavation. Third, some may be concerned that the Dictionary of Chinese Archaeology may be biased toward including
prehistoric sites with a high level of development. We are less concerned about this bias because the Dictionary includes a large number of less-developed sites.\textsuperscript{21} Also, our results would only be overestimated if the Dictionary preferred to include sites closer to the Yellow River and existed after the deluge. However, neither the distance nor the time is a criterion for inclusion in the Dictionary. Lastly, the coverage of civilization characteristics may be incomplete. For example, the literature also suggests that high population concentration (Murdock, 1957) and trade development (Bulliet et al., 2010) are characteristics of early civilizations. However, data on these characteristics are challenging to collect. In comparison, the five characteristics used in our data are also representative and widely used in literature. Moreover, as long as the measurement error is not related to the outbreak of the deluge, it does not affect the consistency of our estimates.

**Is the mapping from sites to grids accurate?** Due to the lack of precise scope of the sites, we cannot directly determine the coverage of grids. Therefore, it is possible that there are measurement errors in mapping sites to grids. In the baseline estimates, we assume that the grid in which the site is excavated and all its neighboring grids are the scope of the actual site. To reduce the measurement error due to possible mis-mapping, we restrict the sample to include the grid in which the site is excavated – the most refined definition of the sites and re-estimate our baseline. Table 4 presents the results, where our baseline results remain robust.

[Insert Table 4 here]

**Con founding political system and climate changes.** Two important confounding factors may affect our results: the political systems of the Zhou civilization established around 1000 BCE and the low-temperature period around 2000 BCE. Regarding the former, the proliferation of political systems after the establishment of Zhou would have affected the development process in different regions of the Yellow River Basin (Qian, 2017; Lv, 2020).

\textsuperscript{21} Of the five characteristics of early civilization, two-thirds of the 1,983 sites in our sample did not produce any of the characteristics.
To address this issue, we exclude the samples after 1000 BCE and re-estimate the results. As shown in Figure 10 (4), the baseline findings remain robust. For the latter, there is scientific evidence that China may have experienced a large-scale low temperature shock around 2000 BCE (Zhu, 1972; Denton and Karlén, 1973). The lower temperatures may have caused migration closer to the river, leading to an overestimation of our results. To address this bias, we conduct a placebo test based on the Yangtze River basin – another cradle of Chinese civilization (see Figure C5 for the location of the Yangtze River). If the common climate shock rather than the deluge facilitated the development of civilizations along the river, then we expect to observe this pattern in the Yangtze River basin as well. We follow the same treatment to estimate the results based on the Yangtze River basin and present the results in Figure 10 (5). As shown, there is no similar pattern in the Yangtze River Basin, indicating that our results are not driven by the common climate shock.

**Clustering standard error and spillover effects.** Last but not least, regarding model selection, the clustering standard error and spillover effects may affect the baseline results. For clustering standard errors, our baseline observation unit is a fine-grained 10 × 10 kilometers grid, leading to possible correlation of the error terms over a large spatial scale. To address this issue, we re-estimate Equation (1) at the 1° × 1° latitude and longitude grid. The results are shown in Figure 10 (1), where our baseline results remain robust. For spillover effects, the presence of migration may lead to a violation of the Stable Unit Treatment Value Assumption (SUTVA) in the DID approach. To address this issue, we use the spillover-robust estimation proposed by Clarke (2017), and exclude the samples 30-70 kilometers away.

---

22 The reason we chose the Yangtze River basin as a placebo is that the Yangtze River is also a cradle of Chinese civilization, and many prehistoric human civilizations originated in the Yangtze River basin. Moreover, the Yangtze River and the Yellow River were similar in terms of geography and climate in the prehistoric period. However, compared to the Yellow River, the Yangtze River was not as prone to flooding (Shelach-Lavi, 2015). Specifically, there is no evidence to prove that the Yangtze River basin experienced floods of the same magnitude as the Yellow River during the same period. Therefore, we can take the development of civilization in the Yangtze River basin during the same period as a plausible counterfactual.

23 Clarke (2017) proposes a data-driven (non ad-hoc) procedure which frames the issue as a problem similar to determining the optimal bandwidth for regression discontinuity design and generates a spillover-robust estimation. This method will determine, endogenously from data, the maximum distance of spillover and the distance bins over which spillovers propagate. We can understand these two concepts through an example. Assuming the spillover effect of deluge may exist even at 70 kilometers away from the Yellow River. At the same time, the spillover effect is larger in distance bin [50, 60] than in [60, 70], and the effect is homogeneous within each distance bin. In this example, the maximum distance of spillover is 20 kilometers, and the optimal distance bin is 10 kilometers. Then, we can estimate the treatment effect unbiased. Using this method, the researchers only have to set two key parameters: Delta and Maxdist. The parameter Delta defines the fineness of grid to be searched when testing for optimal distance bin. The parameter Maxdist defines the maximum possible
from the Yellow River – the areas more affected by migration.\textsuperscript{24} Figures 10 (2) and (3) present the results, where the baseline findings remain robust.

5. Multivariate Origin of Civilizations: Possible Channels

In the previous section, we found that the outbreak of the deluge in 1920 BCE accelerated the development of early civilizations in the Yellow River basin. In this section, we explore the possible causal channels. Classical theories of civilization development focus mainly on three hypotheses: productivity, conflict, and cooperation. According to the productivity hypothesis, the development of productivity led to an increase in agricultural surpluses, which accelerated the emergence of an unproductive elite class and other complex social organizations (Engels, 1884; Childe, 1936; Johnson and Earle, 2000). The conflict hypothesis suggests that population growth and limited resources lead to frequent conflicts in which violence is gradually monopolized and social structures become more complex (Carneiro, 1970; Turchin et al., 2013). The cooperation hypothesis suggests that maintaining stable irrigation requires highly centralized organization, strict discipline, and strong leadership (Wittfogel, 1957).

However, when multiple hypotheses are tested together, the effect of any single factor is unknown. The consensus among archaeologists and historians is that the factors influencing the progress of human civilization are complex and never monolithic (Renfrew, 1972; Lull and Mico, 2011). Service (1975) criticized the view of war, irrigation, and other factors as
distance to test for a particular distance bin. Under the two parameters, this method tests the optimal distance bin ranging from $\Delta_1$, increasing in units of $\Delta_1$, until reaching Maxdist, under the minimum Root Mean Squared Error (RMSE) rule. As the author said, Maxdist does not imply that spillovers cannot be estimated beyond the value set in Maxdist via the iterative estimation procedure, but rather that optimal distance bins will not be considered beyond Maxdist. In this paper, we set $\Delta_1=1$ and Maxdist=10.

\textsuperscript{24} In our study, only the migration across treatment and control groups may affect our results. Given our cutoff of 50 kilometers and the range of migration in prehistoric conditions, we assume the relevant migration took place in regions that are 30-70 kilometers away from the Yellow River. Following Hornbeck and Keniston (2017), we drop these observations in which migration are most plausible to alleviate the spillover effect.
the primary drivers of the origins of civilization as “oversimplified causal theories”. Carneiro (1970) argued that “while warfare may be a necessary condition …..it is not a sufficient one”. In sum, we must view the development of civilization as a process of multiple factors acting simultaneously. In what follows, we empirically test each of these three hypotheses. We also follow the multivariate theory of cultural evolution (Peregrine et al., 2007) and conduct a horse-race test among these possible channels in Section 5.4.

5.1 Cooperation Hypothesis

A popular hypothesis in the face of negative exogenous shocks, inspired by Wittfogel’s (1957) theory of hydraulic empires, is the cooperation hypothesis, which suggests that when faced with increased risk of natural hazards, people cooperate to combat them through collective action (Burton-Chellew et al., 2013; Gross and Böhm, 2020; Vorlaufer and Ivo Steimanis, 2023). In the context of this paper, the deluge of the Yellow River might necessitate cross-regional cooperation, accelerating the development of early civilizations (Allen et al., 2023). If the cooperation hypothesis holds, we expect to observe more collaborative activities post-deluge. Although direct observations of prehistoric activities are not available, we follow Waldinger (2022) to construct the potential cooperation opportunities for each grid. In this context, if the cooperation hypothesis holds, we expect grids with larger cooperation opportunities to be more prone to early civilization characteristics.

Specifically, we measure potential cooperation opportunities in two steps: First, we identify the number of archaeological sites in neighboring grids within 30 kilometers of each grid in the pre-deluge period (2500-2000 BCE) – a shorter distance indicates more potential cooperation opportunities, and divide our sample into two groups. A high-cooperation-opportunity group includes grids with at least one site in the neighboring grids within a 30 kilometers distance, otherwise the grids are included in the low-cooperation-opportunity group. Second, we calculate the average distance from the grid to those neighboring archaeological sites. Figure C6 visualizes our findings. As shown, the sites with civilization
characteristics are primarily located in the high-cooperation-opportunity group. We then re-estimate the baseline results using the new grouping. The results are presented in Table 5.25 The comparison between Column 1 and Column 2 shows that early civilization characteristics are more likely to emerge in the grids with larger cooperation opportunities. Furthermore, adding the intersection of the average distance, Column 3 shows that a smaller average distance to other sites – and thus more opportunities for cooperation – increases the likelihood of the base grid having early civilization characteristics. The results are consistent when evaluating civilization development using intensive margin, as shown in Columns 4-6. Together, our findings support the cooperation hypothesis.

[Insert Table 5 here]

5.2 Conflict Hypothesis

Besides cooperation, conflict may also shape early civilization development (Carneiro, 1970; Tilly, 1985; Turchin et al., 2013). In the context of the deluge, the conflict hypothesis points to fiercer competition and surged conflicts between the early civilizations for limited sources after the deluge. Although we do not have direct access to data on conflict records, if the conflict hypothesis holds, we expect the number of early civilizations in each grid to decrease due to the conquest post-conflict. Similarly, we expect the number of early civilizations for the 30 kilometers radius of a grid to decrease for the same reason. Next, we re-estimate equation (1) using the previous two indicators as our new dependent variables. Specifically, we differentiate early civilizations with archaeological sites in our baseline: separate sites in the same or neighboring grid could belong to the same civilization and thus were likely to

---

25 It is important to note that in this regression, we use the stricter coding method as in the robustness test. We make the adjustment to overcome the overestimation of our results. For example, if a grid has early civilization characteristics in the pre-deluge period, the neighboring grids are also assumed to have early civilization characteristics according to the baseline coding method. In other words, such coding methods may over-assign high-cooperation-opportunity groups with high levels of civilization development, which may lead to an overestimation of our results.
fight as allies in conflicts against other civilizations. Therefore, we count the number of civilizations rather than sites when testing the conflict hypothesis.\textsuperscript{26}

We report the results in Table 6. Column 1 shows that the number of civilizations within the grid did not decrease after the deluge. Column 2, we further control for the interaction terms of the pre-deluge number of civilizations and the time fixed effect to account for the possibility that war-prone grids may have different responses to common events. As shown, the number of early civilizations does not change significantly. Similarly, the results remain robust when we replace the number of within-grid early civilizations with those of nearby grids (within 30 kilometers of the base grid) in Columns 3-4. Taken together, we find no empirical support for the conflict hypothesis based on the distributional dynamics of civilizations in our study.

[Insert Table 6 here]

5.3 Productivity Hypothesis

Another popular hypothesis that scholars explain the development of early civilizations concerns productivity (Engels, 1884; Childe, 1936; Johnson and Earle, 2000). Just as the flooding of the Nile River brought fertile soil, the deluge of the Yellow River could have increased the productivity of the soil and, in turn, facilitated civilization development.

To test the productivity hypothesis, we obtained slope gradient data from Global Agro-Ecological Zones (GAEZ v4.0) for the Yellow River Basin. Since water flow is slower in areas with slower slopes, sediment deposition is more likely to occur and thus more conducive to increasing soil productivity. If this hypothesis holds, we expect to observe that areas with gentler slopes are more likely to form early civilization characteristics. Figure C7

\textsuperscript{26} For example, if there are five archaeological sites within 30 kilometers of a certain grid, three of which belong to the Longshan civilization and two to the Dawenkou civilization, we the assign a value of 2 instead of 5 to the calculation of the number of early civilizations.
visualizes the distribution of slopes in the Yellow River basin, with darker color representing greater slope. Combined with Figure 7(b), we find suggestive evidence that most of these more developed sites were distributed in the southeast and south of the Yellow River basin, where the slope is gentler. Further, we empirically test the hypothesis by adding the slope as interaction terms. Table 7 presents the results. As expected, gentler slopes predict higher levels of early civilization development – confirming the productivity hypothesis.

[Insert Table 7 here]

5.4 Horse Race of Channels

The previous analysis provides supportive evidence for the cooperation and productivity hypotheses. This section conducts a horse-race test of these two channels. Specifically, it is possible that gentler slopes – with more fertile soil – inherently attract more settlement of early civilizations and thus more cooperation opportunities. That is, the productivity hypothesis dominates the cooperation hypothesis. To examine this issue, we re-estimate the results in Table 5 using two subsamples: we use the 25th percentile of the slopes of all grids in the sample as the threshold – a grid is a high-slope area if its average slope exceeds this quartile, and is a low-slope area otherwise. If the productivity hypothesis dominates the cooperation hypothesis, we expect no acceleration in early civilization development for high-slope areas, even those with higher cooperation opportunities.

We present the results in Figure 11. As shown in the left panel, in grids with low cooperation opportunities, the deluge did not accelerate civilization development in high-slope subsamples. Although the coefficients are weakly significant in the low-slope subsamples, the magnitudes are very small. That is, our results are not solely driven by the productivity hypothesis. Moreover, in grids with high cooperation opportunities, the deluge accelerated early civilization development in both high and low slope subsamples, with a stronger effect in low slope subsamples - reflecting the impact of the productivity hypothesis. The results
are robust to both intensive and extensive margin evaluations of early civilization development. Taken together, the results indicate that both productivity and cooperation hypotheses are at work, whereas no single hypothesis dominates the other. As Peregrine et al. (2007) pointed out, “we believe cultural evolution is multivariate (i.e., that several causal factors are involved)”, and “simple univariate models and analyses do not capture the multivariate processes at work”.

[Insert Figure 11 here]

6. Conclusion

This paper uses data from archaeological sites dating from 5000-500 BCE and explores the causal relationship between great floods and the development of civilizations by using the prehistoric Yellow River deluge around 1920 BCE as a natural experiment. We find that after the outbreak of the deluge, areas close to the Yellow River were more likely to produce early civilization characteristics. This suggests that the deluge accelerated civilization development. In the mechanism analysis, we find supportive evidence of the cooperation hypothesis – that the increased cooperation after the deluge accelerated early civilization development; and the productivity hypothesis – that the deluge brought fertile soil for agriculture, whereas no single hypothesis dominates the other. However, we do not find supportive evidence of the conflict hypothesis that increased conflicts following the deluge accelerated early civilization development. As the first empirical investigation that reveals the dynamic evolution of early civilization in prehistory, this paper highlights the multivariate origins of early civilizations and enriches the literature on endogenous institution formation by exploring how external shocks affect the establishment of some of the earliest institutions in human history.
Reference


**Figures and Tables**

**Figure 1. Timeline of Civilization Development**

*Notes:* The figure shows the development of the four major ancient civilizations. We also highlight the emergence of writing and cities for the Mesopotamian and Chinese civilizations. The Mesopotamian civilization developed early Uruk cities around 4000 BCE (Fagan, 2016), and cuneiform writing was created around 3300 BCE (McIntosh, 2008; Bulliet et al., 2010). The earliest cities appeared in the Yellow River Basin around 3300-2800 BCE (Qian, 1999), and the earliest writing appeared in China until 1200 BCE (Boltz, 1986).
Figure 2. The Yellow River Basin

Notes: This figure shows the construction of grids. The geographical data were obtained from the Center for Resource and Environmental Science and Data of the Chinese Academy of Sciences (https://www.resdc.cn/).
Figure 3. Location of the Archaeological Sites in the Sample Area

Notes: This figure shows the distribution of archaeological sites. Among them, the blue lines indicate the Yellow River; the sample area of this paper is the gridded. The black circle indicates the number of sites in the grid, and a larger circle indicates more sites. The geographical data are from the Center for Resource and Environmental Science and Data, Chinese Academy of Sciences (https://www.resdc.cn/). Site data are from the Dictionary of Chinese Archaeology.
Figure 4. Distribution of the Dependent Variables in the Sample Area

Notes: This figure shows the distribution of the intensive margin for early civilization development. The blue lines indicate the Yellow River. The sample area of this paper is the gridded. The black triangle indicates that the site did not produce any early civilization characteristics. The red circle indicates the site produced at least one early civilization characteristic, and bigger circles indicate more characteristics thus higher levels of civilization development. The geographical data are from the Center for Resource and Environmental Science and Data, Chinese Academy of Sciences (https://www.resdc.cn/). Site data are from the Dictionary of Chinese Archaeology.
Figure 5. Comparison with Borcan et al. (2018)

Notes: This figure compares the data from Borcan et al. (2018) with this paper. The black line is the state history index for China constructed by Borcan et al. (2018), the red line indicates the existence of civilization characteristics constructed in this paper, and the blue lines indicate the intensity of civilization characteristics constructed in this paper. The horizontal axis of this figure is the period – each period lasts 500 years; the vertical axis is the mean value of the variable within the period. The upper left corner is the $t$-value obtained from the regression of the state history index for China on the two variables we constructed.
Figure 6. Trend of Early Civilization Characteristics

Notes: This figure illustrates the trend of early civilization characteristics. The red line indicates class differentiation. The blue line indicates public buildings. The yellow line indicates the use of bronze. The green line indicates the emergence of writing. The purple line indicates early cities. The horizontal axis of this figure is the period – each period lasts 500 years; the vertical axis is the mean value of the variable within the period.

Electronic copy available at: https://ssrn.com/abstract=4650064
Figure 7. Distribution of Sites before and after the Deluge

Notes: This figure shows the distribution of sites and their civilization development before and after the deluge. The blue lines indicate the Yellow River. The sample area of this paper is gridded. The black triangle indicates that the site did not produce any early civilization characteristics. The red circle indicates the site produced at least one early civilization characteristic, and larger circles indicate more characteristics and thus higher levels of civilization development. The geographical data are from the Center for Resource and Environmental Science and Data, Chinese Academy of Sciences (https://www.resdc.cn/). Site data are from the Dictionary of Chinese Archaeology.
Figure 8. The Effect of Great Deluge on Civilization Development: Event Study

Notes: The figure depicts the differences in the development of the civilization characteristics between the grids close to the Yellow River and those otherwise before and after the deluge. The markers and connected lines represent the ordinary least squares estimates and 95 percent confidence intervals based on standard errors clustered at the grid level. The dependent variable is the existence of the civilization characteristics. The reference period is 2500-2000 BCE. The regression considers grid fixed effects and period fixed effects.
Figure 9. The Effect of Great Deluge on Civilization Development: Treatment

Intensity

Notes: The figure depicts the differences in the development of the civilization characteristics between the grids within different distance bins and those very far from the Yellow River before and after the deluge. The markers and connected lines represent the ordinary least squares estimates and 95 percent confidence intervals based on standard errors clustered at the grid level. The dependent variables of panels (a) and (b) are the existence of the civilization characteristics and the intensity of the civilization characteristics, respectively. The distance bins are divided every 20 kilometers. The reference group is the grids whose distance to the Yellow River is more than 80 kilometers. The regression considers grid fixed effects and period fixed effects.
Figure 10. The Effect of Great Deluge on Civilization Development: Robustness Test

Notes: The figure depicts the estimations under different robustness tests. The markers and dashed lines represent the ordinary least squares estimates and 95 percent confidence intervals based on standard errors clustered at the grid level. The dependent variables are the existence of the civilization characteristics and the intensity of the civilization characteristics, respectively. The regression considers grid fixed effect, period fixed effect, and grid-specific period trends.
Figure 11. The Mechanism of the Effect of Great Deluge on Civilization Development:

Discussion

Notes: The figure depicts the estimates under different conditions. The markers and dashed lines represent the ordinary least squares estimates and 95 percent confidence intervals based on standard errors clustered at the grid level. The dependent variable is the existence of the civilization characteristics and the intensity of the civilization characteristics, respectively. “Low Cooperation Opportunities” refers to the grids with no other sites within 30 kilometers of neighboring grids in the pre-deluge period (2500-2000 BCE). “High Cooperation Opportunities” refers to grids with at least one site within 30 kilometers of neighboring grids in the pre-deluge period (2500-2000 BCE). “High Slope” refers to grids with slopes on the 25th percentile. “Low Slope” refers to the rest of the grids. The regression considers grid fixed effect, period fixed effect, and grid-specific period trends.

Electronic copy available at: https://ssrn.com/abstract=4650064
<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Source</th>
<th>Obs</th>
<th>Mean</th>
<th>Sd</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deluge</td>
<td>1</td>
<td>63,297</td>
<td>0.645</td>
<td>0.478</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Post</td>
<td>2</td>
<td>63,297</td>
<td>0.333</td>
<td>0.471</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Class differentiation</td>
<td>3</td>
<td>63,297</td>
<td>0.005</td>
<td>0.069</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Public Good</td>
<td>3</td>
<td>63,297</td>
<td>0.007</td>
<td>0.082</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Bronze</td>
<td>3</td>
<td>63,297</td>
<td>0.009</td>
<td>0.094</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Word</td>
<td>3</td>
<td>63,297</td>
<td>0.003</td>
<td>0.055</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>City</td>
<td>3</td>
<td>63,297</td>
<td>0.005</td>
<td>0.070</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>The existence of civilization characteristics</td>
<td>3</td>
<td>63,297</td>
<td>0.012</td>
<td>0.111</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>The intensity of civilization characteristics</td>
<td>3</td>
<td>63,297</td>
<td>0.034</td>
<td>0.331</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

Data Sources:
Table 2. The Effect of Great Deluge on Civilization Development: Extensive Margin

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Deluge×Post</strong></td>
<td>0.029***</td>
<td>0.016***</td>
<td>0.016***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Grid FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Period FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Deluge * Period Trend</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Grid-specific Trend</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Obs.</td>
<td>63,297</td>
<td>63,297</td>
<td>63,297</td>
</tr>
</tbody>
</table>

Notes: The sample consists of 7,033 grids from 5000 BCE to 500 BCE. The dependent variables are the existence of civilization characteristics. Deluge is an indicator that equals one if the grid is close to the Yellow River, with a distance of no more than 50 kilometers. Post is an indicator that equals one after 2000 BCE. Standard errors in parentheses are clustered at the grid level. ***, **, * represent significance levels of 1%, 5%, and 10%, respectively.
<table>
<thead>
<tr>
<th></th>
<th>Intensive Margin</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td>Deluge×Post</td>
<td>0.092***</td>
<td>0.049***</td>
<td>0.049***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td></td>
</tr>
<tr>
<td>Grid FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Period FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Deluge * Period Trend</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td></td>
</tr>
<tr>
<td>Grid-specific Trend</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Obs</td>
<td>63,297</td>
<td>63,297</td>
<td>63,297</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** The sample consists of 7,033 grids from 5000 BCE to 500 BCE. The dependent variables are the intensity of civilization characteristics, respectively. Deluge is an indicator that equals one if the grid is close to the Yellow River, with a distance of no more than 50 kilometers. Post is an indicator that equals one after 2000 BCE. Standard errors in parentheses are clustered at the grid level. ***, **, * represent significance levels of 1%, 5%, and 10%, respectively.
Table 4. The Effect of Great Deluge on Civilization Development: Alternative Coding

<table>
<thead>
<tr>
<th></th>
<th>Extensive Margin</th>
<th>Intensive Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td><strong>Deluge</strong>×<strong>Post</strong></td>
<td>0.008***</td>
<td>0.005***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Grid FE</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Period FE</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Grid-specific Trend</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Obs.</td>
<td>63,297</td>
<td>63,297</td>
</tr>
</tbody>
</table>

Notes: The sample consists of 7,033 grids from 5000 BCE to 500 BCE. The dependent variables are the existence of civilization characteristics and the intensity of civilization characteristics, respectively. **Deluge** is an indicator that equals one if the grid is close to the Yellow River, with a distance of no more than 50 kilometers. **Post** is an indicator that equals one after 2000 BCE. Standard errors in parentheses are clustered at the grid level. ***, **, * represent significance levels of 1%, 5%, and 10%, respectively.
Table 5. The Mechanism of the Effect of Great Deluge on Civilization Development:

<table>
<thead>
<tr>
<th>Cooperation</th>
<th>Extensive Margin</th>
<th>Intensive Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low-coop. (# of site = 0)</td>
<td>High-coop. (# of site &gt; 0)</td>
</tr>
<tr>
<td>Deluge×Post</td>
<td>0.001* (0.001)</td>
<td>0.017*** (0.005)</td>
</tr>
<tr>
<td>(1/Avg Dist)×Post</td>
<td>-0.115 (0.077)</td>
<td>-0.182 (0.136)</td>
</tr>
</tbody>
</table>

Grid FE | YES | YES | YES | YES | YES | YES | YES |
Period FE | YES | YES | YES | YES | YES | YES | YES |
Grid-specific Trend | YES | YES | YES | YES | YES | YES | YES |
Obs. | 50,652 | 12,645 | 63,297 | 50,652 | 12,645 | 63,297 |

Notes: The sample consists of 7,033 grids from 5000 BCE to 500 BCE. The dependent variables are the existence of civilization characteristics and the intensity of civilization characteristics, respectively. Deluge is an indicator that equals one if the grid is close to the Yellow River, with a distance of no more than 50 kilometers. Post is an indicator that equals one after 2000 BCE. Avg Dist is the average distance from the grid to the archaeological sites within 30 kilometers in the pre-shock period (2500-2000 BCE). Standard errors in parentheses are clustered at the grid level. ***, **, * represent significance levels of 1%, 5%, and 10%, respectively.
Table 6. The Mechanism of the Effect of Great Deluge on Civilization Development:

<table>
<thead>
<tr>
<th>War</th>
<th># of civilizations in the grid</th>
<th># of civilizations around 30km</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td><em>Deluge</em>×<em>Post</em></td>
<td>0.005***</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Grid FE</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Period FE</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Grid-specific Trend</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td># of civilizations×Period FE</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Obs. 63,297 63,297 63,297 63,297

Notes: The sample consists of 7,033 grids from 5000 BCE to 500 BCE. The dependent variables are the number of civilizations within the grid and located up to 30 kilometers from the grid, respectively. *Deluge* is an indicator that equals one if the grid is close to the Yellow River, with a distance of no more than 50 kilometers. *Post* is an indicator that equals one after 2000 BCE. In Columns (1) and (2), the # of civilizations is the number of civilizations within the grid in the pre-shock period (2500-2000 BCE); In Columns (3) and (4), the # of civilizations is the number of civilizations located up to 30 kilometers from the grid in the pre-shock period (2500-2000BC). Standard errors in parentheses are clustered at the grid level. ***, **, * represent significance levels of 1%, 5%, and 10%, respectively.
Table 7. The Mechanism of the Effect of Great Deluge on Civilization Development:

<table>
<thead>
<tr>
<th></th>
<th>Extensive Margin</th>
<th>Intensive Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Slope×Deluge×Post</td>
<td>-0.008***</td>
<td>-0.005***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Slope×Post</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Deluge×Post</td>
<td>0.067***</td>
<td>0.040***</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Grid FE</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Period FE</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Grid-specific Trend</td>
<td>NO</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: The sample consists of 7,033 grids from 5000 BCE to 500 BCE. The dependent variables are the number of civilizations within the grid and located up to 30 kilometers from the grid, respectively. Deluge is an indicator that equals one if the grid is close to the Yellow River, with a distance of no more than 50 kilometers. Post is an indicator that equals one after 2000 BCE. Slope is the mean slope of the grid. Standard errors in parentheses are clustered at the grid level. ***, **, * represent significance levels of 1%, 5%, and 10%, respectively.
Online Appendix

Appendix A. Data Construction

The archaeological site data are taken the Dictionary of Chinese Archaeology. This appendix provides a brief introduction to the data construction process. Compiled by scholars from the Institute of Archaeology of the Chinese Academy of Social Sciences, the Dictionary of Chinese Archaeology contains information on all important prehistoric archaeological sites in China, including their names, locations, excavated artifacts, civilizations, and time of existence. The Dictionary is widely used by scholars in archaeology for its professionalism and accuracy. Figure A1 presents the cover of the Dictionary.

Figure A1. The Dictionary of Chinese Archaeology
Figure A2 shows an example of the data structure. In this example, the name of the site is the Dantu Site, which is located in Dantu Village, Chaohe Town, Wulian County, Shandong Province. The archaeological excavation has unearthed a large number of items, such as city walls, pottery, stone tools, and jade tools. Archaeologists have determined that two late Neolithic human civilizations once existed at the site, namely the Dawenkou Civilization and the Longshan Civilization. The site existed about 4,800 years ago.

![Figure A2. The Information of the Archaeological Site](image)

*Notes: The figure is extracted from p.259 of the Dictionary of Chinese Archaeology.*

To make the data collection process transparent to the readers, we document below the detailed steps in compiling the dataset, the complications that arose, and the decisions that were made:

**Step 1.** We manually read the information on the record to determine whether the site was an early human civilization site.
Step 2. We collect detailed information about the site through the *Dictionary of Chinese Archaeology* and news reports and excavation reports published online as supplementary information.

Step 3. We locate the longitude and latitude information of the sites through *Baidu Maps*.

Step 4. Based on the information of the items excavated from the site, we determine whether there were early civilization characteristics, including class differentiation, large public buildings, use of bronze, written records, and cities. Among them, we assign the value of 1 to the variable of class differentiation if the excavations revealed noble burials or human sacrifices, and 0 otherwise. We assign the value of 1 to the variable of large public buildings if temples, warehouses, meeting halls, squares, water facilities, and public roads are found in the excavation. Similarly, we assign the value of 1 to the variable of use to the bronze if a bronze vessel, a bronze statue, or a bronze smith is unearthed; we assign the value of 1 to the variable of written records if writings on pottery, oracle, or bronze are unearthed; and we assign the value of 1 to the variable of cities if archaeologists classify the site as a prehistoric city.

Step 5. We also identify the early human civilization to which the site belongs from the Dictionary. Specifically, if a site had multiple prehistoric civilizations, we enter their information separately.

Step 6. To determine the duration of a site’s existence, we consider the following scenarios:
1. The *Dictionary* directly mentions the duration of the site’s existence. Then, we directly record the information;
2. A vague description is mentioned in the *Dictionary*. For example, if a record mentions that the site existed from the beginning of the Western Zhou dynasty to the end of the Spring and
Autumn Period. In this case, we use 1046 BCE (the beginning of the Western Zhou period) and 476 BCE (the end of the Spring and Autumn period) to track the duration of the site;

3. Only the duration of the civilization is mentioned. For example, it was mentioned that the late Yangshao civilization once existed in a specific site. In this case, we determine the duration of existence according to the existence time of the prehistoric civilization to which it belongs. Specifically, the late Yangshao civilization existed from 3500-2900 BCE, so we enter the time of existence of the site as 3500-2900 BCE.

4. Only the generic periods of prehistoric civilization are mentioned. For example, if the site is recorded in the Dictionary as one in the Middle Neolithic period, then we classify its existence as from 5000 to 3000 BCE. Similarly, if the site belongs to the Late Neolithic period, i.e., the Copper and Stone Age, we classify its existence from 3000 to 2000 BCE. If it belongs to the Bronze Age, we classify its existence from 2000 to 476 BCE.

It should be noted that the coding method is likely to entail possible measurement error, especially when the site’s existence is uncertain. For example, the Linjia site in Gansu province belongs to the Majiayao culture (3300-2050 BCE), with excavated bronze artifacts. Given that the existence of the Linjia site is uncertain, we supplement it with the existence of the Majiayao culture, i.e., 3300-2050 BCE, with the use of bronze in 3300 BCE. However, the earliest bronze artifact found on the site, a bronze knife, has been dated to around 3000 BCE. We then reconcile the differences by following the archaeological facts and adjust the date of bronze use for the Linjia site to 3000 BCE. Similarly, the earliest cities in China were formed after 3300 BCE, the earliest bronzes in China were found around 3000 BCE, and the earliest writing in China was formed after 1200 BCE. If there is a discrepancy between the variable assignment and the above facts, we adjust the variable assignment accordingly. Moreover, in our empirical analysis, we take 500 years as a period –this extended period helps to avoid the measurement error. Meanwhile, the possible measurement error will bias our results only if it is related to the regional and temporal distribution of the deluge at the same time, which is less likely.
Appendix B. Matching Archaeological Sites with Grids

This section shows how to match archaeological sites to the grids. Ideally, we can map the site directly to the exact grid of its location. However, as explained in Section 3.1, there is a technical problem: we do not have precise boundaries for the archaeological site during its existence, which may extend beyond a single grid, resulting in an inaccurate match between the archaeological sites and the grid.

To address this issue, we first match the archaeological site to the grid based on whether its longitude and latitude are within the grid. If the longitude and latitude of the archaeological site are within the grid, we assume that the archaeological site completely covers the grid and neighboring grids. Figure B1 illustrates the process, where a blue grid indicates the presence of a site, and a yellow grid is a neighboring grid. In the baseline regression, if a site falls within the blue grid A, we consider a total of nine grids, i.e., A-I, as the actual coverage by that site. In the robustness checks, we also test for a stricter definition, assuming that the site only exists within grid A. Our results remain robust.
Figure B1. An example of the matching process

Notes: Each grid has a size of 10 × 10 kilometers, where a blue grid indicates the presence of a site, and a yellow grid is a neighboring grid. In the baseline regression, if a site falls within the blue grid A, we consider a total of nine grids, i.e., A-I, as the actual coverage by that site.
Appendix C. Additional Tables and Figures

Table C1. Indicators of Civilization characteristics

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Photograph</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noble tombs</td>
<td><img src="image1.png" alt="Noble Tomb" /></td>
<td>A noble tomb named “Taosi Bei Tomb” in Shanxi Province.</td>
</tr>
<tr>
<td>Burial of living people</td>
<td><img src="image2.png" alt="Burial Tomb" /></td>
<td>A tomb named “Houshi Shang Tomb” in Shanxi Province, which includes the burial of living people.</td>
</tr>
<tr>
<td>Large buildings</td>
<td><img src="image3.png" alt="Large Buildings" /></td>
<td>Palace buildings in Taosi Site in Shanxi Province.</td>
</tr>
</tbody>
</table>
Bronze ware in a tomb named “Wenxi Jiuwutou Tomb” in Shanxi Province.

Inscription in the “Yuanqu Bei Baie Tomb” tomb in Shanxi Province.

Walled city in Taosi Site in Shanxi Province.
Figure C1. The location of the Lajia site

Notes: The figure is retrieved from Wu et al. (2016).
Figure C2. Sample area in Yellow River Basin after Cropping

Notes: This figure shows the sample area after cropping. The grids are those located in the lower reaches of the Jishi Gorge and are within 100 kilometers of the Yellow River. The geographical data were obtained from the Center for Resource and Environmental Science and Data of the Chinese Academy of Sciences (https://www.resdc.cn/).
Figure C3. The Distribution of Treatment and Control Groups in Yellow River Basin

Notes: This figure illustrates the distribution of the treatment and control groups. The striped grids indicate the treatment group, while the blank grids indicate the control group. The geographical data were obtained from the Center for Resource and Environmental Science and Data of the Chinese Academy of Sciences (https://www.resdc.cn/).
Figure C4. The Effect of Great Deluge on State Origin: Event Study

Notes: The figure depicts the differences in the development of the civilization characteristics between the grids close to the Yellow River and those otherwise before and after the deluge. The markers and connected lines represent the ordinary least squares estimates and 95 percent confidence intervals based on standard errors clustered at the grid level. The dependent variable is the intensity of the civilization characteristics. The reference group is the period 2500-2000 BCE. The regression considers grid fixed effects and period fixed effects.
Figure C5 Yangtze River Basin

Notes: This figure shows the construction of grids in the Yangtze River basin. The geographical data were obtained from the Center for Resource and Environmental Science and Data of the Chinese Academy of Sciences (https://www.resdc.cn/).
Figure C6. Visualization of the Cooperation Hypothesis

Notes: The figure presents the distribution of high-cooperation-opportunity grids. In the figure, the red circles indicate the archaeological sites with early civilization characteristics after the deluge. The black grids indicate those with high cooperation opportunities.
**Figure C7. Slope of the Yellow River Basin**

*Notes:* This figure shows the slopes along the Yellow River Basin. Darker shades indicate higher slopes. The slope data is obtained from Global Agro-Ecological Zones ([https://www.fao.org/home/en/](https://www.fao.org/home/en/)).
Table C1. The Effect of Great Deluge on State Origin: Lagged Outcomes

<table>
<thead>
<tr>
<th></th>
<th>Extensive Margin</th>
<th></th>
<th>Intensive Margin</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Deluge×Post</td>
<td>0.005</td>
<td>0.005</td>
<td>0.010</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.007)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Grid FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Period FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Deluge * Period Trend</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Grid-specific Trend</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Obs.</td>
<td>56,264</td>
<td>56,264</td>
<td>56,264</td>
<td>56,264</td>
</tr>
</tbody>
</table>

Notes: The sample consists of 7,033 grids from 5000 BCE to 500 BCE. The dependent variables are the existence of civilization characteristics and the intensity of civilization characteristics in the previous period, respectively. Deluge is an indicator that equals one if the grid is close to the Yellow River, with a distance of no more than 50 kilometers. Post is an indicator that equals one after 2000 BCE. Standard errors in parentheses are clustered at the grid level. *, **, *** represent significance levels of 10%, 5%, and 1%, respectively.