Horses, Battles, and the State*

Motohiro Kumagai†

This draft: September 2022

Abstract

This research explores the effect of horses on the state and battles. It exploits multiple exogenous sources of variation in the adoption of horse riding on the battlefield: (i) the regional variation in the spread of horse-riding technology, (ii) the variation in climatic suitability of horse survival, and (iii) the change in the availability of horses in the Americas during the Columbian Exchange. The research, using these exogenous variations and multiple datasets spanning several millennia, provides robust evidence of the significant impact of horses on the state and battles. Rich historical accounts support the critical role of cavalry in state formation through its influence as a source of military power.

Keywords: State, Nation, Centralization, Hierarchy, Territory, Battle, Conflict, Institution, Geography, Climate, Biogeography, Horse, Weapon

JEL Codes: N00, O10, O43, O44, Q34, Z13

*I thank Guillaume Blanc, Oded Galor, Louis Putterman, Stelios Michalopoulos, and David Weil for their valuable comments and suggestions. I also thank Peter Turchin and James S. Bennet for sharing their code and data, and Shuhei Kitamura and Masahiro Kubo for sharing their data with me. I am grateful to seminar and conference participants at Brown University and EBES. All mistakes are my own.

†Department of Economics, Brown University, Providence RI. Email: motohiro_kumagai@brown.edu

1 Introduction

Today's wealthiest countries have strong centralized states. On the other hand, the world's poorest people tend to live in areas where the state is fragile or failing. This contrast between the rich and the poor also goes far back in historical times. Despite the critical role of the state in developing the wealth of nations, the complexity of the origins has prevented them from being fully explored. State history is shown to explain much variation in economic development today (Putterman and Weil, 2010; Borcan et al., 2018; Michalopoulos and Papaioannou, 2013); therefore, understanding the origins of the state is fundamental to understand the inequality today.

This paper explores the effect of horses on state formation and on historical battles. As many historical studies and narratives demonstrate, horses gave overwhelming military power to people, and thus, those who had access to horses succeeded in conquest and territorial expansion. Centralization and hierarchy, therefore, tended to emerge in regions where horse riding was available for military purposes. As is consistent with this hypothesis, this study establishes the positive impacts of horse riding on the process of both state building and on historical battles.

The exploration of the impact of horse riding on the state and on battles is associated with significant empirical hurdles. First, the observed relationship between them may reflect the effects of state formation and warfare on the invention and adoption of horse riding, rather than the other way around. Second, the effects of institutional, cultural, and geographic characteristics on the joint evolution of equestrianism, the state, and warfare may have governed the observed relationship between these forces. Given these empirical hurdles, a desirable empirical setting requires an exogenous source of variation in the use of cavalry on the battlefield.

Given the potential endogeneity of the use of cavalry, state formation, and battles, this study exploits multiple sources of exogenous variation in adopting horse riding. First, the study utilizes geographical variations in the spread of equestrianism on the battlefield across regions, based on which I conduct two-stage least squares (2SLS) estimates. Second, the study exploits a natural experiment associated with the expansion of the availability of horses in the Americas during the Columbian Exchange. I conduct panel analyses based on observations in the Americas, which experienced the exogenous change in the availability of horses caused by the Exchange.

The instrumental variables are constructed based on plausibly exogenous variations in the spread of horsemanship that emerge from two distinct components. The first component is an exogenous source of variation in the geographical diffusion of horse-riding technology. The second component is an exogenous source of variation in the availability of native horses, conditional on geographical and climatic characteristics.

In light of the importance of the metal bit for the emergence of cavalry on battlefields, the study takes advantage of historical and archaeological accounts regarding the geographical diffusion of the metal bit (Drews, 2004). In particular, it exploits the distance from Tell el-Ajjul, where the oldest horse metal bit was found, as an exogenous source of variation in the spread of horse riding on the battlefield.

The validity of the distance from Tell el-Ajjul as one of the two instrumental variables is enhanced by several additional factors. First, as is consistent with historical and archaeological accounts, the distance from the lower Volga-Don region, where modern horses were first domesticated, is not significantly correlated with the spread of military equestrianism.[1] Second, the association between the distance from Tell el-Ajjul and the spread of military horsemanship is unaffected by the spread of iron, which had particular importance in early civilization. Third, this association is likewise unaffected by the distance from pristine sites of agricultural transition. Tell el-Ajjul is located in Gaza, close to places of early civilization. Therefore, accounting for the distance from the independent sites of agricultural transition alleviates the concern that the association between horse riding and the distance from Tell el-Ajjul reflects the relationship with the early civilization. Fourth, distances from major centers of economic and political power at roughly the same time as the invention of the first metal bit are not significantly correlated with the spread of equestrianism in war. Finally, the distance from Tell el-Ajjul appears to be orthogonal to economic development and historical battles before the invention of the metal bit.

The second component of the instrumental variables is motivated by the hypothesis that there were more horses available in regions where horses were native and where the environment was suitable to their survival. Thus, conditional on geographical characteristics, the analysis exploits exogenous variation in the prehistorical distribution of horses and variation in climatic suitability for their survival. These features identify bioclimatic conditions conducive to the availability of horses in prehistoric times and thus for the emergence of horse riding. In particular, I construct an index of the historical availability of horses that captures both extensive and intensive margins.

One may be concerned that the association between this index and the diffusion of horse riding on the battlefield captures only agriculture and historical trade. However, the estimate of the index of the availability of horses is stable and remains highly significant if multiple measures of agriculture and historical trade are added. These variables are the average and standard deviation of land productivity, the average and standard deviation of caloric suitability, years elapsed since the agricultural transition, the distance from the closest long-distance historical trade route, and a dummy of historical transport mammals. The index,

[1] As describe in detail in the empirical methodology section, the primary purpose of horse domestication was to eat their meat. Although some people rode horses before the invention of a metal bit, it was recreational, ritual, or athletic riding, which never led to secure riding in warfare (Drews, 2004.)

therefore, predicts the diffusion of horse riding on battlefields through a different path than agriculture and historical trade.

Moreover, the research exploits a natural experiment associated with the Columbian Exchange to mitigate possible concerns about omitted variable bias. It uses the exogenous change in the availability of horses in the Americas resulting from the Columbian Exchange. In the Americas, horses became extinct during the Pleistocene and Holocene transition; hence, horses were unavailable until the Columbian Exchange. European immigrants brought horses to the Americas, and horses spread across the continents. Therefore, only after the Columbian Exchange horses became available, giving rise to the exogenous variation in the index of the availability of horses.

This study establishes the impact of horses on state formation and warfare in multiple layers. The first layer is based on cross-sectional data on the location of ancient cities from two different data sets. With these data and the instrumental variables, the analysis shows a positive association between time since the emergence of cavalry and the presence of ancient cities.

The second layer employs data on the history of the state compiled by Borcan et al. (2018). This data is based on the present-day boundaries of 159 countries and provides a comprehensive history of state formation. Using this data and the instrumental variables, the analysis shows a significant, positive association between time since the emergence of cavalry and the history of the state across modern countries.

The third layer is based on the *Ethnographic Atlas*, which covers the cultural and institutional features of more than 1,200 precolonial societies worldwide. I use multiple variables as proxies for the state in traditional societies. These variables are jurisdictional hierarchy beyond local communities, social stratification, and size of local communities. With this data and the instrumental variables, I demonstrate the positive relationships between time since the emergence of cavalry and the state in preindustrial societies.

The fourth layer again utilizes data by Borcan et al. (2018). They also provide a flow index of the state for every half-century from 3500 BCE to 2000 CE. Instead of using instrumental variables, I exploit the exogenous variation in the availability of horses in the Americas resulting from the Columbian Exchange. The Columbian Exchange is what caused the exogenous change in the availability of horses in the Americas. As is consistent with the cross-sectional results, the panel regressions confirm that a greater change in the availability of horses had a positive impact on state formation.

The fifth layer uses data compiled by Mayshar et al. (2022), who collected information on radiocarbon estimates dating various prehistorical archaeological sites from Whitehouse and Whitehouse (1975). Following their methodology, I conduct difference-in-differences estimates. This analysis shows that the Neolithic transition led to more visible traces of complex hierarchical societies in areas where agriculture was more likely to start with horse riding, controlling for the distances from the closest agricultural center and from the closest long-distance trade route.

In the following layers, I explore the impact of horse-riding on historical battles, exploiting the *World Historical Battles Database*. This database, recently compiled by Kitamura (2021), provides rich information on conflicts covering the entire recorded history and the entire world. The sixth layer of the analysis uses this data in a cross-sectional setting and conducts 2SLS regressions. Based on virtual countries (cells) at the 1×1 cell level, it demonstrates the significant positive association between horse-riding and historical battles.

The last layer utilizes the exogenous variation in the availability of horses in the Americas resulting from the Columbian Exchange. I run panel regressions by creating virtual countries at the 1×1 level. As is consistent with the cross-sectional results, the analysis establishes that an exogenous increase in the availability of horses caused an increase in the historical frequency of battles.

The empirical analysis accounts for a wide range of confounding geographical and climatic characteristics and continent-/country-/time-fixed effects. Moreover, to better identify the effect of horses on the state through their impact on historical battles, it controls for several measures of agriculture and historical trade. First, it accounts for the average and standard deviation of land productivity as well as the distance from the closest center of agriculture. When possible, it further controls for dependence on agriculture, intensity of agriculture, and time since agricultural transition. Second, it accounts for the distance from the closest historical long-distance trade route as well as the presence of transport mammals. When possible, it further controls for motifs related to trade. Furthermore, in light of the concern that the diffusion of horse-riding technology captures that of iron, the study also accounts for the spread of iron. Reassuringly, the established association between horses and both state formation and battles is unaffected by the inclusion of these variables.

In summary, the empirical analyses provide repeated evidence that horses had a significant effect on state formation and on historical battles. This finding is supported by rich historical documents and narratives, giving further credence to the positive impact of horses on the state and warfare.

The paper relates to several strands of the literature. First, it connects to the literature exploring the prehistoric origins of inequality. Diamond (1997) argues that differences in the availability of domesticable plants and animals led to differences in the timing of the agricultural transition, which then caused variations in the wealth of nations. Hibbs and Olsson (2004) and Olsson and Hibbs (2005) empirically test the Diamond thesis and find that the timing of the agricultural transition explains non-trivial parts of prosperity before modernity.

Second, the study is related to the literature on the origins of early forms of hierarchy and the state. Fenske (2014) demonstrates that ethnic groups whose homelands are characterized by ecological diversity developed centralized hierarchy. Borcan et al. (2018) find a hump-shaped relationship between state history and current GDP per capita. Link (2021) shows that regions with transport mammals developed long-distance trade and social hierarchy. Mayshar et al. (2022) challenge the traditional surplus hypothesis. They find that cereals are easily stored and appropriated, thus facilitating taxation and contributing to the emergence of hierarchy.

Finally, this research contributes to the literature on the deep determinants of conflicts. Conventionally, scholars have examined the association between ethnolinguistic fragmentation and conflicts. Early studies about the effects of ethnic, linguistic and religious fractionalization on conflicts is inconclusive (Collier and Hoeffler, 1998; Collier and Hoeffler, 2004; Fearon and Laitin, 2003). Montalvo and Reynal-Querol (2005) and Esteban et al. (2012) demonstrate that polarization leads to the risk of civil conflict. Arbatlı et al. (2020) show that population diversity is a determinant of civil war. My study contributes to this line of research by providing empirical evidence that the availability of horses increased the risk of battles.

2 Historical Evidence

The importance of horses for warfare cannot be overstated. For several thousand years in the past, horses were the single most important instrument of battles. The main use of horses was a shock weapon and they were the ancient and medieval equivalent of modern planes and tanks. Horses gave warriors superior height and speed. The combination of these forces almost always overwhelmed opponents (Chamberlin, 2010). Moreover, they could pull supply wagons and artillery. Once horse-riding became possible, they were also used for reconnaissance and scouting (Ebrey et al., 2006).

War wagons and chariots played an important role in battles for nearly thousand years from about 2000 BCE. The first chariot battle for which we have a record was the battle of Megiddo in northern Palestine in 1469 BCE (Keegan, 2011). The oldest manual on training horses for chariot warfare was written around 1350 BCE by the Hittite horsemaster, Kikkuli and the existence of this text is a testament to the importance of horses as an weapon in ancient times (Chamberlin, 2010).

From about 1000 BCE the importance of chariot started to decline and horseback riding had appeared in battle fields. The earliest record of extensive use of cavalry was by the Assyrians from 900 BCE (Law, 2018). Riding appeared in Babylon in 1200 BC and cavalry appeared in armies some time later. The Etruscans, an ancient Italy, used horse-riding around 700 BCE. The Persians had cavalry in their armies at least in 500 BCE. Romans depended on cavalry by the time of Roman empire (Ellis, 2004). During the Warring States period (403-221 BCE), the Chinese began to use cavalry against rival states (Ebrey et al., 2006).

Alexander the Great is one of the best example of the first great horse warriors (information in this paragraph is from Chamberlin, 2010). Alexander was born in Macedonia in Europe and came to power in 336 BCE. When he took over, he inherited a strong army of 400 light horse scouts and over 3300 heavy horse cavalry from his father Philip. In addition to these horses, he had obtained Thessalian, Thracian and Persian cavalry totaling over 6000 by the time he began his exceptional campaign. During the campaign, he fought many battles including the battle of the Granicus in 334 BCE, the Battle of Issus in 333 BCE and the battle near the Tigris River at Gaugamela in 331 BCE. He continued to expand his territory and gather horses from opponents he defeated. When Alexander died in 323 BCE, he had established the largest empire the world ever known, stretching from Alexandria to Ferghana and from Macedonia to India. He accomplished this feat all with horses.

During the Late Antiquity and the Middle Ages the Byzantine Empire, which was heir to the Roman Empire, was the most powerful economic, cultural and military force in the Eastern Europe (information in this paragraph come from Hyland, 1996). The empire's economy was largely directed to producing high-quality horses and the Theodisian code includes many laws about horses such as breeding, acquisition by the military, fodder requisition, and so forth. By the beginning of the Middle Ages, the armies largely depended on horses and infantry was no longer the most important. The empire had more than 150000 cavalry in its later years.

Another great example of horses and conquest is the Mongol armies of Genghis Khan and his successors, originating in Mongolia in East Asia (information in this paragraph is taken from Hyland, 1996 and Ellis, 2004). Genghis Khan united the tribes for the first time in history and began a campaign. During the campaign, he and his successors conquered the Kwarezmian Empire in 1220, Russia in 1237, Baghdad in 1258 and China in 1279, finally establishing Yuan dynasty. The Mongol armies amplified by incorporating subjugated peoples and it enabled them to continue the expansion. This conquest was not possible without horses. For example, when they fought with the Kwarezmian, the army numbered 150000, of which 140000 were cavalry. During 1221, Genghis Khan went from Bamian to Ghanza via Kabul in only two days, covering 130 miles. In 1241, the army invaded Hungary and marched 180 miles in only three days. This phenomenal mobility was the key of the Mongol's successful conquest and their sturdy horses gave this extraordinary speed.

Horses also constituted the great part of the Muslim military (information in this paragraph is from Hyland, 1996 and Ellis, 2004). In 622, the Arabs began migration out of the

Arabian peninsula and horses were crucial in battles such as Yarmūk in 636, Siffīn in 636 and Quādissiya in 637. The central part of Muslim armies was Mamluks who were originally slaves and freed to serve in the Muslim army. At that time, mobile tactics, advanced breeding of horses and detailed training manuals were available and they made Mamluks cavalry a strong fighting force. Another feature of the Arab armies is that non-Arabs were important to the armies. Although such warriors were not allowed to ride horses in the first, the use of horses by them gradually became common. Under the Ummayads (659-750) and the Abassids (750-1258), non-Arab mercenary cavalry occupied a great part of the Muslim armies. The Ottoman Empire, founded in 1299 heavily depended on horses. Although they had gunpowder weapons, the role of gunpowder was limited and the most effective weapon were cavalry that made up the bulk of the armies.

In China, under the Shang and Chou dynasties (1700 BCE - 1100 BCE), chariots were the key weapon, and many northern states introduced light cavalry into their armies in the Warring States era from 402 to 221 BCE (Ellis, 2004). Cavalry were particularly important because they did reconnaissance, pursed fleeing soldiers, cut supply lines and pillaged the countryside, and hence states that incorporated horses into their armies obtained military advantage (Ebrey et al., 2006). During the whole period until the beginning of the nineteenth century except for the Mongol conquest in 1279, China successfully repelled nomadic threats from the northern frontiers. Although their tactics varied depending on periods, horses were constantly important weapon. During the Tang Dynasty (618-907), for example, Chinese armies usually were composed of many cavalries. The number of available horses at that time was at least 300000 (Ellis, 2004).

In Africa horses were not native to the continent, and horses were introduced from Asia (information in this paragraph come from Law, 2018 unless other documents are cited). Horses and war chariots were first introduced into Egypt possibly by the Hyksos. The historical heritage indicate that they became common in Egypt from 1600 BCE. Tuthmose III, who was an Egyptian military pharaoh, established a standing army with chariotry and infantry and fought the first chariot battle at Megiddo in Syria in 1460 against allies of the Hyksos (Chamberlin, 2010). Many other empires in West Africa have originated through conquest by invading bands of horsemen. The Ghana Empire, the Songhai state, the Hausa kingdoms and the Zaghawa and Saifawa states were established around the middle centuries of the first millennium CE. From the twelfth to the fifteenth centuries, the Mossi-Dagomba group of kingdoms, the Bariba states, the Nupe and Jukun kingdoms, the Yoruba states and Benin emerged. In the sixteenth century, the kingdom of Gonja was founded. It was horses that gave invaders their military advantage over the peoples whom they conquered.

3 Data

In this section I describe several data that are utilized to explore the association between horses and state/battle. The first four subsection describe data on the state, followed by descriptions about data on historical battles and horse-riding in battle fields.

3.1 Ancient Cities

To capture hierarchical complexity back in time at the fine scale, I draw on data on the location of ancient cities, following Mayshar et al. (2022). The first data is Degroff (2009), which provides information on the location of cities and towns that were founded before 400 CE. The second data is Reba et al. (2016), which provide information on the location of urban settlements from 3700 BCE to 2000 CE. As in Mayshar et al. (2022), I select two points in time: 500 BCE and 450 CE. Using these database, I construct a dummy variable that takes 1 if there is an ancient city in a cell, and 0 if otherwise. I also create a variable of log distance to the nearest ancient city for each cell. I conduct the analysis of ancient cities, making a virtual country at the 1×1 cell level. Summary statistics is in Table B1.

3.2 State History and State Index

Borcan et al. (2018) provide the dataset on the state for the area of 159 modern-day countries for every half-century from 3500 BCE to 2000 CE. For each country and half century, they create an index of the state (*State Index*) by combining three dimensions such as hierarchy, autonomy and territory. The database also provides a stock variable of the state (*State History*). This variable is constructed by aggregating *State Index* over the entire period and thus it is the comprehensive index of the cumulative state history. I use *State History* for a cross-country analysis, while I utilize *State Index* between 1000 CE and 2000 CE for a country panel analysis. Summary statistics are in Table B2 and B3.

3.3 Ethnographic Data

The most comprehensive ethnographic data is the *Ethnographic Atlas* (EA) by Murdock (1967). This atlas represents 1267 societies from around the world and all groups are observed before industrialization or European contact. The sample is global, with an emphasis on North American and African groups. European groups are under-represented. Giuliano and Nunn (2018) extend the original data, adding several ethnic groups from Europe. This increases the sample to 1309 ethnic groups. The database contains information on cultural, institutional and economic characteristics. As proxies of the state, I use the variables "v31" (mean size of communities), "v33" (jurisdictional hierarchy beyond local community", and "v66" (class stratification). To calculate geographical variables, I create a 50 km buffer zone with the geocoordinate of an ethnic group being the centroid. Then I aggregate values in cells within the buffer area. Summary statistics is in Table B4.

3.4 Radiocarbon-Dated Archaeological Sites

Whitehouse and Whitehouse (1975) provide global data on radiocarbon-dated prehistoric archaeological sites. Mayshar et al. (2022) georeference these sites and classify them according to whether or not they predate the Neolithic transition. As a result, they obtain a list of 825 sites that belong to pretransition years and 3309 sites that belong to the posttransition years. I draw on the data compiled by Mayshar et al. (2022). Based on their data, I create a virtual country at the 1×1 cell level. Each cell includes information on the number of pre-Neolithic and post-Neolithic sites. In the empirical analysis, I either use all ruins, or only sites of prehistoric settlements.[2] Summary statistics is in Table B5.

3.5 Historical Battles

The *World Historical Battles Database* (WHBD) is a recently compiled database that provides rich information on conflicts covering the entire recorded history and the entire world (Kitamura, 2021). The main source of this database are crowd-based sources and according to the author, the overall coverage (temporal and spatial) is the largest among the existing historical conflict data that are publicly available. The spatial coverage is global and temporal coverage is from 2500 BCE until present. It includes 7741 battles with their geolocation and year of occurrence. With its rich information, I create a virtual country at the 1×1 cell level. For each cell, I calculate an indicator of battles, the count of battles and log distance to the nearest battle point. The characteristics of this database enables to examine the association between horses and battles going back to older history at the finer scales than previously available data on historical conflicts.

3.6 Time since Cavalry Emergence

To capture the historical use of horses in battles, I construct a measure of time elapsed since cavalry emergence, using data from Turchin et al. (2016) and Turchin et al. (2021). They collect a variety of sources about mounted warfare and map the spread of horse-riders in military operations. Importantly, their data is not about when people started riding horses,

[2] The atlas classifies these ruins according to 10 different categories: (1) undifferentiated sites and findspots; (2) settlements; (3) funerary monuments; (4) religious monuments; (5) caves and rock shelters; (6) cave art and rock reliefs; (7) hoards and votive deposits; (8) mineral sources; (9) mineral workings; and (10) sites that combine several of the above categories.

but when horse-riding became systematically used in warfare. In particular, they do not include cases where only a tiny proportion of the army used horses. To be included in their data, cavalry had to constitute at least five percent of the overall army. I first create a 1×1 raster file of time elapsed since cavalry emergence, and then aggregate values at arbitrary scales such as country, ethnic homeland, etc. Distribution of the number of years since cavalry emergence as of 2000 CE is depicted in Figure 1.

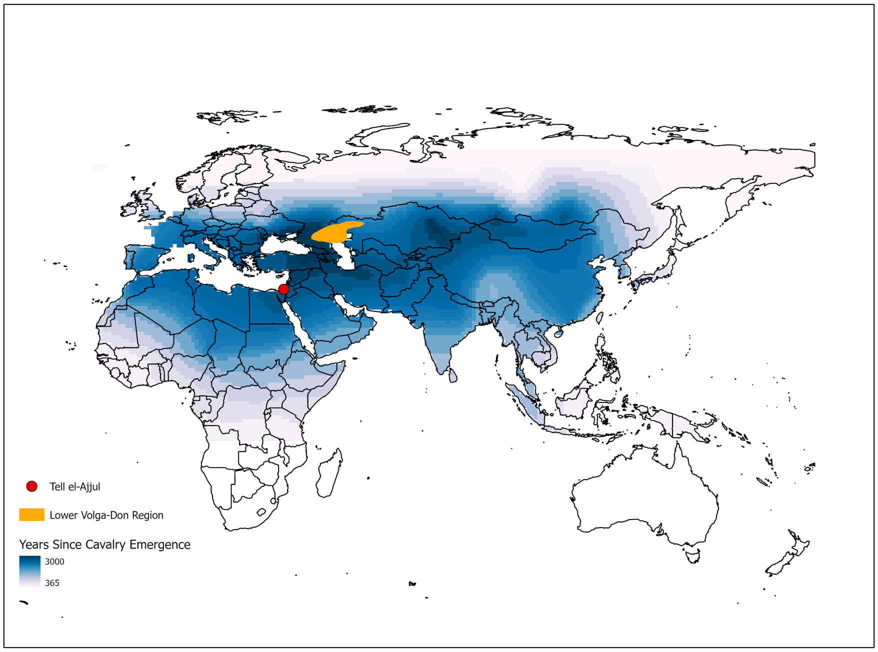

Figure 1: Time Elapsed Since Cavalry Emergence, the Location of Tell el-Ajjul and the Lower Volga-Don Region

4 Empirical Methodology: Instrumental Variable Approach

The relationship between horses and state/battle may reflect the impact of riding on the battlefield on state formation and combat, as well as the impact of socioeconomic development on the demand for and supply of horses. In addition, institutional, geographic, and cultural characteristics may influence the co-evolution of horseback riding and state/battle, and may govern the association of these characteristics. Hence, in light of the potential endogeneity of horses and state/battle, this research exploits exogenous variation in the adoption of horse-riding to establish the causal effect of horses on state formation through their effect on battle. In what follows, I present the two components of the instrumental variables. This research further exploits the natural experiment associated with the Columbian Exchange. It also use panel data on radiocarbon-dated prehistoric and posthistoric archaeological sites and conduct a difference-in-difference regression. These empirical methodology are described in the section of empirical result.

4.1 Spread of Horse-Riding in Battle Fields

The first component is motivated by the historical account of the gradual diffusion of horse-riding starting about 1000 BCE (information in this subsection come from Drews (2004) unless other documents are cited). The invention of a metal bit enabled people to safely control horses, allowing them to ride horses in battle fields. Reflecting the role of the metal bit in horse-riding, the study exploits the distance from Tell el-Ajjul, where the first metal bit was found, as an instrument for years elapsed since cavalry emergence.

Secure riding is a prerequisite if a rider wants to concentrate his attention on his weapons and his opponents. Before the invention of a metal bit, secure riding was impossible and thus horses were not ridden in battle fields. A bit is consisted of a mouthpiece and two cheekpieces. A rider directs a horse to left or right and brings it to a stop through this instrument. However, if a horse worked the snaffle back to its premolars and then clamped them down, the bridling system becomes neutralized, and the rider loses his control on the horse. This happens much more often if an organic bit is used because it is easily frayed by wear. With a metal bit, the rider does not need to worry about this. As Bokovenko (2000) concludes, the development of a more reliable type of bronze bridle enabled the mastery of horse-riding. Tell el-Ajjul, located in the Gaza Stripe, is the place where a bronze bit was first invented in the fifteenth century BC (Figure 1 shows the location). Therefore, I exploit the distance from Tell el-Ajjul to capture the geographical diffusion process of horse-riding in battle fields.

Horse-riding on the battle field started much after horses were domesticated. Archaeological evidence support that the villagers' purpose in domesticating horses was to eat them. Horse bones occupy most parts of the bones found at a number of sites such as Repin on the Don (early third-millennium) and Botai in northeastern Kazakhstan (3500-3000 BC). By 3000 BC, horsemeat was regularly eaten west of the Dnieper: a tenth of the bones found in Late Tripolye settlements were horse bones, and horses were probably domestic. At Csepel-Háros in Hungary, horses were the most common food animal by the middle of the second millennium. A number of evidence for the consumption of horsemeat after domestication indicates that domestication of horses in itself does not necessarily predict the emergence of horse-riding. Rather, these evidence suggest a possibility of the existence of a more important factor resulting in horse-riding. The lower Volga-Don region is recently identified as the place where modern horses were first domesticated (Librado et al., 2021), as shown in

Figure 1.

Table 1: The Geographical Diffusion of Cavalry from Tell el-Ajjul

	Time since Cavalry Emergence								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log Dist. to Tell el-Ajjul	-322.922***	-389.139***	-285.839***	-438.931***		-405.925***		-360.024***	-341.509***
	(72.284)	(54.041)	(45.785)	(100.714)		(71.268)		(91.912)	(68.625)
Log Dist. to the Lower Volga-Don					-64.475	-24.988			-16.464
					(63.414)	(56.442)			(56.428)
Time since Iron Emergence (400 CE)							117.289***	73.968***	70.938***
							(29.213)	(25.939)	(25.636)
Continent FE		✓	✓						
Geographical Controls			✓	✓	✓	✓	✓	✓	✓
Country FE				✓	✓	✓	✓	✓	✓
Adjusted R^2	0.196	0.439	0.660	0.784	0.767	0.785	0.775	0.789	0.790
Observations	7260	7260	7260	7260	7260	7260	7260	7260	7260

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Dependent variable is years elapsed since cavalry emergence. Distances are taken logarithm. Continent dummies are Africa, Asia, and Europe. Geographical controls are log distance from the closest agricultural center, latitude, longitude, latitude $\times$ longitude, average of elevation, average of land productivity, standard deviation of elevation, standard deviation of land productivity, island dummy, average of temperature, average of precipitation, standard deviation of temperature and standard deviation of precipitation. *** p<0.01, ** p<0.05, * p<0.10.

To examine the association between horse-riding on the battle field and the distance to the Tell el-Ajjul, I conduct several empirical exercises, constructing virtual countries (cells) at the 1×1 cell level. Indeed, in line with the historical account, the distribution of years elapsed since cavalry emergence is indicative of a geographical diffusion process from Tell el-Ajjul (Table 1). As reported in column 1, there is a highly significant negative unconditional association between time since cavalry emergence and the distance to Tell el-Ajjul. This association is robust to accounting for continent fixed effects (column 2), a number of geoclimatic characteristics (column 3), and country fixed effects rather than continent fixed effects (column 4). Consistent with the historical account that domestication of horses in itself is not conducive to horse-riding on the battle field, the diffusion pattern of cavalry is not significantly correlated with the distance to the lower Volga-Don region, where modern horses were first domesticated (column 5). As reported in column 6, the highly significant negative association between cavalry diffusion and the distance from Tell el-Ajjul is unaffected when the distance from the lower Volga-Don region is accounted for. Column 7 shows the positive association between years elapsed since cavalry and iron emerged. The relationship between diffusion of cavalry and the distance from Tell el-Ajjul is stable and highly significant if spread of iron is added (column 8). This indicates that the association between cavalry emergence and a bronze bit is independent of that between cavalry emergence and iron spread. Column 9 adds all the variables and it shows that the diffusion of cavalry is strongly associated with the distance from Tell el-Ajjul.

Table 2: The Geographical Diffusion of Cavalry: Inclusion of the Agricultural Origins

	Time since Cavalry Emergence					
	(1)	(2)	(3)	(4)	(5)	(6)
Log Dist. to Tell el-Ajjul	-386.032***	-289.910***	-437.183***	-414.358***	-403.685***	-363.786***
	(49.096)	(81.377)	(45.169)	(43.587)	(47.657)	(69.656)
Log Dist. to Near East		-74.700*				-73.924*
		(38.051)				(38.040)
Log Dist. to Northern China			-82.524**			-75.129**
			(38.648)			(31.225)
Log Dist. to Southern China				-39.322		-14.492
				(29.809)		(13.538)
Log Dist. to West African Sub-Sahara					35.747**	32.136*
					(17.983)	(16.381)
Country FE	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.736	0.739	0.747	0.739	0.738	0.752
Observations	8506	8506	8506	8506	8506	8506

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Dependent variable is years elapsed since cavalry emergence. Distance are taken logarithm. *** p<0.01, ** p<0.05, * p<0.10.

The plausibility of the use of the distance from Tell el-Ajjul as an instrumental variable for time since cavalry emergence is further enhanced by a few additional empirical findings. One concern is the geographical closeness between Tell el-Ajjul and places of early civilization. In light of the critical role of agricultural transition in early civilization, Table 2 examines the robustness to the inclusion of the distance from independent agricultural transition places.[3] Column 1 shows the negative significant relationship between the distance from Tell el-Ajjul and time since cavalry emergence, accounting for country fixed effects. This association is very stable if the distance from Near East, Northern China, Southern China and West African Sub-Sahara are sequentially added (columns 2-5). Furthermore, as is shown in column 6, the estimate of the distance from Tell el-Ajjul is unaffected if these variables are added together.

[3] Purugganan and Fuller (2009) identify seven accepted regions of pristine agricultural transition across the world. Because the measure of cavalry diffusion is available only in the Old World, I use four places of independent agricultural centers in the Old World.

Table 3: The Geographical Diffusion of Cavalry: Insignificance of Distance from Other Places

	Time since Cavalry Emergence										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Log Dist. to Tell el-Ajjul	-386.032***	-276.799***	-422.527***	-265.091***	-471.631**	-442.446***	-264.495***	-406.681***	-254.960**	-444.204**	-441.905***
	(49.096)	(92.982)	(114.332)	(99.068)	(211.849)	(39.658)	(98.910)	(100.398)	(105.471)	(195.269)	(39.008)
Log Dist to Eridu		-125.270					-134.540				
		(80.716)					(83.706)				
Log Dist. to Itjtawy			38.691					26.043			
			(118.022)					(103.786)			
Log Dist. to Susa				-135.682					-142.373		
				(86.918)					(93.028)		
Log Dist to Knossos					94.105					68.086	
					(234.245)					(222.951)	
Log Dist to Erligang						-151.544*					-150.444
						(88.025)					(92.381)
Log Dist. to Agricultural Frontier							-22.002	-20.417	-21.727	-19.607	-0.690
							(15.810)	(15.299)	(15.788)	(13.585)	(11.843)
Country FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.736	0.738	0.736	0.739	0.737	0.747	0.741	0.739	0.742	0.739	0.747
Observations	8506	8506	8506	8506	8506	8506	8506	8506	8506	8506	8506

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Dependent variable is years elapsed since cavalry emergence. Distance are taken logarithm. *** p<0.01, ** p<0.05, * p<0.10.

As established in Table 3, time since cavalry emergence is not significantly correlated with the distances from other major cities that existed at the roughly same period when a metal bit was invented. These cities are Eridu (the city of Sumer civilization), Itjtawy (the capital of Egypt in the middle kingdom), Susa (the capital of Elam civilization), Knossos (the site of Bronze Age Aegean civilization), and Erlingang (the site of Bronze Age civilization in China). This result lends credence to the unique role of the invention of a metal bit at Tell el-Ajjul in the diffusion of horse-riding in battle fields.

Moreover, in contrast to its pivotal role after the invention of a metal bit, the distance to Tell el-Ajjul is not significantly correlated with economic development and battle occurrence before the invention. As is shown in Table 4, the distance to Tell el-Ajjul is not significantly related to population count in 2000 BC (columns 1 and 2), population density in 2000 BC (columns 3 and 4), urbanization in 2000 BC (columns 5 and 6), and the log distance to the closest battle before 2000 BC.

4.2 Historical Horse Index

The second component of the instrumental variables is motivated by the hypothesis that in a region where horses were native and environment was suitable for their survival, there were more available horses. Thus, the analysis exploits variation in prehistorical distribution of

Table 4: Orthogonality of Distance from Tell el-Ajjul and Development before the Invention of the Metal Bit

	Population Count		Population Density		Urbanization		Log.Dist to Battle	
	(1) 2000 BCE	(2) 2000 BCE	(3) 2000 BCE	(4) 2000 BCE	(5) 2000 BCE	(6) 2000 BCE	(7) 2000 BCE	(8) 2000 BCE
Log Dist. to Tell el-Ajjul	14.024	-23.415	0.158	-0.306	0.828	-0.186	0.013	0.162
	(21.716)	(29.328)	(0.307)	(0.404)	(0.719)	(0.459)	(0.110)	(0.140)
Log Dist. to Agricultural Frontier	-25.611*	-14.287*	-0.363*	-0.208*	-0.410**	0.022	0.172***	0.121***
	(13.168)	(7.939)	(0.187)	(0.112)	(0.185)	(0.269)	(0.050)	(0.040)
Country FE	✓	✓	✓	✓	✓	✓	✓	✓
Geoclimatic Controls		✓		✓		✓		✓
Adjusted R^2	0.507	0.564	0.497	0.558	0.186	0.194	0.758	0.809
Observations	9495	9495	9495	9495	9495	9495	9495	9495

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. *** p<0.01, ** p<0.05, * p<0.10.

horses as well as variation in climatic suitability for their survival. These features identify bioclimatic conditions that would be conductive for horse availability in prehistoric times and thus the emergence of riding horses. In particular, I construct an index of historical horse availability that captures both extensive and intensive margin.

To create the historical horse index, I use two data sources. The first data is the Phylogenetic Atlas of Mammal Macroecology (PHYLACINE) database developed by Faurby et al. (2018). This atlas provides information on spatial distribution for all 5,831 known mammal species that lived since the beginning of the Late Pleistocene, which roughly corresponds to a period between 130,000 years ago until present. One of the advantages of the atlas is that it provides global maps of present natural ranges at 110 km $\times$ 110 km grid size, which are estimates of where species would live without human influences. I use a predicted distribution of horses (*Equus ferus*), which identifies the presence of horses in historic times for each cell. The map of distribution of Equus ferus is shown in Table A1.[4]

The second data comes from Naundrup and Svenning (2015), which provide a map of climatically suitable habitat for horses (*Equus ferus*) worldwide at a 10-km resolution. This suitability index is created by first obtaining climatic dimensions that predict wild-horse presence well. Then, based on these selected climatic characteristics, cells that are climatically suitable habitat for horses are identified. These process are conducted using well-established

[4]It is well known that in the Americas horses went extinct during the Pleistocene and Holocene transition (Naundrup and Svenning, 2015). This period is earlier than agricultural transition in these continents, and hence even earlier than the emergence of the state. The period of analyses is the one when the state began to form and thereafter. Therefore, I assign 0 to cells in the Americas so that the measure well captures the initial level of horse availability in the period of analysis.

models in conservation management.[5] The suitability index of horses generated this way is a result of nonlinear transformation of climatic features and wild-living horse presence. The map of distribution of Equus ferus is shown in Table A2.

Using these two datasets, I construct a historical horse index (HHI), which is calculated according to the following formula:

$$HHI_i = HorseSuit_i \times \mathbb{1}_{i \in Exist}, \tag{1}$$

where $HorseSuit_i$ is climatic suitability for horses in a cell i and $\mathbb{1}_{i \in Exist}$ is a dummy variable that takes 1 if horses exist in a cell i, and 0 otherwise. Therefore, this measure captures both an extensive margin and intensive margin of horse availability. If horses are not available, then this measure takes 0 because climatic suitability for horses does not make sense in the cell. If horses are available, then the measure takes values associated with climatic suitability. Figure 2 depicts a global map of the HHI. Importantly, both underlying data are predictions, not actual values, and hence the HHI is plausibly orthogonal to human influences. This feature alleviates the concern of reverse causality from state/battle to horse use.

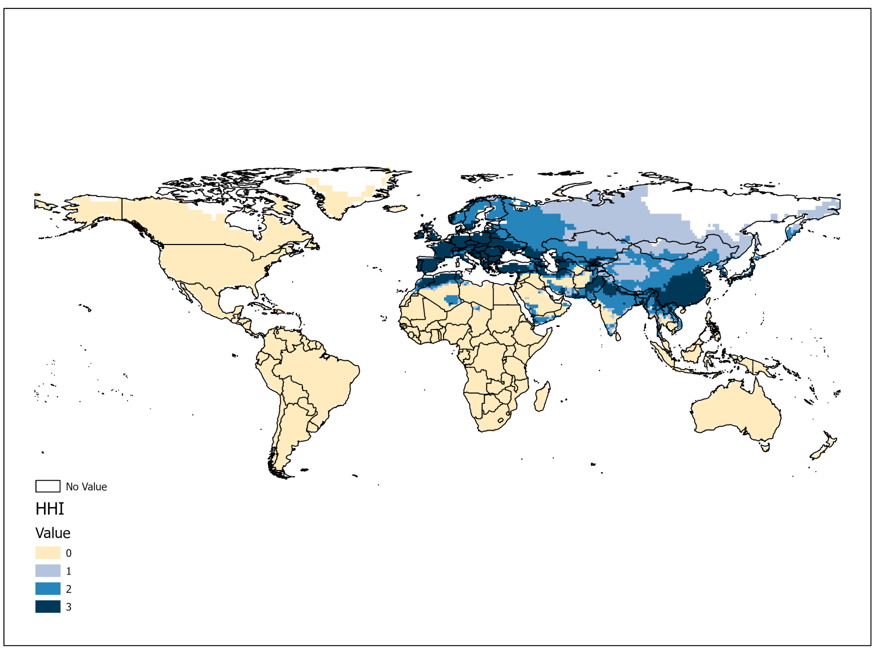

Figure 2: Historical Horse Index

In line with the proposed hypothesis, the distribution of years since cavalry emergence is positively associated with the HHI. As reported in column 1 of Table 5, unconditionally, there

[5] These models are species distribution models (SDM) and Maximum Entropy (MAXENT) modeling.

Table 5: Historical Horse Index and Time Elapsed Since Cavalry Emergence

	Time since Cavalry Emergence					
	(1)	(2)	(3)	(4)	(5)	(6)
Historical Horse Index	122.274***	128.251***	107.218***	104.833***	111.153***	77.405***
	(27.098)	(37.690)	(24.500)	(26.211)	(17.332)	(22.519)
Log Dist. to Tell el-Ajjul			-378.221***	-376.391***	-325.261***	-465.619***
			(44.967)	(46.188)	(34.109)	(108.591)
Log Dist. to Agricultural Frontier				-8.309	12.057	-6.174
				(11.051)	(8.201)	(14.093)
Continent FE		✓	✓	✓	✓	
Georaphical Controls				✓	✓	✓
Country FE						✓
Adjusted R^2	0.089	0.249	0.498	0.499	0.682	0.791
Observations	7416	7416	7416	7416	7416	7416

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distance are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. Geographical controls are latitude, longitude, latitude $\times$ longitude, average of elevation, average of land productivity, standard deviation of elevation, standard deviation of land productivity, island dummy, average of temperature, average of precipitation, standard deviation of temperature and standard deviation of precipitation. *** p<0.01, ** p<0.05, * p<0.10.

exists a highly significant positive association. Column 2 adds continent fixed effects, and it does not alter the coefficient of the HHI. Column 3 includes the distance to Tell el-Ajjul, and coefficients of the HHI and the distance to Tell el-Ajjul are both highly significant and remain expected signs. Furthermore, columns 4, 5 and 6 show that this highly significant association between time since cavalry emergence and the HHI is not weakened when the distance to the closest agricultural origin, geoclimatic features, and country fixed effects are accounted for. Reassuringly, the estimated coefficient of the HHI are very stable across specifications, remaining significant at the 1% level.

One may concern that the association between the HHI and the diffusion of cavalry may be governed by agriculture and historical trade. However, as is shown in Table 6, this is not the case. It shows that the coefficient of the HHI is robust to the inclusion of measures of agriculture and historical trade such as average caloric suitability, standard deviation of caloric suitability, time elapsed since agriculture transition[6], the distance from the closest long-distance historical trade and a dummy of transport mammals. Column 1 accounts

[6]The data on time since agricultural transition at the cell level is taken from Currie et al. (2020).

Table 6: Historical Horse Index, Agriculture and Historical Trade

	Time since Cavalry Emergence					
	(1)	(2)	(3)	(4)	(5)	(6)
Historical Horse Index	55.765***	55.078**	50.993***	53.453***	55.792***	48.283**
	(19.452)	(21.573)	(18.501)	(19.455)	(19.556)	(19.730)
Log Dist. to Tell el-Ajjul	-311.563***	-309.634***	-269.211***	-292.706***	-312.580***	-252.927***
	(80.495)	(87.404)	(68.505)	(80.928)	(80.404)	(74.670)
Land Productivity (Avg.)	9.776*	7.913	7.019	7.779	9.903*	3.946
	(5.869)	(10.036)	(6.591)	(5.831)	(5.901)	(9.098)
Land Productivity (Std)	7.808	8.493	7.706	6.748	7.821	7.507
	(5.436)	(5.676)	(5.798)	(5.541)	(5.468)	(6.204)
Caloric Suitability (Avg.)		6.631				5.635
		(39.667)				(34.593)
Caloric Suitability (Std)		-4.743				-5.045
		(7.747)				(7.804)
Time since Agricultural Transition			64.535			64.906
			(47.455)			(45.763)
Log Dist. to Trade Routes				-13.316***		-13.097***
				(5.057)		(4.702)
Transport Mammal Dummy					-3.716	-15.494
					(27.188)	(26.172)
Country FE	✓	✓	✓	✓	✓	✓
Other Controls	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.887	0.887	0.890	0.889	0.887	0.892
Observations	3686	3686	3686	3686	3686	3686

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distance are taken logarithm. Independent variables except for distances and island dummy are standardized. Other controls are latitude, longitude, latitude $\times$ longitude, average of elevation, standard deviation of elevation, island dummy, average of temperature, average of precipitation, standard deviation of temperature and standard deviation of precipitation. *** p<0.01, ** p<0.05, * p<0.10.

for basic controls and country fixed effects for comparison, including average and standard deviation of land productivity. Column 2-5 sequentially control for these additional variables. Column 6 adds all the variables together. Reassuringly, the HHI is stable across specifications and highly significant, establishing that the HHI predicts the diffusion of cavalry through a different path from agriculture and historical trade.

4.3 Empirical Model: Instrumental Variable Approach

I conduct an instrumental variable approach with the distance to Tell el-Ajjul and the HHI as instruments for years elapsed since cavalry emergence.

The second stage provides a cross-section estimate of the relationship between time since cavalry emergence and measures of state/battle:

$$y_i = \alpha_0 + \alpha_1 TSCE_i + X_i^{'}\gamma + \delta_i + \epsilon_i. \tag{2}$$

Here y_i represents a measure related to the state or historical battle of observation i, $TSCE_i$ is years elapsed since cavalry emergence of observation i, $X_i^{'}$ is a vector of geographical and climatic characteristics of observation i, δ_i is a vector of fixed effects and ϵ_i is an error term.

In the first stage, $TSCE_i$, time since cavalry emergence of observation i is instrumented by the log distance from Tell el-Ajjul, D_i, as well as the HHI of observation i, HHI_i:

$$TSCE_i = \beta_0 + \beta_1 D_i + \beta_2 HHI_i + X_i^{'}\omega + \delta_i + \eta_i. \tag{3}$$

Here $X_i^{'}$ is the same vector of geogclimatic features of observation i used in the second stage, and η_i is an error term.

5 Empirical Result: The Association between the State and Horses

5.1 Ancient Cities and Horses

In this subsection, I explore the association between ancient cities and horses. The existence of a city or a large settlement are proxies for hierarchy and socioeconomic development (Mayshar et al., 2022). I draw on the data from Degroff (2009), which depicts cities found before 400 CE as well as the data from Reba et al. (2016), which refers to the location of large settlements from classical antiquity (450 CE) and preclassical antiquity up to 500 BCE. For this analysis, I create a virtual country at the 1 × 1 cell size.

Table 7 establishes the effect of horse-riding on the presence of ancient cities. The unconditional OLS relationship between years elapsed since cavalry appearance and presence of ancient cities is positive and significant (column 1). This association is stable if continent fixed effects, geographical characteristics and climatic features are included (columns 2-4). Adding country fixed effects instead of continent fixed effects diminishes the coefficient of time since cavalry emergence, but it is still statistically highly significant (column 5). The IV estimate in column 6 suggests that time since cavalry emergence had a positive and

Table 7: The Association between Ancient Cities and Horse-Riding

	Presende of Ancient Cities					
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS
Time since Cavalry Emergence (400 CE)	0.080***	0.089***	0.076***	0.095***	0.038***	0.100**
	(0.018)	(0.020)	(0.027)	(0.029)	(0.013)	(0.040)
Log Dist. to Agricultural Frontier			-0.019***	-0.021***	-0.009	-0.006
			(0.007)	(0.006)	(0.007)	(0.007)
Latitude			-0.163***	-0.055	0.061	0.033
			(0.058)	(0.064)	(0.078)	(0.091)
Longitude			-0.233**	-0.220***	0.062	0.108
			(0.100)	(0.081)	(0.078)	(0.074)
Latitude × Longitude			0.085	0.187***	-0.039	-0.075
			(0.103)	(0.066)	(0.072)	(0.071)
Elevation (Avg.)			-0.018	-0.027*	-0.036**	-0.061**
			(0.013)	(0.015)	(0.018)	(0.026)
Land Productivity (Avg.)			0.047***	0.051***	0.026***	0.020**
			(0.017)	(0.018)	(0.008)	(0.010)
Elevation (Std)			0.045***	0.036***	0.018**	0.010
			(0.017)	(0.013)	(0.007)	(0.007)
Land Productivity (Std)			-0.005	-0.002	0.003	0.000
			(0.007)	(0.006)	(0.004)	(0.004)
Island Dummy			0.095	0.073	0.202	0.168
			(0.274)	(0.267)	(0.221)	(0.232)
Temperature (Avg.)				-0.036	-0.052	-0.157*
				(0.045)	(0.063)	(0.088)
Precipitation (Avg.)				-0.005	-0.029**	-0.030
				(0.013)	(0.014)	(0.019)
Temperature (Std)				-0.167***	-0.086**	-0.141**
				(0.042)	(0.040)	(0.066)
Precipitation (Std)				-0.022	0.016	0.027*
				(0.013)	(0.010)	(0.016)
Continent FE		✓	✓	✓		
Country FE					✓	✓
First-F						11.606
J-Test (p-value)						0.209
Adjusted R^2	0.080	0.098	0.232	0.255	0.440	
Observations	7406	7406	7406	7406	7406	7406

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1 × 1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

significant impact on the presence of ancient cities. The coefficient is larger than the OLS estimate and there are some possible explanation. Firstly, it may reflect measurement error in years elapsed since cavalry emergence. Secondly, the IV estimate may correct omitted variable bias. For example, it is possible that regions that had powerful chariotry developed the state and that they depended too much on chariotry, thus delaying a shift to cavalry. Another possibility is the characteristics of nomadic pastoralists. In general, nomadic societies depended on horses more than agricultural societies, while agricultural societies tended to develop the state earlier. In both cases, the OLS estimate is downward biased. Furthermore, the value of the first stage F-statistics and p-value of the J-test suggest the validity of the IV.

5.1.1 Robustness: Trade

Horses were also important for long-distance trade and hence it is possible that the association between the presence of ancient cities and cavalry is confounded by historical trade. To address this concern, I account for the distance from the closest trade routes.[7] Table C1 shows that the result is unaffected by the inclusion of this variable, indicating that the relationship between the ancient cities and horse-riding is not driven by historical trade. Moreover, transport mammals are another factor that may have contributed to state formation through their effect on trade (Link, 2021). Therefore, to better identify the effect of horse-riding in battle fields on the presence of ancient cities, I account for a transport mammal dummy. Table C2 shows that the result is unaffected by the inclusion of a transport mammal dummy, indicating that the relationship between the ancient cities and horse-riding is not driven by transport mammals. These results show that horse-riding in battle fields has an impact on the presence of ancient cities independent from historical long-distance trade.

5.1.2 Robustness: Agriculture

Horses were also important for agriculture and hence it is possible that the association between the ancient cities and cavalry is confounded by agricultural activities. Although I already account for measures of agricultural productivity and the distance from the nearest agricultural origin in the basic result, I further add time elapsed since agricultural transition.[8] Table C3 shows that the result is unaffected by the inclusion of this variable, indicating that the relationship between the presence of ancient cities and horse-riding in battle fields is not driven by agriculture.

[7] This data on historical long-distance trade routes is used in previous studies by Michalopoulos et al. (2018) and Link (2021)

[8] The data on time since agricultural transition at the cell level is taken from Currie et al. (2020).

5.1.3 Robustness: Other Tests

Iron had a particular importance in civilization and the spread of a bronze bit may capture that of iron. Therefore, Table C4 accounts for time since iron emergence. It shows that the result is unaffected by the inclusion of spread of iron and hence it demonstrates that the result is not driven by iron. Mayshar et al. (2022) establish that cereal crops rather than roots and tubers were conducive to development of the state. Hence, I account for a dummy of the existence of wild relatives of cereals. Table C5 shows that the result is unaffected by the inclusion of wild relatives of cereals, indicating that the relationship between the presence of ancient cities and horse-riding in battle fields is not driven by cereals.

I also use data on ancient cities from Reba et al. (2016) as dependent variables. Table C6 shows the result, using an ancient city dummy constructed from their data as a dependent variable. Although there is no association between the indicator of ancient cities and horse-riding, it may be because the number of observations in this data is small and hence there is little variation available. Therefore, instead of a dummy variable, I use the log distance to the nearest ancient city as a dependent variable in Table C7. Reassuringly, it shows the significant association between the distance to the closest ancient city and time elapsed since cavalry emerged.

5.2 State History and Horses

In this subsection, I explore the association between *State History* and horse-riding in battle fields, exploiting data from Borcan et al. (2018). The index of state history is an accumulative score that captures a country's history of hierarchy, autonomy and territory. Therefore, it allows to examine the effect of horse-riding in battle field on the state in general across countries.

Table 8 establishes the highly significant positive association between *State History* and years elapsed since cavalry emergence. The unconditional relationship between *State History* and time since cavalry is shown in column 1. The estimated coefficient is very robust to accounting for continent fixed effects (column 2), geographical features (column 3), and climatic characteristics (column 4). The IV estimate in column 5 is larger than the OLS, and it is highly significant. A large value of the first stage F-statistics and p-value of J-test are suggestive of the validity of the IV estimate, showing that time since cavalry emergence has affected the formation of state across countries.

5.2.1 Robustness: Trade

Horses were also important for long-distance trade and hence it is possible that the association between *State History* and cavalry is confounded by historical trade. To address this concern, I account for the distance from the closest trade routes. Table C8 shows that the result is unaffected by the inclusion of this variable, indicating that the relationship between *State History* and horse-riding is not driven by historical trade. Moreover, transport mammals are another factor that may have contributed to state formation through their effect on trade (Link, 2021). Therefore, to better identify the effect of horse-riding in battle fields on *State History*, I account for a transport mammal dummy. Table C9 shows that the result is unaffected by the inclusion of a transport mammal dummy, indicating that the relationship between *State History* and horse-riding is not driven by transport mammals. These results show that horse-riding in battle fields has an impact on *State History* independent from historical long-distance trade.

5.2.2 Robustness: Agriculture

Horses were also important for agriculture and hence it is possible that the association between *State History* and cavalry is confounded by agricultural activities. Although I already account for measures of agricultural productivity and the distance from the nearest agricultural origin in the basic result, I further add time elapsed since agricultural transition. Table C10 shows the result. Although OLS estimate becomes insignificant by the inclusion of the timing of agricultural transition, 2SLS estimate is highly significant, suggesting that the basic result is not driven by agriculture.

5.2.3 Robustness: Other Tests

Iron had a particular importance in civilization and the spread of a bronze bit may capture that of iron. Therefore, Table C11 accounts for time since iron emergence. It shows that the result is unaffected by the inclusion of spread of iron and hence it demonstrates that the result is not driven by iron. Mayshar et al. (2022) establish that cereal crops rather than roots and tubers were conducive to development of the state. Hence, I account for a dummy of the existence of wild relatives of cereals. Table C12 shows that the result is unaffected by the inclusion of wild relatives of cereals, indicating that the relationship between *State History* and horse-riding in battle fields is not driven by cereals.

5.3 Ethnographical Records and Horses

In this subsection, I explore the association between proxies of the state in ethnographical records and horse-riding on the battle field, drawing on the data from the *Ethnographic Atlas*. In particular, I examine the effect of time since cavalry emergence on the degree of centralization, social stratification and size of communities. The analysis establishes that horse-riding in battle fields had had a persistent impact on pre-industrial ethnographical outcomes.

Table 9 establishes that years since cavalry emergence had affected centralization (columns 1 and 2), social stratification (columns 3 and 4) and community size (columns 5 and 6). For each dependent variable, both the OLS and IV estimates are positive and significant, and the coefficient of the IV is larger than the OLS. Furthermore, the first stage F-statistics and p-value of J-test are large enough to suggest the validity of the IV estimates.

5.3.1 Robustness: Trade

Horses were also important for long-distance trade and hence it is possible that the association between the proxies of the state and cavalry is confounded by historical trade. To address this concern, I account for the distance from the closest trade routes. Table C13 shows that the IV estimates are unaffected by the inclusion of this variable, indicating that the relationship between the proxies of the state and horse-riding is not driven by historical trade. Moreover, transport mammals are another factor that may have contributed to state formation through their effect on trade (Link, 2021). Therefore, to better identify the effect of horse-riding in battle fields on the proxies of the state, I account for a transport mammal dummy. Table C14 shows that the IV estimates are unaffected by the inclusion of a transport mammal dummy, indicating that the relationship between proxies of the state and horse-riding is not driven by transport mammals. Further to address the concern about endogeneity and identification, I account for trade related motifs taken from Michalopoulos and Xue (2021).[9] Table C15 shows that the result is unaffected by the inclusion of the trade related motifs, indicating that the relationship between proxies of the state and horse-riding is not driven by trade. These results show that horse-riding in battle fields has an impact on the proxies of the state independent from historical long-distance trade.

5.3.2 Robustness: Agriculture

Horses were also important for agriculture and hence it is possible that the association between the proxies of the state and cavalry is confounded by agricultural activities. Although I already account for measures of agricultural productivity and the distance from the nearest agricultural origin in the basic result, I further add measures of dependence on agriculture and agricultural intensity. Table C16 and C17 show that the IV estimates are unaffected

[9]They show that motifs are strongly associated with the natural environments and institutional settings. For example, the presence of earthquake related motifs is significantly higher in earthquake regions. Similar associations are found for other environmentally determined variables such as the presence of storms and lightnings, or information about different modes of economic production. Therefore, motifs can be used as proxies of the experiences of ethnic groups.

by the inclusion of dependence on agriculture and agricultural intensity, respectively, indicating that the relationship between proxies of the state and horse-riding is not driven by agriculture.

5.3.3 Robustness: Other Tests

Iron had a particular importance in civilization and the spread of a bronze bit may capture that of iron. Therefore, Table C18 accounts for time since iron emergence. By the inclusion of iron, the association between the degree of centralization and cavalry diffusion becomes insignificant while hierarchy and community size remain significant relationship with cavalry diffusion. Mayshar et al. (2022) establish that cereal crops rather than roots and tubers were conducive to development of the state. Hence, I account for a measure of cereal advantage over roots and tubers. Table C19 shows that the IV estimates are unaffected by the inclusion of cereal advantage over roots and tubers, indicating that the relationship between proxies of the state and horse-riding is not driven by cereals.

5.4 Panel Data Based on the Columbian Exchange: State Index and Horses

The natural experiment associated with the Columbian Exchange provides exogenous change in the HHI in the Americas. This index is constructed such that if a horse does not exist in a cell, then that cell takes a value 0 because climatic suitability for horses does not make sense in the cell. Therefore, all the cells in the Americas where horses became extinct in prehistoric times have values 0 before the Columbian Exchange. European immigrants brought horses to the Americas and they spread across the continents (Chamberlin, 2010). Therefore, after the Columbian Exchange, cells in the Americas take values identical with the suitability index. Because this index measures how suitable an area is for horse survival, the change in the HHI captures the change in the availability of horses.

Database by Borcan et al. (2018) provide panel data on the index of the state every half century since 3500 BCE to 2000 CE. Hence, I restrict samples to the Americas and conduct panel analyses, using exogenous variation in the HHI resulting from the Columbian Exchange. Accounting for country and time fixed effects as well as utilizing the exogenous variation in the HHI avoid a concern about omitted variable bias.

The benchmark sample used in this analysis comprises 27 countries in the Americas. Following Mayshar et al. (2022), I use the year since 1000 CE, with observations available every half century. The association between *State Index* and the HHI is estimated according to the following specification:

$$State_{it} = \alpha_0 + \alpha_1 HHI_{i,t-1} + \gamma_i + \delta_t + \epsilon_{it}, \tag{4}$$

where $State_{it}$ is *State Index* of country i in year t; HHI_{it} takes 0 if $t-1 < 1500CE$, and takes a value associated with climatic suitability for horses if $t-1 \geq 1500CE$; γ_i is country fixed effects; and δ_t is year fixed effects.

Table 10 demonstrates that change in the HHI has a positive impact on *State Index*. Column 1 shows the unconditional positive significant association. This association is unaffected by adding change in agricultural suitability (column 2) and country fixed effects (column 3). Adding time fixed effects in column 4 decreases the magnitude of the coefficient and increases standard error, resulting in an insignificant estimate. This is not surprising because most countries in the America had not developed the state before 1500 CE. After the contact with Europe, they started to develop complicated societies. Indeed, the average of *State Index* is 4.24 before 1500 CE and 29.5 after 1500 CE. Moreover, the score linearly increases over time. Therefore, time fixed effects absorb much variation in *State Index*. The coefficient of the HHI becomes significant again when including geographical controls (columns 4-7). It is robust to further controls of climatic features in column 8. Accounting for time-variant geoclimatic effects on *StateIndex* and the HHI as well as country and time fixed effects, the availability of horses has affected on state formation of countries in the Americas.

The identification assumption is that until 1500 CE, *State Index* did not have systematically different trends across the set of countries that experienced different shocks in the HHI. Here, I show the existence of parallel trends before 1500 CE. In particular, I regress *State Index* on the change in the HHI generated by the Columbian Exchange interacted with a linear trend before 1500 CE. Table 11 shows the result of this exercise. The coefficient on the interaction term indicates whether *State Index* in countries that experienced a larger change in the HHI were on a different linear trend before the Columbian Exchange. Reassuringly, the coefficient is always insignificant. It suggests that before 1500 CE, countries in the Americas had not experienced systematically different trends relating *State Index* and the HHI.

5.5 Difference-In-Differences Using Radiocarbon-Dated Prehistoric and Posthistoric Archaeological Sites

In this subsection, I examine the association between prehistoric archaeological sites and horses, using the data and methodology developed by Mayshar et al. (2022). They georeference archaeological sites reported by Whitehouse and Whitehouse (1975) and classify them according to whether or not they predate the Neolithic transition. Based on this data, I create 1×1 virtual countries (cells) in which every cell includes information on the number of pre- and post-Neolithic sites. Unlike the previous exercise that uses cross-sectional variation in ancient cities, I run the following difference-in-differences regressions:

$$y_{it} = \alpha_0 D_i \times P_t + \alpha_1 HHI_i \times P_t + \eta_i + \eta_t + \epsilon_{it}, \tag{5}$$

where i indicates cell i, t indicates whether the site predates the Neolithic transition, η_i and η_t are cell and period fixed effects, respectively, P_t is a dummy variable that identifies archaeological sites dating to after the Neolithic transition, D_i is the log distance from Tell el-Ajjul of cell i, HHI_i is the historical horse index of cell i, and ϵ_{it} is an error term.

Column 1 of Table 12 shows that the distance from Tell el-Ajjul is associated with a decrease in the probability of finding post-Neolithic sites, in comparison to finding pre-Neolithic sites. It also shows that the HHI is related to an increase in the probability of finding post-Neolithic sites relative to pre-Neolithic sites. This result confirms that the Neolithic transition led to more visible traces of human societies in areas that were closer to Tell el-Ajjul and/or areas where the HHI was higher. To better identify the effect of horses in battle fields on the presence of post-Neolithic sites, column 2 adds the distance from the closest agricultural origin and the distance from the nearest long-distance trade route as well as the distance from the lower Volga-Don region, where the modern horses were first domesticated.[10] The inclusion of these distances would remove the associations with agriculture and trade from the estimated coefficients of the distance from Tell el-Ajjul and the HHI. The estimates of the distance to Tell el-Ajjul and the HHI are unaffected by the inclusion of these variables. The results reported in the rest of Table 12 confirm the results in columns 1 and 2, with different dependent variables: the number of archaeological sites (columns 3 and 4), the presence of a prehistoric settlement (columns 5 and 6), and the number of prehistorical settlements (columns 7 and 8).

The same qualitative results are obtained when accounting for a dummy variable that indicates wild relatives of cereals (Table C20), hence indicating that the basic result is not driven by the availability of cereals. Table C21 adds a transport mammal dummy, and this inclusion does not alter the basic result. Therefore, the availability of transport mammals does not drive the association between archaeological sites and horses.

[10] As is described in the empirical methodology section, the purpose of horse domestication was to eat their meat and did not lead to horse-riding in battle fields. On the other hand, the invention of a metal bit enabled horse-riding on the battle field. Therefore, accounting for the distance from the lower-Volga Don regions would separate the effects of horse domestication and horse-back riding.

6 Empirical Result: The Association between Historical Battles and Horses

In this section, I explore the association between historical battles and horses, drawing on the *World Historical Battles Database* (WHBD) compiled by Kitamura (2021). The *WHBD* covers older historical battles in the entire world than other previous databases on conflicts. Furthermore, it provides rich information on each battle such as geolocation, battle result, etc. These characteristics of the *WHBD* enable to examine the relationship between historical battles and horses going back to older time at granular scales than previous studies. Therefore, I create 1×1 virtual countries (cells) to take advantage of the dataset. First, I examine cross-sectional association between historical battles and horse-riding on the battle field. Then, I use observations in the Americas and exploit the exogenous change in the HHI resulting from the Columbian Exchange. This natural experiment further allows to rule out unobservable characteristics that are systematically correlated with historical battles and the HHI.

6.1 Historical Battles and Horses: Cross-Virtual Country

To examine the association between historical battles and horse-riding in battle fields across virtual countries, I run the 2SLS regression described by (2) and (3). Table 13 establishes the negative significant association between the log distance to the closest battle before 900 CE and time elapsed since cavalry emergence (Figure A3 shows the distribution of battles before 900 CE).[11][12] As reported in column 1, there is a highly significant negative unconditional association between time since cavalry emergence and the distance to the nearest battle. This association is stable when accounting for continent fixed effects (column 2), geographical characteristics (column 3), climatic features (column 4), and country fixed effects instead of continent fixed effects (column 5). As is shown in column 6, the IV estimate suggests that years elapsed since cavalry emerged had a positive impact on historical battles, accounting for a number of geoclimatic characteristics and country fixed effects.

[11] I choose 900 CE because of the possibility that gunpowder had affected the use of cavalry. Gunpower was first invented in China and used for warfare around 904 CE and then it spread across Eurasia by the end of the 13th century (Buchanan, 2006; Andrade, 2016). Therefore, it would be more difficult to examine the effect of cavalry emergence on battles if battles after 900 CE were added.

[12] If a dummy variable of the occurrence of battles instead of the distance is used, the estimate becomes insignificant once controlling for country fixed effects. However, this is because there is very limited variation in the dummy variable. Indeed, 95.4% of the observations take 0 and thus, accounting for country fixed effects absorbs almost all the variation available. Consistent with this explanation, there is significant association between historical battles and horse-riding on the battle field when using the distance from the closest battle or controlling for continent fixed effects instead of country fixed effects.

6.1.1 Robustness

In this subsection, I conduct a series of robustness tests. Mayshar et al. (2022) establish that cereal crops rather than roots and tubers were conducive to development of the state. Given the possible correlation between the state and conflicts, I account for a dummy of the existence of wild relatives of cereals. Table C22 shows that the result is unaffected by the inclusion of wild relatives of cereals, indicating that the relationship between historical battles and horse-riding is not driven by cereals.

Transport mammals are another factor that may have contributed to state formation through their effect on trade (Link, 2021). Given the possible correlation between the state and trade, I account for a transport mammal dummy. Table C23 shows that the result is unaffected by the inclusion of a transport mammal dummy, indicating that the relationship between historical battles and horse-riding is not driven by transport mammals.

Regarding the possibility that trade confounds conflicts and horse-riding, I further account for historical trade routes. Table C24 shows that the result is unaffected by the inclusion of the distance from historical trade routes, when further accounting for the distance to the lower Volga-Don region.

Horses were also important for agriculture and hence it is possible that the association between conflicts and the existence of cavalry is confounded by agriculture. Although I already account for measures of agricultural suitability and the distance from the nearest agricultural origin in the basic result, I further add time elapsed since agricultural transition. Table C25 shows that the result is unaffected by the inclusion of time elapsed since agricultural transition, when further accounting for the distance to the lower Volga-Don region.

Iron had a particular importance in battles and the spread of a bronze bit may capture that of iron. Therefore, Table C26 accounts for time since iron emergence. It shows that the result is unaffected by the inclusion of spread of iron and hence it demonstrates that the result is not driven by iron.

6.2 Historical Battles and Horses: Natural Experiment Associated with the Columbian Exchange

To examine the association between historical battles and horses, I exploit the exogenous variation in the HHI caused by the Columbian Exchange. The empirical specification is the same as (4) except for dependent variables and the unit of analysis. Dependent variables in this subsection are measures about historical battles such as an indicator of battles and the number of battles. The unit of analysis is a 1×1 virtual country (cell) and it is much granular than a country that is used in the previous section. The exogenous change in the HHI as well as the panel nature of the data alleviate the concern about omitted variable bias.

Table 14 establishes a highly significant positive impact of the HHI on the presence of historical battles. As is reported in column 1, the unconditional association between historical battles and the HHI is highly significant and positive. The inclusion of land productivity (column 2) and cell fixed effects (column 3) does not affect the coefficient of the HHI. Column 4 accounts for time fixed effects and the magnitude of the coefficient of the HHI increases, remaining highly significant. The rest of Table 14 sequentially controls for time-variant geographical and climatic characteristics. Reassuringly, the coefficient of the HHI is unaffected by these inclusion. The estimate is highly significant in all the specification, indicating the positive effect of horse availability on historical battles.

The identification assumption is that until 1500 CE, battles did not have systematically different trends across the set of cells that experienced different shocks in the HHI. Here, I show the existence of parallel trends before 1500 CE. In particular, I regress the presence of battles on the change in the HHI generated by the Columbian Exchange interacted with a linear trend before 1500 CE. Table 15 shows the result of this exercise. The coefficient on the interaction term indicates whether battles in cells that experienced a larger change in the HHI were on a different linear trend before the Columbian Exchange. Reassuringly, the coefficient is small and insignificant. It suggests that before 1500 CE, cells in the Americas had not experienced systematically different trends relating conflicts and the HHI.

The same qualitative results are obtained when using the number of battles instead of the dummy variable. Higher change in the HHI resulting from the Columbian Exchange is positively associated with more battles as is shown in Table C27. Moreover, Table C28 shows that the coefficient on the interaction term between the HHI and time trend before 1500 CE is not significantly related to the number of battles. Thus, it indicates the parallel trend before the Columbian Exchange holds.

To sum up, the empirical analyses using the natural experiment associated with the Columbian Exchange in this subsection give a strong evidence of the positive impact of horses on historical battles.

7 Concluding Remarks

This research explores the effect of horses on state formation and historical battle. Exploiting exogenous geographical variation in the adoption of horse-riding, the study establishes that consistent with rich historical narratives, the years since cavalry emergence had a positive impact on state/battle. The analysis provide repeated evidence, utilizing multiple data sets and conducting several empirical methodology such as an instrumental variable approach, the natural experiment associated with the Columbian Exchange, and difference-in-differences estimates.

Although this study demonstrates the important role of horses in the emergence of the state and the occurrence of battles, there are other factors that contributed to state formation. In the context of military weapons, gun powder is another critical factor that shaped state history. Therefore, one extension of this research is to explore the association between horses and gun powder (substitute or complement?) as well as their joint impact on state formation. History of weapon would have persistent impacts on the wealth of nations.

Table 8: The Association between State History and Horse-Riding

	State History				
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) 2SLS
Time since Cavalry Emergence	372.528***	323.593***	297.471***	321.349***	564.333***
	(43.335)	(58.618)	(69.568)	(103.683)	(122.574)
Log Dist. to Agricultural Frontier			-40.367	-35.044	-46.998
			(24.984)	(30.980)	(29.224)
Latitude			540.458***	578.354***	377.584**
			(177.321)	(188.973)	(187.510)
Longitude			878.611***	745.205**	802.823***
			(295.371)	(292.538)	(264.258)
Latitude × Longitude			-525.876***	-262.735	-214.126
			(181.122)	(200.940)	(198.141)
Elevation (Avg.)			-190.523**	-166.434	-316.780**
			(80.797)	(156.128)	(158.281)
Land Productivity (Avg.)			7.149	-21.411	-86.953
			(67.547)	(76.345)	(71.762)
Elevation (Std)			220.293***	183.054**	185.208**
			(74.320)	(88.849)	(84.951)
Land Productivity (Std)			7.864	27.272	-18.393
			(79.483)	(89.329)	(77.502)
Island Dummy			32.917	70.517	269.086
			(275.952)	(268.444)	(252.181)
Temperature (Avg.)				15.930	-342.543
				(257.157)	(266.388)
Precipitation (Avg.)				-209.284*	-337.600***
				(121.338)	(126.525)
Temperature (Std)				-306.408*	-542.750***
				(174.136)	(190.667)
Precipitation (Std)				48.416	203.535*
				(110.777)	(110.156)
Continent FE		✓	✓	✓	✓
First Stage F-Statistics					41.228
J-Test (p-value)					0.684
Adjusted R^2	0.403	0.430	0.535	0.556	
Observations	94	94	94	94	94

Note: OLS regressions with robust standard errors. The unit of analysis is a country, restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table 9: The Association between the State Proxies and Horse-Riding

	Centralization		Social Stratification		Community Size	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
Time since Cavalry Emergence	0.147*	0.347**	0.133**	0.239***	0.516*	0.956***
	(0.078)	(0.132)	(0.064)	(0.039)	(0.272)	(0.337)
Log Dist. to Agricultural Frontier	-0.012	-0.027	-0.021	-0.029	-0.316**	-0.324**
	(0.067)	(0.064)	(0.034)	(0.032)	(0.147)	(0.151)
Latitude	0.035	-0.175	-0.379	-0.482	1.141*	0.790
	(0.463)	(0.499)	(0.298)	(0.334)	(0.642)	(0.463)
Longitude	-0.295	-0.310	-0.428	-0.421	-0.632	-0.528
	(0.563)	(0.524)	(0.272)	(0.258)	(0.791)	(0.762)
Latitude × Longitude	-0.957***	-0.835**	-0.286	-0.232	-0.409	-0.201
	(0.327)	(0.325)	(0.207)	(0.196)	(0.684)	(0.590)
Elevation (Avg.)	-0.067	-0.151	-0.046	-0.092	0.203	0.036
	(0.089)	(0.104)	(0.058)	(0.063)	(0.299)	(0.259)
Land Productivity (Avg.)	0.184**	0.185***	0.044	0.045	0.626***	0.616***
	(0.071)	(0.062)	(0.041)	(0.037)	(0.170)	(0.158)
Elevation (Std)	-0.126*	-0.155**	0.009	-0.007	-0.338**	-0.408***
	(0.064)	(0.070)	(0.029)	(0.028)	(0.137)	(0.145)
Land Productivity (Std)	-0.123**	-0.140**	-0.058	-0.068	0.175	0.140
	(0.056)	(0.058)	(0.087)	(0.086)	(0.133)	(0.140)
Island Dummy	-0.198	-0.227	-1.157***	-1.177***	0.000	0.000
	(0.209)	(0.168)	(0.204)	(0.169)	(.)	(.)
Temperature (Avg.)	-0.477*	-0.703***	-0.316**	-0.439**	0.529	0.036
	(0.277)	(0.246)	(0.122)	(0.159)	(0.703)	(0.526)
Precipitation (Avg.)	-0.152*	-0.157**	-0.216***	-0.218**	0.168	0.202
	(0.089)	(0.071)	(0.077)	(0.085)	(0.288)	(0.314)
Temperature (Std)	0.101	-0.076	0.064	-0.027	-0.417	-0.825*
	(0.271)	(0.210)	(0.142)	(0.137)	(0.400)	(0.434)
Precipitation (Std)	0.008	0.044	0.179***	0.200***	-0.026	0.054
	(0.074)	(0.059)	(0.042)	(0.039)	(0.195)	(0.161)
Continent FE	✓	✓	✓	✓	✓	✓
First Stage F-Statistics		20.129		20.332		18.919
J-Test (p-value)		0.153		0.169		0.451
Adjusted R^2	0.206		0.122		0.274	
Observations	595	595	550	550	278	278

Note: OLS regressions with robust standard errors clustered at the language family level. The unit of analysis is an ethnic group, restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table 10: State Index and the Historical Horse Index: The Natural Experiment Associated with the Columbian Exchange

	State Index							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Historical Horse Index	6.416***	6.601***	5.686***	0.001	1.015	3.358**	3.401**	4.301***
	(0.331)	(0.334)	(0.786)	(1.523)	(1.155)	(1.481)	(1.412)	(1.365)
Land Suitability		-2.048***	15.316*	1.462	3.792	15.003***	13.976***	10.842
		(0.684)	(7.755)	(4.081)	(7.830)	(4.238)	(4.281)	(7.314)
Country FE			✓	✓	✓	✓	✓	✓
Time FE				✓	✓	✓	✓	✓
Time × Geolocation					✓	✓	✓	✓
Time × Elevation						✓	✓	✓
Time × Island							✓	✓
Time × Climate								✓
Adjusted R^2	0.454	0.463	0.604	0.824	0.831	0.882	0.882	0.888
Observations	452	452	452	452	452	452	452	452

Note: OLS regressions with robust standard errors clustered at the country level. The unit of observation is the territory delimited by modern country borders every 50 years, spanning since 1000 CE to 2000 CE. Only countries in the Americas are used. Geolocation are latitude, longitude and latitude × longitude. Elevation are average of elevation and standard deviation of elevation. Island is a dummy variable. Climate are average temperature, average precipitation, standard deviation of temperature and standard deviation of precipitation. All the variables except for the dependent variable are standardized. *** p<0.01, ** p<0.05, * p<0.10.

Table 11: State Index and the Historical Horse Index: Pre-Trend

	State Index							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Change HHI × Time	4.546	4.584	10.339	33.458	18.046	9.673	9.439	22.152
	(8.595)	(8.649)	(8.766)	(23.360)	(14.281)	(21.173)	(20.920)	(20.111)
Change Land Suit. × Time		0.090	3.220	8.613	10.039	6.265	6.941	18.903
		(2.469)	(4.842)	(5.356)	(7.951)	(9.211)	(10.013)	(15.851)
Country FE			✓	✓	✓	✓	✓	✓
Time FE				✓	✓	✓	✓	✓
Geolocation × Year					✓	✓	✓	✓
Elevation × Year						✓	✓	✓
Island × Year							✓	✓
Climate × Year								✓
Adjusted R^2	0.002	-0.002	0.880	0.880	0.886	0.887	0.887	0.893
Observations	220	220	220	220	220	220	220	220

Note: OLS regressions with robust standard errors clustered at the country level. The unit of observation is the territory delimited by modern country borders every 50 years, spanning since 1000 CE to 1500 CE. Only countries in the Americas are used. Geolocation are latitude, longitude and latitude × longitude. Elevation are average of elevation and standard deviation of elevation. Island is a dummy variable. Climate are average temperature, average precipitation, standard deviation of temperature and standard deviation of precipitation. All the variables except for the dependent variable are standardized. *** p<0.01, ** p<0.05, * p<0.10.

Table 12: Difference-In-Difference Analysis: Archaeological Sites and Horses

	Ancient Site (Dummy)		Ancient Site (Count)		Ancient Settlement (Dummy)		Ancient Settlement (Count)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After Agr. × Log Dist. to Tell el-Ajjul	-0.039***	-0.107***	-0.154***	-0.465***	-0.035**	-0.100***	-0.119**	-0.345***
	(0.014)	(0.030)	(0.054)	(0.117)	(0.016)	(0.033)	(0.054)	(0.109)
After Agr. × Historical Horse Index	0.036***	0.062***	0.126**	0.253***	0.034**	0.058***	0.100**	0.192***
	(0.013)	(0.020)	(0.051)	(0.082)	(0.013)	(0.019)	(0.043)	(0.067)
After Agr. × Log Dist. to the Neolithic Origin		-0.009		-0.009		-0.009		-0.017
		(0.007)		(0.022)		(0.008)		(0.021)
After Agr. × Log Dist. to Trade Routes		-0.003		0.013		-0.007		0.002
		(0.008)		(0.025)		(0.009)		(0.024)
After Agr. × Log Dist. to Volga-Don		0.123***		0.489***		0.128***		0.378***
		(0.028)		(0.117)		(0.027)		(0.100)
Cell FE	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.240	0.257	0.222	0.246	0.194	0.218	0.166	0.188
Observations	25552	25552	25552	25552	25552	25552	25552	25552

Note: The table reports difference-in-differences OLS regression. The unit of observation is the 1×1 virtual country (cell) before or after the Neolithic transition. Robust standard errors are clustered at the country level. All the independent variables except for the distances are standardized. *** p<0.01, ** p<0.05, * p<0.10.

Table 13: The Association between Historical Battles and Horse-Riding

	Log Dist. to the Closest Battle					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	OLS	OLS	2SLS
Time since Cavalry Emergence (900 CE)	-0.999***	-0.978***	-0.729***	-0.849***	-0.594***	-0.798**
	(0.124)	(0.130)	(0.131)	(0.167)	(0.116)	(0.372)
Log Dist. to Agricultural Frontier			0.200**	0.237***	0.196**	0.188**
			(0.093)	(0.082)	(0.091)	(0.091)
Latitude			-0.046	-0.055	-1.450	-1.379
			(0.480)	(0.441)	(1.112)	(1.222)
Longitude			1.191*	0.630	-2.461*	-2.567*
			(0.687)	(0.652)	(1.307)	(1.312)
Latitude × Longitude			0.101	0.175	2.464**	2.544**
			(0.594)	(0.548)	(1.231)	(1.271)
Elevation (Avg.)			0.067	0.374**	0.339**	0.433
			(0.046)	(0.148)	(0.166)	(0.266)
Land Productivity (Avg.)			-0.263	-0.271	-0.122	-0.094
			(0.173)	(0.174)	(0.147)	(0.166)
Elevation (Std)			-0.295***	-0.281***	-0.155***	-0.135***
			(0.089)	(0.074)	(0.036)	(0.051)
Land Productivity (Std)			0.087	0.063	0.015	0.023
			(0.059)	(0.053)	(0.029)	(0.026)
Island Dummy			0.879*	0.785	-0.088	-0.032
			(0.458)	(0.565)	(0.291)	(0.353)
Temperature (Avg.)				1.116**	0.624	0.987
				(0.485)	(0.587)	(0.958)
Precipitation (Avg.)				0.475***	0.395	0.387
				(0.151)	(0.250)	(0.263)
Temperature (Std)				1.308***	0.355	0.547
				(0.455)	(0.378)	(0.557)
Precipitation (Std)				0.028	-0.128	-0.148
				(0.133)	(0.132)	(0.137)
Continent FE		✓	✓	✓		
Country FE					✓	✓
First-F						13.672
J-Test (p-value)						0.174
Adjusted R^2	0.295	0.365	0.490	0.523	0.730	
Observations	7406	7406	7406	7406	7406	7406

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table 14: The Natural Experiment Associated with the Columbian Exchange and the Association between Historical Battles and the Historical Horse Index

	Presence of Historical Battles							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Historical Horse Index	0.0054***	0.0054***	0.0049***	0.0089***	0.0085***	0.0086***	0.0086***	0.0085***
	(0.0015)	(0.0015)	(0.0014)	(0.0025)	(0.0026)	(0.0027)	(0.0027)	(0.0024)
Land Productivity		0.0004	0.0000	0.0001*	0.0002*	0.0002*	0.0002*	0.0001*
		(0.0008)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Cell FE			✓	✓	✓	✓	✓	✓
Time FE				✓	✓	✓	✓	✓
Time × Geolocation					✓	✓	✓	✓
Time × Elevation						✓	✓	✓
Time × Island							✓	✓
Time × Climate								✓
Adjusted R^2	0.021	0.021	0.022	0.056	0.057	0.057	0.057	0.058
Observations	96529	96529	96529	96529	96529	96529	96529	96529

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country in the Americas every 50 years, spanning since 1000 CE to 2000 CE. Geolocation are latitude, longitude and latitude × longitude. Elevation are average of elevation and standard deviation of elevation. Island is a dummy variable. Climate are average temperature, average precipitation, standard deviation of temperature and standard deviation of precipitation. All the variables except for the dependent variable are standardized. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 15: Pre-Trend in the Association between Historical Battles and the Historical Horse Index

	Presence of Historical Battles							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Change HHI × Year	0.0003	0.0003	0.0009	0.0024	0.0004	0.0001	0.0001	0.0004
	(0.0003)	(0.0002)	(0.0009)	(0.0022)	(0.0023)	(0.0019)	(0.0019)	(0.0029)
Change Land Suit. × Year		0.0001	0.0008	0.0009	0.0007	0.0007	0.0007	0.0006
		(0.0001)	(0.0010)	(0.0012)	(0.0012)	(0.0013)	(0.0013)	(0.0013)
Cell FE			✓	✓	✓	✓	✓	✓
Time FE				✓	✓	✓	✓	✓
Time × Geolocation					✓	✓	✓	✓
Time × Elevation						✓	✓	✓
Time × Island							✓	✓
Time × Climate								✓
Observations	39600	39600	39600	39600	39600	39600	39600	39600

Note: OLS regressions with robust standard errors clustered at the country level. The unit of observation is 1 × 1 virtual country every 50 years, spanning since 1000 CE to 1500 CE. Only virtual countries in the Americas are used. Geolocation are latitude, longitude and latitude × longitude. Elevation are average of elevation and standard deviation of elevation. Island is a dummy variable. Climate are average temperature, average precipitation, standard deviation of temperature and standard deviation of precipitation. All the variables except for the dependent variable are standardized. *** p<0.01, ** p<0.05, * p<0.10.

References

Andrade, Tonio, "The Gunpowder Age," in "The Gunpowder Age," Princeton University Press, 2016.

Arbatlı, Cemal Eren, Quamrul H. Ashraf, Oded Galor, and Marc Klemp, "Diversity and conflict," *Econometrica*, 2020, *88* (2), 727–797. Publisher: Wiley Online Library.

Binford, Lewis R., *Constructing frames of reference: an analytical method for archaeological theory building using ethnographic and environmental data sets*, University of California Press, 2019.

Bokovenko, Nikolay A., "The origins of horse riding and the development of ancient Central Asian nomadic riding harnesses," *Kurgans, ritual sites, and settlements: Eurasian Bronze and Iron Age*, 2000, *870*, 304–310. Publisher: BAR International Series.

Borcan, Oana, Ola Olsson, and Louis Putterman, "State history and economic development: evidence from six millennia," *Journal of Economic Growth*, 2018, *23* (1), 1–40. Publisher: Springer.

Buchanan, Brenda J., *Gunpowder, explosives and the state: a technological history*, Ashgate Publishing, Ltd., 2006.

Chamberlin, J. Edward, *Horse: How the horse has shaped civilizations*, Vintage Canada, 2010.

Collier, Paul and Anke Hoeffler, "On economic causes of civil war," *Oxford economic papers*, 1998, *50* (4), 563–573. Publisher: Oxford University Press.

_ **and** _ , "Greed and grievance in civil war," *Oxford economic papers*, 2004, *56* (4), 563–595. Publisher: Oxford University Press.

Currie, Thomas E., Peter Turchin, Edward Turner, and Sergey Gavrilets, "Duration of agriculture and distance from the steppe predict the evolution of large-scale human societies in Afro-Eurasia," *Humanities and Social Sciences Communications*, 2020, *7* (1), 1–8. Publisher: Palgrave.

Degroff, Daniel, "Ancient Cities DB," 2009.

Diamond, Jared, *Guns, germs, and steel: the fates of human societies*, New York: Norton, 1997.

Drews, Robert, *Early riders: The beginnings of mounted warfare in Asia and Europe*, Routledge, 2004.

Ebrey, P., A. Walthall, and J. Palais, "Pre-modern East Asia: To 1800: A Cultural," *Social, and Political History*, 2006.

Ellis, John, *Cavalry: History of mounted warfare* number 32, Pen and Sword, 2004.

Esteban, Joan, Laura Mayoral, and Debraj Ray, "Ethnicity and conflict: An empirical study," *American Economic Review*, 2012, *102* (4), 1310–42.

Faurby, Søren, Matt Davis, Rasmus Ø Pedersen, Simon D. Schowanek, Alexandre Antonelli, and Jens-Christian Svenning, "PHYLACINE 1.2: The phylogenetic atlas of mammal macroecology," *Ecology*, 2018, *99* (11), 2626. Publisher: John Wiley & Sons, Ltd.

Fearon, James D. and David D. Laitin, "Ethnicity, insurgency, and civil war," *American political science review*, 2003, *97* (1), 75–90. Publisher: Cambridge University Press.

Fenske, James, "Ecology, trade, and states in pre-colonial Africa," *Journal of the European Economic Association*, 2014, *12* (3), 612–640. Publisher: Oxford University Press.

Galor, Oded and Ömer Özak, "The agricultural origins of time preference," *American Economic Review*, 2016, *106* (10), 3064–3103.

Giuliano, Paola and Nathan Nunn, "Ancestral characteristics of modern populations," *Economic History of Developing Regions*, 2018, *33* (1), 1–17. Publisher: Taylor & Francis.

Goldewijk, Kees Klein, Arthur Beusen, Jonathan Doelman, and Elke Stehfest, "Anthropogenic land use estimates for the Holocene–HYDE 3.2," *Earth System Science Data*, 2017, *9* (2), 927–953. Publisher: Copernicus GmbH.

Hibbs, Douglas A. and Ola Olsson, "Geography, biogeography, and why some countries are rich and others are poor," *Proceedings of the national Academy of sciences*, 2004, *101* (10), 3715–3720. Publisher: National Acad Sciences.

Hyland, Ann, *The medieval warhorse: From Byzantium to the Crusades*, Sutton Pub Limited, 1996.

Keegan, John, *A history of warfare*, Random House, 2011.

Kitamura, Shuhei, "World Historical Battles Database, Version 1.1," *Unpublished manuscript*, 2021.

Law, Robin, *The horse in West African history: the role of the horse in the societies of pre-colonial West Africa*, Routledge, 2018.

Librado, Pablo, Naveed Khan, Antoine Fages, Mariya A. Kusliy, Tomasz Suchan, Laure Tonasso-Calvière, Stéphanie Schiavinato, Duha Alioglu, Aurore Fromentier, and Aude Perdereau, "The origins and spread of domestic horses from the Western Eurasian steppes," *Nature*, 2021, *598* (7882), 634–640. Publisher: Nature Publishing Group.

Link, Andreas, "Beasts of Burden, Trade, and Hierarchy: The Long Shadow of Domestication," *Mimemo*, 2021.

Mayshar, Joram, Omer Moav, and Luigi Pascali, "The Origin of the State: Land Productivity or Appropriability?," *Journal of Political Economy*, 2022, *130* (4), 1091–1144. Publisher: The University of Chicago Press Chicago, IL.

Michalopoulos, Stelios, Alireza Naghavi, and Giovanni Prarolo, "Trade and Geography in the Spread of Islam," *The Economic Journal*, 2018, *128* (616), 3210–3241. Publisher: Wiley Online Library.

_ and Elias Papaioannou, "Pre-colonial ethnic institutions and contemporary African development," *Econometrica*, 2013, *81* (1), 113–152. Publisher: Wiley Online Library.

_ and Melanie Meng Xue, "Folklore," *The Quarterly Journal of Economics*, 2021, *136* (4), 1993–2046. Publisher: Oxford University Press.

Montalvo, José G. and Marta Reynal-Querol, "Ethnic polarization, potential conflict, and civil wars," *American economic review*, 2005, *95* (3), 796–816.

Murdock, George Peter, "Ethnographic atlas: a summary," *Ethnology*, 1967, *6* (2), 109–236. Publisher: JSTOR.

Naundrup, Pernille Johansen and Jens-Christian Svenning, "A geographic assessment of the global scope for rewilding with wild-living horses (Equus ferus)," *PloS one*, 2015, *10* (7), e0132359. Publisher: Public Library of Science San Francisco, CA USA.

Olsson, Ola and Douglas A. Hibbs, "Biogeography and long-run economic development," *European Economic Review*, 2005, *49* (4), 909–938. Publisher: Elsevier.

Purugganan, Michael D. and Dorian Q. Fuller, "The nature of selection during plant domestication," *Nature*, 2009, *457* (7231), 843–848. Publisher: Nature Publishing Group.

Putterman, Louis and David N. Weil, "Post-1500 population flows and the long-run determinants of economic growth and inequality," *The Quarterly journal of economics*, 2010, *125* (4), 1627–1682. Publisher: MIT Press.

Ramankutty, Navin, Jonathan A. Foley, John Norman, and Kevin McSweeney, "The global distribution of cultivable lands: current patterns and sensitivity to possible climate change," *Global Ecology and biogeography*, 2002, *11* (5), 377–392. Publisher: Wiley Online Library.

Reba, Meredith, Femke Reitsma, and Karen C. Seto, "Spatializing 6,000 years of global urbanization from 3700 BC to AD 2000," *Scientific data*, 2016, *3* (1), 1–16. Publisher: Nature Publishing Group.

Turchin, Peter, Daniel Hoyer, Andrey Korotayev, Nikolay Kradin, Sergey Nefedov, Gary Feinman, Jill Levine, Jenny Reddish, Enrico Cioni, and Chelsea Thorpe, "Rise of the war machines: Charting the evolution of military technologies from the Neolithic to the Industrial Revolution," *PloS one*, 2021, *16* (10), e0258161. Publisher: Public Library of Science San Francisco, CA USA.

_ **, Thomas E. Currie, and Edward AL Turner**, "Mapping the spread of mounted warfare," *Cliodynamics*, 2016, *7* (2).

Whitehouse, David and Ruth D. Whitehouse, "Archaeological atlas of the world," 1975. Publisher: Thames and Hudson.

Appendix A. Figures

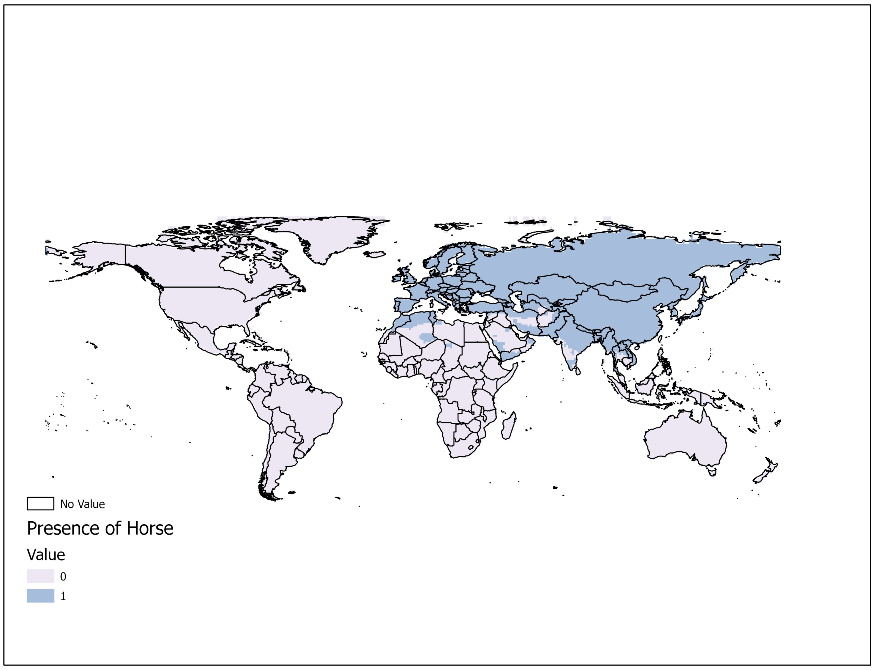

Figure A1: Distribution of Equus ferus in Prehistoric Time

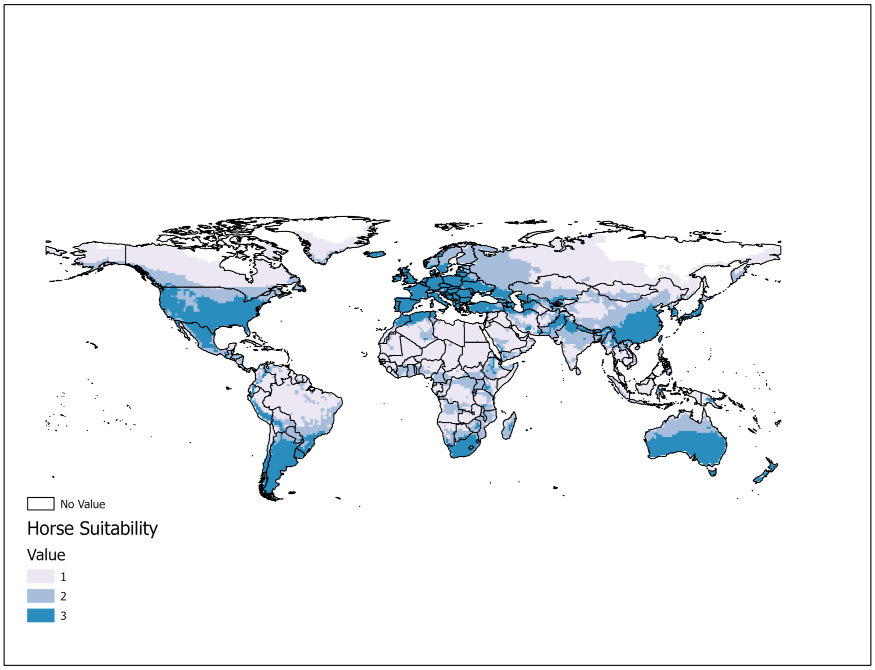

Figure A2: Distribution of Climatic Horse Suitability

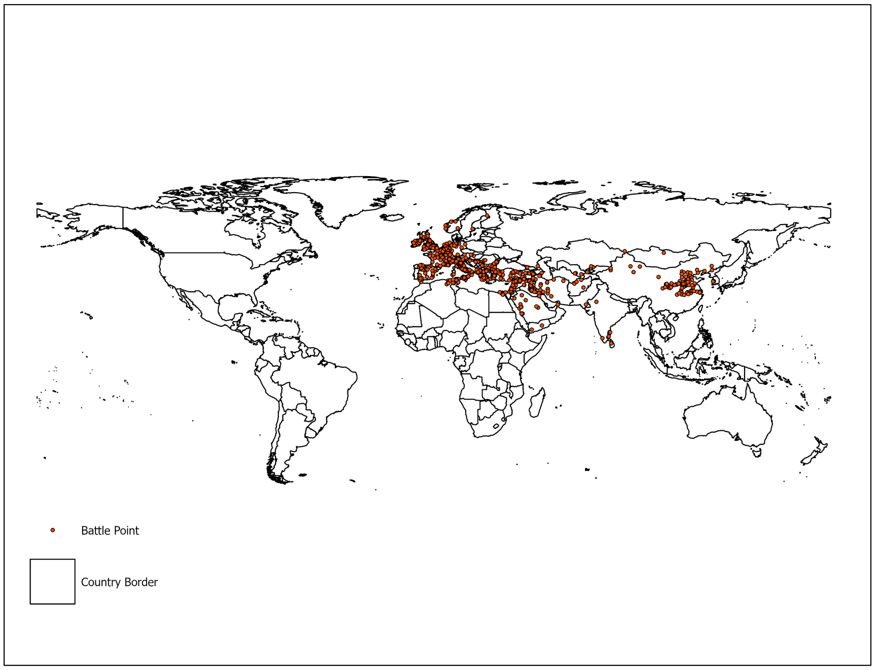

Figure A3: Distribution of Historical Battles Before 900 CE

Appendix B. Table

Table B1: Summary Statistics: Virtual Country

	Mean	SD	Min	Max	N
Dependent Variable					
City before 400 CE (DeGroff, 2009)	0.05	0.21	0.00	1.00	18414
City before 500 BCE (Reba et al., 2016)	0.00	0.07	0.00	1.00	18414
City before 450 CE (Reba et al., 2016)	0.01	0.08	0.00	1.00	18414
Log Dist. to City before 450 CE (Reba et al., 2016)	7.55	1.16	0.00	9.25	18414
Log Dist. to City before 500 BCE (Reba et al., 2016)	7.39	1.29	0.00	9.25	18414
Log Dist. to the Closest Battle before 1000 CE	7.35	1.79	0.00	9.46	18414
Empire Density	8.09	8.85	0.00	31.00	4287
Population Count in 2000 BC	34.87	123.99	0.00	3971.73	17722
Population Density in 2000 BC	0.49	1.73	0.00	64.15	17722
Urbanization in 2000 BC	0.78	10.37	0.00	535.70	17722
Independent Variables					
Log Dist. to Tell el-Ajjul	8.65	0.64	3.97	9.71	16099
Historical Horse Index	0.72	1.02	0.00	3.00	13321
Time since Cavalry Emergence as of 400 CE	350.83	426.03	0.00	1400.00	8727
Time since Cavalry Emergence as of 450 CE	376.95	445.94	0.00	1450.00	8727
Time since Cavalry Emergence as of 900 BCE	634.89	623.86	0.00	1900.00	8727
Control Variables					
Latitude	32.92	32.91	-58.43	83.34	18414
Longitude	16.95	87.64	-179.93	179.63	18414
Latitude × Longitude	248.32	4424.41	-12795.41	12787.17	18414
Elevation (Avg.)	588.20	788.86	-3042.50	5746.12	18136
Land Productivity (Avg.)	0.23	0.29	0.00	1.00	15762
Elevation (Std)	136.38	163.99	0.00	1746.37	18136
Land Productivity (Std)	0.04	0.06	0.00	0.42	15762
Island	0.01	0.12	0.00	1.00	18414
Temperature (Avg.)	8.12	15.02	-27.61	29.96	16713
Precipitation (Avg.)	59.38	57.41	0.00	607.31	16713
Temperature (Std)	8.41	5.31	0.17	22.99	16713
Precipitation (Std)	42.28	39.10	0.00	436.39	16713
Transport Mammal Dummy	0.37	0.48	0.00	1.00	16194
Wild Relatives of Cereals (Dummy)	0.09	0.29	0.00	1.00	18414
Log Dist. to Agricultural Frontier	7.44	1.46	0.00	9.21	18414
Log Dist. to the Lower Volga-Don	8.40	0.92	0.00	9.64	16099
Log Dist. to Near East	8.54	0.86	0.00	9.79	18414
Log Dist. to Northern China	8.70	0.83	0.00	9.88	18414
Log Dist. to Southern China	8.73	0.86	0.00	9.87	18414
Log Dist. to West African Sub-Sahara	8.61	1.07	0.00	9.82	18414
Log Dist. to Itjtawy	8.70	0.64	0.00	9.83	18414
Log Dist. to Susa	8.68	0.65	0.00	9.86	18414
Log Dist. to Knossos	8.67	0.63	0.00	9.82	18414
Log Dist. to Erligang	8.82	0.67	0.00	9.90	18414
Log Dist. to Eridu	8.68	0.65	0.00	9.86	18414
Log Dist. to Trade Routes	7.25	1.92	0.00	9.46	18414
Time since Agricultural Transition	4859.53	2059.38	414.05	11091.54	4287
Time since Iron Emergence as of 400 CE	646.58	334.30	0.00	1780.00	7704
Time since Iron Emergence as of 900 BCE	1134.55	361.04	0.00	2280.00	7704

Table B2: Summary Statistics: Cross-Country

	Mean	SD	Min	Max	N
Dependent Variable					
State History	786.00	579.24	56.06	2,495.01	159
Independent Variables					
Dist. to Tell el-Ajjul	8.27	0.88	4.16	9.69	168
Historical Horse Index	1.05	1.28	0.00	3.00	168
Time since Cavalry Emergence as of 2000 CE	1,655.62	812.89	365.00	3,000.00	127
Control Variables					
Log Dist. to the Neolithic Origin	6.78	2.26	0.00	9.15	234
Latitude	17.38	26.27	-54.28	78.22	201
Longitude	9.94	70.29	-176.17	174.78	201
Latitude × Longitude	395.26	1705.83	-7215.90	5120.53	201
Elevation (Avg.)	404.14	657.12	-2066.00	3059.91	226
Land Productivity (Avg.)	0.40	0.26	0.00	0.96	171
Elevation (Std)	351.14	329.48	0.00	1859.54	226
Land Productivity (Std)	0.17	0.10	0.00	0.40	171
Island	0.33	0.47	0.00	1.00	224
Temperature (Avg.)	17.83	8.93	-17.85	28.31	184
Precipitation (Avg.)	93.18	65.27	2.86	275.39	184
Temperature (Std)	4.46	3.30	0.37	15.07	184
Precipitation (Std)	66.30	44.77	3.71	193.50	184
Log Dist. to Trade Route	6.42	1.87	0.00	9.23	168
Transport Mammal Dummy	0.56	0.50	0.00	1.00	197
Agricultural Transition Timing	0.50	0.50	0.00	1.00	151
Neolithic Transition Timing	4847.58	2510.06	362.00	10500.00	168
hspace0.25cm Time since Iron Emergence as of 2000 CE	2230.65	419.15	1000.00	2957.51	125

Table B3: Summary Statistics: Country Panel

	Mean	SD	Min	Max	N
Dependent Variable					
State Index	17.89	18.61	0.00	50.00	460
Independent Variables					
Historical Horse Index (Lagged)	0.72	0.96	0.00	3.00	452
Land Productivity (Lagged)	9,445.72	3,134.59	1,659.75	17,994.58	460
Control Variables					
Latitude	3.65	21.80	-34.90	45.41	460
Longitude	-74.09	12.56	-99.13	-47.93	460
Latitude × Longitude	-433.18	1530.19	-3437.54	2363.67	460
Elevation (Avg.)	600.80	374.81	89.55	1482.12	460
Elevation (Std)	611.07	431.63	59.58	1594.36	460
Island	0.13	0.34	0.00	1.00	460
Temperature (Avg.)	20.23	7.71	-7.02	25.81	460
Precipitation (Avg.)	134.17	61.58	37.39	254.51	460
Temperature (Std)	2.65	2.98	0.66	13.03	460
Precipitation (Std)	88.98	41.32	20.57	178.13	460

Table B4: Summary Statistics: Cross-Ethnic Group

	Mean	SD	Min	Max	N
Dependent Variable					
Centralization	1.97	1.12	1.00	5.00	1173
Social Stratification	1.81	0.86	1.00	3.00	1127
Mean Size of Community	3.66	2.28	1.00	8.00	620
Independent Variables					
Log Dist. to Tell el-Ajjul	8.60	0.69	4.08	9.54	1163
Historical Horse Index	0.35	0.86	0.00	3.00	1139
Time since Cavalry Emergence	1136.02	669.30	75.00	2930.00	718
Log Dist. to Agricultural Frontier	7.05	1.53	0.00	9.00	1163
Latitude	15.50	22.74	-55.00	78.00	1309
Longitude	3.16	84.42	-178.00	179.00	1309
Latitude $\times$ Longitude	-696.29	2634.36	-11147.00	11682.00	1309
Elevation (Avg.)	690.18	698.14	-1732.00	5477.84	1252
Land Productivity (Avg.)	0.36	0.29	0.00	1.00	1189
Elevation (Std)	197.54	189.05	0.00	1333.05	1252
Land Productivity (Std)	0.07	0.07	0.00	0.40	1189
Island	0.00	0.06	0.00	1.00	1309
Temperature (Avg.)	18.92	9.17	-15.95	29.92	1214
Precipitation (Avg.)	99.56	66.97	0.62	474.63	1214
Temperature (Std)	4.06	3.68	0.18	20.67	1214
Precipitation (Std)	75.56	46.81	0.88	284.88	1214
Cereal Advantage	47044.89	27385.39	0.00	159304.53	1192
Transport Mammal Dummy	0.26	0.44	0.00	1.00	1225
Time since Iron Emergence	2032.72	388.17	40.00	2964.18	672

Table B5: Summary Statistics: Radiocarbon-Dated Archaeological Sites

	Mean	SD	Min	Max	N
Dependent Variable					
Ancient Ruin (Dummy)	0.05	0.22	0.00	1.00	30754
Ancient Ruin (Count)	0.11	0.70	0.00	23.00	30754
Ancient Settlement (Dummy)	0.04	0.19	0.00	1.00	30754
Ancient Settlement (Count)	0.08	0.59	0.00	21.00	30754
Independent Variables					
Log Dist. to Tell el-Ajjul	8.64	0.64	3.97	9.71	30222
Historical Horse Index	0.71	1.02	0.00	3.00	25604
Control Variables					
Log Dist. to the Neolithic Origin	7.35	1.53	0.00	9.17	30754
Log Dist. to Trade Routes	7.21	1.84	0.00	9.31	30222
Log Dist. to Volga-Don	1.76	0.58	0.00	2.79	30222
Transport Mammal Dummy	0.39	0.49	0.00	1.00	29760
Wild Relatives of Cereals (Dummy)	0.19	0.39	0.00	1.00	30754
Wild Relatives of Roots and Tubers (Dummy)	0.08	0.27	0.00	1.00	30754
Wild Relatives of Cereals, Roots and Tubers (Dummy)	0.05	0.22	0.00	1.00	30754

Appendix C. Table

Table C1: The Association between Ancient Cities and Horse-Riding: Trade routes

	Presende of Ancient Cities					
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS
Time since Cavalry Emergence (400 CE)	0.035**	0.050***	0.052**	0.073***	0.027**	0.075**
	(0.014)	(0.016)	(0.025)	(0.026)	(0.011)	(0.037)
Log Dist. to Trade Routes	-0.042***	-0.050***	-0.032***	-0.031***	-0.023***	-0.020***
	(0.007)	(0.008)	(0.006)	(0.007)	(0.005)	(0.005)
Log Dist. to Agricultural Frontier			-0.019***	-0.021***	-0.009	-0.007
			(0.006)	(0.005)	(0.007)	(0.007)
Latitude			-0.147***	-0.052	0.047	0.029
			(0.052)	(0.063)	(0.063)	(0.074)
Longitude			-0.215**	-0.238***	0.087	0.117*
			(0.091)	(0.076)	(0.069)	(0.061)
Latitude × Longitude			0.096	0.217***	-0.055	-0.079
			(0.098)	(0.067)	(0.064)	(0.059)
Elevation (Avg.)			-0.016	-0.023	-0.038**	-0.057**
			(0.013)	(0.014)	(0.017)	(0.024)
Land Productivity (Avg.)			0.037**	0.040**	0.023***	0.019**
			(0.017)	(0.017)	(0.008)	(0.009)
Elevation (Std)			0.037**	0.028**	0.012	0.007
			(0.017)	(0.013)	(0.008)	(0.007)
Land Productivity (Std)			-0.004	-0.003	0.002	-0.000
			(0.008)	(0.007)	(0.004)	(0.004)
Island Dummy			0.103	0.076	0.196	0.172
			(0.282)	(0.279)	(0.228)	(0.235)
Temperature (Avg.)				-0.039	-0.092	-0.165**
				(0.043)	(0.058)	(0.077)
Precipitation (Avg.)				0.023	-0.021	-0.023
				(0.015)	(0.014)	(0.017)
Temperature (Std)				-0.149***	-0.085**	-0.126**
				(0.038)	(0.035)	(0.058)
Precipitation (Std)				-0.031**	0.013	0.021
				(0.013)	(0.009)	(0.014)
Continent FE		✓	✓	✓		
Country FE					✓	✓
First-F						11.341
J-Test (p-value)						0.174
Adjusted R^2	0.146	0.180	0.261	0.281	0.451	
Observations	7406	7406	7406	7406	7406	7406

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C2: The Association between Ancient Cities and Horse-Riding: Transport Mammals

	Presende of Ancient Cities					
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS
Time since Cavalry Emergence (400 CE)	0.072***	0.076***	0.074***	0.092***	0.040***	0.108**
	(0.019)	(0.019)	(0.024)	(0.027)	(0.012)	(0.041)
Transport Mammal Dummy	0.025	0.057*	0.010	0.018	-0.017	-0.050**
	(0.017)	(0.031)	(0.032)	(0.035)	(0.015)	(0.025)
Log Dist. to Agricultural Frontier			-0.018***	-0.021***	-0.009	-0.007
			(0.007)	(0.006)	(0.007)	(0.007)
Latitude			-0.162***	-0.061	0.055	0.014
			(0.058)	(0.060)	(0.084)	(0.098)
Longitude			-0.231**	-0.226***	0.053	0.076
			(0.094)	(0.076)	(0.084)	(0.082)
Latitude × Longitude			0.087	0.195***	-0.034	-0.054
			(0.089)	(0.062)	(0.075)	(0.075)
Elevation (Avg.)			-0.018	-0.029**	-0.034*	-0.056**
			(0.012)	(0.013)	(0.018)	(0.026)
Land Productivity (Avg.)			0.046***	0.049***	0.027***	0.024***
			(0.017)	(0.016)	(0.007)	(0.008)
Elevation (Std)			0.044***	0.035***	0.017**	0.009
			(0.016)	(0.013)	(0.007)	(0.007)
Land Productivity (Std)			-0.004	-0.001	0.003	0.001
			(0.007)	(0.006)	(0.004)	(0.004)
Island Dummy			0.095	0.071	0.201	0.166
			(0.272)	(0.266)	(0.221)	(0.233)
Temperature (Avg.)				-0.045	-0.047	-0.146*
				(0.049)	(0.063)	(0.086)
Precipitation (Avg.)				-0.001	-0.032**	-0.038**
				(0.014)	(0.013)	(0.019)
Temperature (Std)				-0.169***	-0.084**	-0.135**
				(0.045)	(0.042)	(0.067)
Precipitation (Std)				-0.022*	0.017*	0.028*
				(0.013)	(0.010)	(0.016)
Continent FE		✓	✓	✓		
Country FE					✓	✓
First-F						12.912
J-Test (p-value)						0.229
Adjusted R^2	0.080	0.102	0.228	0.252	0.436	
Observations	7395	7395	7395	7395	7395	7395

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances, island dummy, transport mammal dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C3: The Association between Ancient Cities and Horse-Riding: Time since Agricultural Transition

	Presende of Ancient Cities					
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS
Time since Cavalry Emergence (400 CE)	0.063***	0.058***	0.100***	0.142***	0.070***	0.262**
	(0.016)	(0.018)	(0.028)	(0.039)	(0.021)	(0.105)
Time since Agricultural Transition	0.120***	0.124***	0.086***	0.079***	-0.029	-0.072*
	(0.027)	(0.022)	(0.023)	(0.022)	(0.025)	(0.039)
Log Dist. to Agricultural Frontier			-0.013**	-0.016**	-0.010	-0.003
			(0.005)	(0.006)	(0.008)	(0.008)
Latitude			-0.168**	-0.104	0.139	0.087
			(0.081)	(0.083)	(0.108)	(0.162)
Longitude			-0.129	-0.145*	-0.014	0.212
			(0.090)	(0.075)	(0.120)	(0.150)
Latitude × Longitude			0.015	0.209**	0.026	0.106
			(0.119)	(0.091)	(0.128)	(0.175)
Elevation (Avg.)			-0.024*	-0.080***	-0.072***	-0.122***
			(0.014)	(0.023)	(0.024)	(0.041)
Land Productivity (Avg.)			0.039***	0.033***	0.033***	0.025**
			(0.012)	(0.011)	(0.010)	(0.010)
Elevation (Std)			0.027**	0.026***	0.014	0.003
			(0.010)	(0.010)	(0.010)	(0.009)
Land Productivity (Std)			0.003	-0.001	0.009**	0.005
			(0.008)	(0.007)	(0.005)	(0.005)
Island Dummy			0.461	0.450	0.437	0.420
			(0.355)	(0.336)	(0.333)	(0.314)
Temperature (Avg.)				-0.194***	-0.095	-0.283**
				(0.060)	(0.084)	(0.120)
Precipitation (Avg.)				-0.039***	-0.053***	-0.071**
				(0.014)	(0.016)	(0.028)
Temperature (Std)				-0.314***	-0.203***	-0.447***
				(0.087)	(0.065)	(0.133)
Precipitation (Std)				-0.008	0.008	0.025
				(0.011)	(0.010)	(0.025)
Continent FE		✓	✓	✓		
Country FE					✓	✓
First-F						15.538
J-Test (p-value)						0.909
Adjusted R^2	0.207	0.255	0.297	0.314	0.429	
Observations	3680	3680	3680	3680	3680	3680

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1 × 1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C4: The Association between Ancient Cities and Horse-Riding: Spread of Iron

	Presende of Ancient Cities					
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS
Time since Cavalry Emergence (400 CE)	0.044**	0.053***	0.056**	0.076***	0.042***	0.143***
	(0.018)	(0.016)	(0.027)	(0.024)	(0.011)	(0.050)
Time since Iron Emergence (400 CE)	0.059***	0.065***	0.041***	0.051***	0.009	-0.018
	(0.019)	(0.019)	(0.015)	(0.015)	(0.010)	(0.013)
Log Dist. to Agricultural Frontier			-0.021***	-0.024***	-0.008	-0.004
			(0.007)	(0.006)	(0.007)	(0.007)
Latitude			-0.069	-0.022	0.097	0.058
			(0.057)	(0.065)	(0.083)	(0.115)
Longitude			-0.106	-0.160*	0.026	-0.049
			(0.091)	(0.089)	(0.111)	(0.113)
Latitude × Longitude			-0.102	0.150	0.088	0.305*
			(0.086)	(0.120)	(0.138)	(0.173)
Elevation (Avg.)			-0.009	-0.041**	-0.037*	-0.066**
			(0.010)	(0.017)	(0.022)	(0.030)
Land Productivity (Avg.)			0.048**	0.046**	0.026***	0.023***
			(0.019)	(0.018)	(0.007)	(0.008)
Elevation (Std)			0.040***	0.036***	0.018**	0.003
			(0.014)	(0.013)	(0.007)	(0.008)
Land Productivity (Std)			-0.001	-0.003	0.003	0.001
			(0.007)	(0.006)	(0.004)	(0.004)
Island Dummy			0.072	0.056	0.198	0.153
			(0.266)	(0.267)	(0.223)	(0.240)
Temperature (Avg.)				-0.110**	-0.048	-0.161
				(0.053)	(0.080)	(0.100)
Precipitation (Avg.)				0.004	-0.032**	-0.039*
				(0.016)	(0.015)	(0.021)
Temperature (Std)				-0.220***	-0.154***	-0.294***
				(0.051)	(0.044)	(0.079)
Precipitation (Std)				-0.033**	0.012	0.025
				(0.016)	(0.010)	(0.018)
Continent FE		✓	✓	✓		
Country FE					✓	✓
First-F						11.507
J-Test (p-value)						0.316
Adjusted R^2	0.098	0.125	0.249	0.266	0.437	
Observations	6886	6886	6886	6886	6886	6886

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C5: The Association between Ancient Cities and Horse-Riding: Wild Relatives of Cereals

	Presende of Ancient Cities					
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS
Time since Cavalry Emergence (400 CE)	0.080***	0.092***	0.079***	0.093***	0.038***	0.099**
	(0.017)	(0.019)	(0.027)	(0.029)	(0.013)	(0.039)
Wild Relatives of Cereals (Dummy)	0.157***	0.175***	0.076**	0.068**	0.043**	0.045**
	(0.035)	(0.035)	(0.031)	(0.026)	(0.018)	(0.018)
Log Dist. to Agricultural Frontier			-0.018**	-0.020***	-0.009	-0.006
			(0.007)	(0.006)	(0.008)	(0.007)
Latitude			-0.141**	-0.050	0.067	0.040
			(0.057)	(0.063)	(0.078)	(0.090)
Longitude			-0.197**	-0.193**	0.070	0.114
			(0.093)	(0.078)	(0.078)	(0.076)
Latitude × Longitude			0.063	0.167***	-0.045	-0.080
			(0.094)	(0.063)	(0.072)	(0.071)
Elevation (Avg.)			-0.018	-0.026*	-0.036**	-0.060**
			(0.013)	(0.014)	(0.018)	(0.026)
Land Productivity (Avg.)			0.043**	0.049***	0.024***	0.018*
			(0.018)	(0.018)	(0.008)	(0.010)
Elevation (Std)			0.042***	0.035***	0.018**	0.010
			(0.015)	(0.013)	(0.007)	(0.007)
Land Productivity (Std)			-0.007	-0.004	0.002	-0.000
			(0.007)	(0.006)	(0.004)	(0.004)
Island Dummy			0.081	0.067	0.198	0.165
			(0.263)	(0.257)	(0.215)	(0.225)
Temperature (Avg.)				-0.035	-0.047	-0.150*
				(0.044)	(0.062)	(0.086)
Precipitation (Avg.)				0.001	-0.025*	-0.026
				(0.014)	(0.014)	(0.018)
Temperature (Std)				-0.157***	-0.083**	-0.136**
				(0.040)	(0.041)	(0.065)
Precipitation (Std)				-0.031**	0.013	0.023
				(0.012)	(0.010)	(0.015)
Continent FE		✓	✓	✓		
Country FE					✓	✓
First-F						11.686
J-Test (p-value)						0.178
Adjusted R^2	0.117	0.144	0.239	0.260	0.442	
Observations	7406	7406	7406	7406	7406	7406

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances, island dummy, dummy of wild relatives of cereals are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C6: The Association between Ancient Cities and Horse-Riding: Another Data (Dummy)

	Dummy Anc. Cities (500 BCE)		Dummy And. Cities (450 CE)	
	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
Time since Cavalry Emergence (500 BCE)	0.001	0.007		
	(0.002)	(0.012)		
Time since Cavalry Emergence (450 CE)			0.005	0.017
			(0.003)	(0.016)
Log Dist. to Agricultural Frontier	-0.007*	-0.006	-0.005	-0.005
	(0.003)	(0.004)	(0.004)	(0.004)
Latitude	0.034**	0.041**	0.026	0.020
	(0.016)	(0.018)	(0.020)	(0.026)
Longitude	0.007	0.023	0.005	0.013
	(0.021)	(0.034)	(0.020)	(0.027)
Latitude $\times$ Longitude	0.004	-0.008	0.006	-0.000
	(0.018)	(0.028)	(0.020)	(0.028)
Elevation (Avg.)	0.001	0.000	-0.003	-0.008
	(0.003)	(0.004)	(0.005)	(0.009)
Land Productivity (Avg.)	0.002	0.002	0.004	0.002
	(0.002)	(0.002)	(0.003)	(0.003)
Elevation (Std)	-0.001	-0.002	-0.002	-0.003
	(0.001)	(0.002)	(0.002)	(0.002)
Land Productivity (Std)	0.002	0.002	0.002	0.001
	(0.001)	(0.002)	(0.002)	(0.002)
Island Dummy	0.004	-0.001	-0.000	-0.006
	(0.005)	(0.012)	(0.008)	(0.013)
Temperature (Avg.)	0.014	0.011	0.004	-0.016
	(0.009)	(0.012)	(0.014)	(0.032)
Precipitation (Avg.)	-0.007	-0.008	-0.010	-0.010
	(0.005)	(0.005)	(0.006)	(0.007)
Temperature (Std)	-0.006	-0.008	-0.014	-0.024
	(0.006)	(0.008)	(0.010)	(0.021)
Precipitation (Std)	0.006**	0.007**	0.007*	0.008
	(0.003)	(0.003)	(0.004)	(0.005)
Country FE	✓	✓	✓	✓
First-F		5.681		12.086
J-Test (p-value)		0.133		0.343
Adjusted R^2	0.099		0.098	
Observations	7436	7436	7436	7436

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C7: The Association between Ancient Cities and Horse-Riding: Another Data (Distance)

	Log Dist. to Anc. Cities (500 BCE)		Log Dist. to And. Cities (450 CE)	
	(1) OLS	(2) IV	(3) OLS	(4) IV
Time since Cavalry Emergence (500 BCE)	-0.035	-0.484**		
	(0.039)	(0.239)		
Time since Cavalry Emergence (450 CE)			-0.293***	-0.603***
			(0.067)	(0.201)
Log Dist. to Agricultural Frontier	0.170**	0.137	0.136*	0.123
	(0.082)	(0.106)	(0.081)	(0.076)
Latitude	-2.179***	-2.657***	-1.289**	-1.155*
	(0.711)	(0.708)	(0.559)	(0.600)
Longitude	-2.319***	-3.414***	-1.836**	-2.044**
	(0.807)	(1.064)	(0.808)	(0.883)
Latitude × Longitude	1.653**	2.481***	1.403*	1.567*
	(0.716)	(0.936)	(0.727)	(0.826)
Elevation (Avg.)	-0.314***	-0.262**	-0.091	0.036
	(0.103)	(0.117)	(0.076)	(0.116)
Land Productivity (Avg.)	-0.126**	-0.085	-0.122*	-0.091
	(0.059)	(0.077)	(0.067)	(0.085)
Elevation (Std)	0.012	0.082**	-0.023	0.015
	(0.020)	(0.040)	(0.026)	(0.030)
Land Productivity (Std)	-0.009	0.026	0.000	0.013
	(0.018)	(0.025)	(0.020)	(0.021)
Island Dummy	-0.159	0.159	0.017	0.174
	(0.180)	(0.328)	(0.170)	(0.217)
Temperature (Avg.)	-1.434***	-1.221***	-0.703**	-0.176
	(0.455)	(0.447)	(0.290)	(0.399)
Precipitation (Avg.)	0.119	0.149	0.194	0.193
	(0.079)	(0.102)	(0.133)	(0.149)
Temperature (Std)	-0.223	-0.097	-0.003	0.271
	(0.178)	(0.257)	(0.237)	(0.319)
Precipitation (Std)	-0.156***	-0.223***	-0.129**	-0.176**
	(0.045)	(0.063)	(0.059)	(0.087)
Country FE	✓	✓	✓	✓
First-F		5.681		12.086
J-Test (p-value)		0.051		0.182
Adjusted R^2	0.747		0.743	
Observations	7436	7436	7436	7436

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C8: The Association between State History and Horse-Riding: Historical Long-Distance Trade Routes

	State History				
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) 2SLS
Time since Cavalry Emergence	287.729***	230.942***	158.793*	207.422*	416.948**
	(71.157)	(77.307)	(84.769)	(111.739)	(177.631)
Log Dist. to Trade Route	-76.829	-86.835	-145.664***	-122.528**	-79.016
	(48.878)	(52.838)	(53.254)	(55.617)	(60.036)
Latitude			577.322***	623.263***	467.620**
			(168.324)	(191.801)	(205.005)
Longitude			752.972***	654.454**	726.772***
			(268.723)	(278.351)	(260.723)
Latitude × Longitude			-485.443***	-260.060	-227.188
			(161.516)	(191.200)	(185.604)
Elevation (Avg.)			-255.852***	-219.507	-305.270*
			(84.811)	(161.725)	(156.383)
Land Productivity (Avg.)			-16.534	-39.954	-78.973
			(62.549)	(70.520)	(68.792)
Elevation (Std)			263.139***	224.224**	211.102**
			(70.694)	(90.807)	(84.644)
Land Productivity (Std)			6.865	19.657	-9.413
			(75.865)	(86.603)	(75.615)
Island Dummy			123.947	126.126	244.542
			(245.999)	(251.568)	(244.977)
Temperature (Avg.)				15.701	-233.642
				(242.910)	(265.036)
Precipitation (Avg.)				-131.557	-248.441*
				(130.407)	(145.911)
Temperature (Std)				-257.577	-439.365**
				(168.148)	(197.540)
Precipitation (Std)				25.974	141.875
				(106.888)	(119.940)
Continent FE		✓	✓	✓	✓
First Stage F-Statistics					32.528
J-Test (p-value)					0.574
Adjusted R^2	0.415	0.442	0.578	0.584	
Observations	94	94	94	94	94

Note: OLS regressions with robust standard errors. The unit of analysis is a country, restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C9: The Association between State History and Horse-Riding: Transport Mammals

	State History				
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) 2SLS
Time since Cavalry Emergence	266.020***	243.107***	288.091***	321.500***	570.120***
	(61.475)	(60.970)	(78.320)	(107.538)	(121.925)
Transport Mammal Dummy	398.377***	378.263***	51.836	-1.335	-96.225
	(114.171)	(132.384)	(170.241)	(197.223)	(170.192)
Log Dist. to Agricultural Frontier			-37.596	-35.095	-50.422
			(28.760)	(33.766)	(30.900)
Latitude			520.108**	578.993**	427.892*
			(203.318)	(232.992)	(229.499)
Longitude			835.380**	746.229**	875.400***
			(345.212)	(358.293)	(311.258)
Latitude × Longitude			-508.728**	-263.088	-240.621
			(199.272)	(221.786)	(215.618)
Elevation (Avg.)			-185.942**	-166.346	-307.214*
			(84.431)	(160.043)	(160.466)
Land Productivity (Avg.)			9.830	-21.486	-90.942
			(65.695)	(76.293)	(70.654)
Elevation (Std)			219.788***	183.003**	181.486**
			(75.169)	(90.520)	(84.313)
Land Productivity (Std)			2.385	27.469	-3.258
			(77.624)	(89.336)	(78.795)
Island Dummy			49.126	70.179	240.501
			(274.089)	(261.101)	(259.922)
Temperature (Avg.)				16.392	-301.650
				(280.323)	(291.430)
Precipitation (Avg.)				-209.252*	-332.543***
				(122.859)	(128.740)
Temperature (Std)				-306.179	-521.241**
				(186.973)	(202.600)
Precipitation (Std)				48.411	199.881*
				(111.407)	(110.568)
Continent FE		✓	✓	✓	✓
First Stage F-Statistics					37.911
J-Test (p-value)					0.610
Adjusted R^2	0.443	0.463	0.529	0.550	
Observations	94	94	94	94	94

Note: OLS regressions with robust standard errors. The unit of analysis is a country, restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances, island dummy, transport mammal dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C10: The Association between State History and Horse-Riding: Time since Agricultural Transition

	State History				
	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	OLS	OLS	2SLS
Time since Cavalry Emergence	120.735	96.028	54.154	113.551	426.240**
	(77.530)	(77.745)	(76.537)	(107.252)	(189.113)
Time since Agricultural Transition	346.401***	401.253***	456.087***	415.438***	245.802*
	(82.838)	(118.321)	(111.621)	(121.058)	(141.084)
Latitude			603.325***	644.080***	423.556*
			(171.508)	(202.695)	(229.429)
Longitude			987.055***	932.734***	882.674***
			(276.700)	(279.132)	(251.771)
Latitude × Longitude			-497.702***	-373.663*	-266.859
			(170.147)	(192.069)	(205.812)
Elevation (Avg.)			-126.540	-112.184	-282.173*
			(77.777)	(150.483)	(171.169)
Land Productivity (Avg.)			-6.519	-37.005	-97.301
			(63.057)	(70.056)	(73.989)
Elevation (Std)			123.662*	107.581	147.468*
			(67.059)	(85.671)	(85.613)
Land Productivity (Std)			71.452	65.597	0.600
			(67.956)	(80.817)	(75.712)
Island Dummy			-44.513	14.954	215.094
			(250.409)	(262.328)	(251.628)
Temperature (Avg.)				-6.634	-340.325
				(228.164)	(279.801)
Precipitation (Avg.)				-115.845	-251.976*
				(130.380)	(140.163)
Temperature (Std)				-165.668	-448.535**
				(163.232)	(222.678)
Precipitation (Std)				75.418	200.511*
				(103.016)	(106.946)
Continent FE		✓	✓	✓	✓
First Stage F-Statistics					17.306
J-Test (p-value)					0.148
Adjusted R^2	0.491	0.507	0.611	0.604	
Observations	92	92	92	92	92

Note: OLS regressions with robust standard errors. The unit of analysis is a country, restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C11: The Association between State History and Horse-Riding: Spread of Iron

	State History				
	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	OLS	OLS	2SLS
Time since Cavalry Emergence	263.091**	235.408*	128.443	166.714	574.131***
	(103.591)	(120.119)	(113.244)	(138.409)	(207.802)
Time since Iron Emergence	134.887	115.084	225.133**	180.746*	-13.211
	(107.095)	(108.197)	(93.518)	(97.665)	(122.437)
Log Dist. to Agricultural Frontier			-48.952*	-39.894	-46.569
			(25.167)	(30.653)	(29.721)
Latitude			613.271***	639.927***	374.327*
			(179.294)	(188.688)	(207.951)
Longitude			1057.489***	902.339***	790.981***
			(292.509)	(291.068)	(281.638)
Latitude × Longitude			-555.963***	-327.287*	-209.709
			(167.580)	(195.083)	(214.019)
Elevation (Avg.)			-174.485**	-133.692	-318.243*
			(68.216)	(148.651)	(170.436)
Land Productivity (Avg.)			13.787	-9.215	-87.439
			(65.652)	(74.645)	(74.609)
Elevation (Std)			211.123***	175.767**	185.727**
			(65.507)	(83.914)	(85.930)
Land Productivity (Std)			5.480	29.668	-18.285
			(74.134)	(87.001)	(77.529)
Island Dummy			58.031	90.851	266.370
			(231.633)	(249.204)	(253.366)
Temperature (Avg.)				72.055	-344.425
				(257.550)	(289.310)
Precipitation (Avg.)				-166.569	-339.928**
				(133.097)	(142.196)
Temperature (Std)				-217.134	-547.812**
				(174.058)	(236.288)
Precipitation (Std)				33.967	203.630*
				(118.729)	(112.332)
Continent FE		✓	✓	✓	✓
First Stage F-Statistics					14.949
J-Test (p-value)					0.670
Adjusted R^2	0.398	0.425	0.557	0.566	
Observations	93	93	93	93	93

Note: OLS regressions with robust standard errors. The unit of analysis is a country, restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C12: The Association between State History and Horse-Riding: Wild Relatives of Cereals

	State History				
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) 2SLS
Time since Cavalry Emergence	280.265***	250.288***	258.780***	273.144**	517.303***
	(69.464)	(71.241)	(77.514)	(104.797)	(128.005)
Wild Relatives of Cereals (Dummy)	274.064*	290.858*	209.394	252.985*	167.763
	(146.720)	(148.642)	(127.572)	(131.806)	(127.924)
Log Dist. to Agricultural Frontier			-37.809	-33.208	-45.039
			(24.980)	(30.555)	(28.686)
Latitude			484.637***	525.478***	354.967**
			(181.979)	(184.282)	(177.778)
Longitude			848.030***	669.255**	748.886***
			(295.782)	(293.314)	(270.540)
Latitude × Longitude			-499.372***	-173.655	-158.067
			(174.182)	(198.948)	(195.101)
Elevation (Avg.)			-180.639**	-156.004	-300.543*
			(82.778)	(158.293)	(158.161)
Land Productivity (Avg.)			-1.489	-28.107	-87.330
			(66.320)	(74.219)	(69.718)
Elevation (Std)			207.449***	168.593*	175.485**
			(75.210)	(89.514)	(84.202)
Land Productivity (Std)			12.028	33.909	-11.161
			(78.131)	(87.822)	(74.801)
Island Dummy			70.351	70.267	256.610
			(245.628)	(234.143)	(225.295)
Temperature (Avg.)				16.113	-320.198
				(260.300)	(264.840)
Precipitation (Avg.)				-166.207	-301.080**
				(128.516)	(136.283)
Temperature (Std)				-354.680**	-560.109***
				(176.271)	(186.919)
Precipitation (Std)				-5.898	157.901
				(114.416)	(118.023)
Continent FE		✓	✓	✓	✓
First Stage F-Statistics					45.535
J-Test (p-value)					0.983
Adjusted R^2	0.424	0.450	0.543	0.570	
Observations	94	94	94	94	94

Note: OLS regressions with robust standard errors. The unit of analysis is a country, restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances, island dummy, dummy of wild relatives of cereals are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C13: The Association between Proxies of the State and Horse-Riding: Historical Long-Distance Trade Routes

	Centralization		Social Stratification		Community Size	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
Time since Cavalry Emergence	0.102	0.297*	0.099	0.189***	0.295	0.623*
	(0.081)	(0.153)	(0.061)	(0.059)	(0.291)	(0.360)
Log Dist. to Trade Route	-0.055*	-0.028	-0.041*	-0.028	-0.219***	-0.164**
	(0.029)	(0.027)	(0.020)	(0.026)	(0.072)	(0.074)
Log Dist. to Agricultural Frontier	-0.019	-0.029	-0.026	-0.030	-0.336**	-0.336**
	(0.066)	(0.064)	(0.034)	(0.033)	(0.153)	(0.154)
Latitude	0.130	-0.099	-0.309	-0.408	1.414**	1.129**
	(0.478)	(0.556)	(0.314)	(0.386)	(0.646)	(0.542)
Longitude	-0.323	-0.322	-0.441	-0.432	-0.795	-0.690
	(0.565)	(0.527)	(0.278)	(0.262)	(0.792)	(0.781)
Latitude × Longitude	-0.909**	-0.827**	-0.257	-0.226	-0.240	-0.153
	(0.331)	(0.331)	(0.208)	(0.204)	(0.728)	(0.662)
Elevation (Avg.)	-0.032	-0.122	-0.019	-0.061	0.362	0.219
	(0.099)	(0.120)	(0.060)	(0.075)	(0.296)	(0.260)
Land Productivity (Avg.)	0.184**	0.184***	0.043	0.044	0.582***	0.587***
	(0.072)	(0.064)	(0.042)	(0.038)	(0.155)	(0.152)
Elevation (Std)	-0.140**	-0.158**	-0.002	-0.011	-0.432***	-0.452***
	(0.067)	(0.071)	(0.028)	(0.027)	(0.126)	(0.132)
Land Productivity (Std)	-0.119**	-0.136**	-0.055	-0.063	0.224*	0.190
	(0.053)	(0.054)	(0.084)	(0.080)	(0.124)	(0.132)
Island Dummy	-0.207	-0.228	-1.168***	-1.179***	0.000	0.000
	(0.202)	(0.168)	(0.196)	(0.170)	(.)	(.)
Temperature (Avg.)	-0.379	-0.623**	-0.239*	-0.355*	0.925	0.521
	(0.307)	(0.294)	(0.125)	(0.200)	(0.706)	(0.555)
Precipitation (Avg.)	-0.063	-0.111	-0.148*	-0.171	0.524*	0.456
	(0.099)	(0.074)	(0.085)	(0.108)	(0.297)	(0.287)
Temperature (Std)	0.145	-0.030	0.108	0.026	-0.140	-0.462
	(0.268)	(0.200)	(0.135)	(0.134)	(0.419)	(0.461)
Precipitation (Std)	-0.023	0.023	0.158***	0.180***	-0.151	-0.070
	(0.076)	(0.061)	(0.036)	(0.039)	(0.202)	(0.172)
Continent FE	✓	✓	✓	✓	✓	✓
First Stage F-Statistics		15.173		15.332		12.690
J-Test (p-value)		0.147		0.130		0.325
Adjusted R^2	0.209		0.124		0.290	
Observations	595	595	550	550	278	278

Note: OLS regressions with robust standard errors clustered at the language family level. The unit of analysis is an ethnic group, restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C14: The Association between Proxies of the State and Horse-Riding: Transport Mammals

	Centralization		Social Stratification		Community Size	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
Time since Cavalry Emergence	0.012	0.245*	0.058	0.185***	0.461	1.022**
	(0.082)	(0.119)	(0.093)	(0.057)	(0.300)	(0.403)
Transport Mammal Dummy	0.587***	0.428***	0.328	0.235	0.231	-0.292
	(0.120)	(0.118)	(0.193)	(0.164)	(0.437)	(0.541)
Log Dist. to Agricultural Frontier	-0.007	-0.023	-0.018	-0.026	-0.320**	-0.319*
	(0.069)	(0.066)	(0.035)	(0.032)	(0.151)	(0.156)
Latitude	-0.229	-0.364	-0.514	-0.578*	1.059	0.895
	(0.397)	(0.399)	(0.314)	(0.334)	(0.705)	(0.621)
Longitude	-0.598	-0.531	-0.600*	-0.543*	-0.711	-0.428
	(0.517)	(0.463)	(0.307)	(0.290)	(0.832)	(0.839)
Latitude × Longitude	-0.748**	-0.684**	-0.161	-0.143	-0.346	-0.281
	(0.307)	(0.285)	(0.223)	(0.207)	(0.727)	(0.612)
Elevation (Avg.)	-0.066	-0.149	-0.048	-0.093	0.198	0.042
	(0.082)	(0.089)	(0.061)	(0.059)	(0.291)	(0.264)
Land Productivity (Avg.)	0.188***	0.187***	0.047	0.047	0.629***	0.613***
	(0.061)	(0.056)	(0.036)	(0.034)	(0.170)	(0.157)
Elevation (Std)	-0.141**	-0.166**	0.007	-0.009	-0.334**	-0.413**
	(0.060)	(0.064)	(0.031)	(0.029)	(0.134)	(0.150)
Land Productivity (Std)	-0.108*	-0.129**	-0.050	-0.062	0.178	0.136
	(0.056)	(0.058)	(0.090)	(0.088)	(0.136)	(0.138)
Island Dummy	-0.261	-0.272*	-1.185***	-1.197***	0.000	0.000
	(0.199)	(0.158)	(0.183)	(0.156)	(.)	(.)
Temperature (Avg.)	-0.540*	-0.744***	-0.353**	-0.465***	0.496	0.081
	(0.275)	(0.234)	(0.131)	(0.155)	(0.682)	(0.514)
Precipitation (Avg.)	-0.049	-0.082	-0.159**	-0.178**	0.207	0.153
	(0.094)	(0.070)	(0.072)	(0.082)	(0.320)	(0.346)
Temperature (Std)	0.207	0.005	0.098	-0.003	-0.388	-0.858*
	(0.249)	(0.205)	(0.141)	(0.132)	(0.422)	(0.490)
Precipitation (Std)	-0.013	0.028	0.163***	0.189***	-0.039	0.070
	(0.082)	(0.059)	(0.042)	(0.035)	(0.195)	(0.176)
Continent FE	✓	✓	✓	✓	✓	✓
First Stage F-Statistics		25.030		26.830		18.995
J-Test (p-value)		0.199		0.217		0.423
Adjusted R^2	0.228		0.132		0.272	
Observations	595	595	550	550	278	278

Note: OLS regressions with robust standard errors clustered at the language family level. The unit of analysis is an ethnic group, restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances, island dummy and transport mammal dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C15: The Association between Proxies of the State and Horse-Riding: Trade Related Motifs

	Centralization		Social Stratification		Community Size	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	2SLS	OLS	2SLS	OLS	2SLS
Time since Cavalry Emergence	0.225*	0.330**	0.136*	0.221***	0.354	0.896**
	(0.116)	(0.139)	(0.068)	(0.039)	(0.288)	(0.372)
Log Trade Related Motifs	0.074	0.069	-0.030	-0.035	-0.070	-0.103
	(0.069)	(0.070)	(0.078)	(0.082)	(0.365)	(0.373)
Log Dist. to Agricultural Frontier	-0.022	-0.031	-0.030	-0.037	-0.323**	-0.340**
	(0.067)	(0.064)	(0.034)	(0.033)	(0.152)	(0.160)
Latitude	-0.252	-0.374	-0.468	-0.562	1.435**	0.928*
	(0.541)	(0.594)	(0.317)	(0.356)	(0.688)	(0.490)
Longitude	-0.401	-0.428	-0.411	-0.422	-0.233	-0.222
	(0.595)	(0.595)	(0.271)	(0.272)	(0.759)	(0.798)
Latitude × Longitude	-0.593	-0.515	-0.219	-0.164	-1.044	-0.675
	(0.383)	(0.410)	(0.225)	(0.233)	(0.730)	(0.697)
Elevation (Avg.)	-0.100	-0.138	-0.053	-0.085	0.120	-0.046
	(0.078)	(0.083)	(0.065)	(0.061)	(0.299)	(0.247)
Land Productivity (Avg.)	0.168**	0.169**	0.021	0.023	0.540***	0.536***
	(0.070)	(0.065)	(0.040)	(0.037)	(0.164)	(0.155)
Elevation (Std)	-0.087	-0.100*	0.018	0.007	-0.323**	-0.398**
	(0.054)	(0.055)	(0.032)	(0.032)	(0.138)	(0.144)
Land Productivity (Std)	-0.124**	-0.132**	-0.057	-0.064	0.176	0.145
	(0.057)	(0.056)	(0.089)	(0.087)	(0.136)	(0.139)
Island Dummy	-0.077	-0.091	-1.236***	-1.252***	0.000	0.000
	(0.177)	(0.156)	(0.190)	(0.160)	(.)	(.)
Temperature (Avg.)	-0.499*	-0.601***	-0.315**	-0.401***	0.279	-0.235
	(0.247)	(0.212)	(0.118)	(0.129)	(0.790)	(0.573)
Precipitation (Avg.)	-0.191*	-0.193**	-0.245***	-0.246***	-0.068	-0.032
	(0.099)	(0.089)	(0.085)	(0.088)	(0.313)	(0.325)
Temperature (Std)	0.037	-0.049	0.113	0.046	-0.399	-0.893
	(0.284)	(0.229)	(0.162)	(0.146)	(0.612)	(0.637)
Precipitation (Std)	0.031	0.048	0.221***	0.236***	0.193	0.275
	(0.083)	(0.074)	(0.058)	(0.050)	(0.236)	(0.208)
Continent FE	✓	✓	✓	✓	✓	✓
First Stage F-Statistics		16.969		17.803		13.751
J-Test (p-value)		0.142		0.181		0.496
Adjusted R^2	0.144		0.117		0.246	
Observations	562	562	516	516	253	253

Note: OLS regressions with robust standard errors clustered at the language family level. The unit of analysis is an ethnic group, restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Trade related motifs is taken logarithm after being divided by the total number of motifs in the ethnic group and added 0.01. Independent variables except for distances, island dummy, and log of trade related motifs are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C16: The Association between Proxies of the State and Horse-Riding: Dependence on Agriculture

	Centralization		Social Stratification		Community Size	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
Time since Cavalry Emergence	0.092	0.255**	0.086	0.173***	0.300	0.593*
	(0.073)	(0.107)	(0.051)	(0.053)	(0.191)	(0.298)
Log Dist. to Agricultural Frontier	0.002	-0.010	-0.022	-0.028	-0.232**	-0.240*
	(0.060)	(0.058)	(0.029)	(0.026)	(0.112)	(0.117)
Latitude	-0.153	-0.325	-0.385	-0.470	0.832	0.607
	(0.397)	(0.394)	(0.246)	(0.280)	(0.503)	(0.493)
Longitude	-0.276	-0.309	-0.318	-0.321	0.188	0.222
	(0.474)	(0.439)	(0.248)	(0.236)	(0.528)	(0.508)
Latitude × Longitude	-0.663*	-0.564*	-0.213	-0.170	-0.540	-0.400
	(0.341)	(0.323)	(0.212)	(0.202)	(0.533)	(0.502)
Elevation (Avg.)	-0.123*	-0.188**	-0.092	-0.127**	-0.131	-0.232
	(0.064)	(0.073)	(0.067)	(0.059)	(0.236)	(0.242)
Land Productivity (Avg.)	0.154***	0.157***	0.018	0.020	0.344**	0.344**
	(0.039)	(0.034)	(0.033)	(0.031)	(0.155)	(0.153)
Elevation (Std)	-0.112*	-0.136**	0.032	0.018	-0.159	-0.208
	(0.060)	(0.063)	(0.034)	(0.037)	(0.138)	(0.132)
Land Productivity (Std)	-0.148***	-0.160***	-0.077	-0.084	0.125	0.104
	(0.052)	(0.055)	(0.083)	(0.082)	(0.144)	(0.152)
Island Dummy	0.187	0.149	-1.063***	-1.086***	0.000	0.000
	(0.204)	(0.191)	(0.124)	(0.100)	(.)	(.)
Temperature (Avg.)	-0.563**	-0.745***	-0.373***	-0.472***	0.171	-0.148
	(0.228)	(0.204)	(0.118)	(0.138)	(0.400)	(0.378)
Precipitation (Avg.)	-0.218	-0.225*	-0.247***	-0.249***	-0.183	-0.163
	(0.135)	(0.119)	(0.051)	(0.054)	(0.207)	(0.205)
Temperature (Std)	0.131	-0.021	0.072	-0.004	-0.199	-0.481
	(0.269)	(0.235)	(0.151)	(0.136)	(0.329)	(0.334)
Precipitation (Std)	-0.017	0.016	0.152***	0.171***	-0.168	-0.105
	(0.079)	(0.065)	(0.034)	(0.029)	(0.167)	(0.143)
Dependence on Agriculture	✓	✓	✓	✓	✓	✓
Continent FE	✓	✓	✓	✓	✓	✓
First Stage F-Statistics		17.587		17.395		15.846
J-Test (p-value)		0.461		0.650		0.958
Adjusted R^2	0.275		0.154		0.417	
Observations	595	595	550	550	278	278

Note: OLS regressions with robust standard errors clustered at the language family level. The unit of analysis is an ethnic group, restricted on observations in the Old World because the data on horse-riding is not available in the New World. Dependence on agriculture ("v5") is a categorical variable taking 10 categories. Hence, the estimated coefficient and associated standard error are excluded from the table to save space. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C17: The Association between Proxies of the State and Horse-Riding: Agricultural Intensity

	Centralization		Social Stratification		Community Size	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	2SLS	OLS	2SLS	OLS	2SLS
Time since Cavalry Emergence	0.014	0.220**	0.053	0.152***	0.286	0.712*
	(0.080)	(0.093)	(0.063)	(0.048)	(0.259)	(0.360)
Log Dist. to Agricultural Frontier	-0.005	-0.019	-0.019	-0.026	-0.216**	-0.226**
	(0.056)	(0.052)	(0.024)	(0.022)	(0.095)	(0.101)
Latitude	-0.233	-0.435	-0.504**	-0.595**	0.007	-0.295
	(0.328)	(0.325)	(0.214)	(0.236)	(0.570)	(0.527)
Longitude	-0.451	-0.482	-0.477**	-0.480**	-1.865***	-1.795***
	(0.446)	(0.410)	(0.203)	(0.199)	(0.553)	(0.428)
Latitude × Longitude	-0.455	-0.332	-0.045	0.005	0.985*	1.178***
	(0.292)	(0.283)	(0.173)	(0.167)	(0.477)	(0.376)
Elevation (Avg.)	-0.159**	-0.237***	-0.102	-0.140**	-0.144	-0.290
	(0.059)	(0.078)	(0.060)	(0.058)	(0.269)	(0.286)
Land Productivity (Avg.)	0.109**	0.114***	-0.005	-0.001	0.429***	0.429***
	(0.042)	(0.038)	(0.036)	(0.034)	(0.124)	(0.121)
Elevation (Std)	-0.133*	-0.160**	0.010	-0.004	-0.298**	-0.358***
	(0.065)	(0.070)	(0.032)	(0.030)	(0.120)	(0.107)
Land Productivity (Std)	-0.093*	-0.111*	-0.043	-0.052	0.205	0.165
	(0.048)	(0.055)	(0.078)	(0.078)	(0.131)	(0.146)
Island Dummy	0.002	-0.043	-1.050***	-1.076***	0.000	0.000
	(0.156)	(0.127)	(0.187)	(0.162)	(.)	(.)
Temperature (Avg.)	-0.572**	-0.790***	-0.385***	-0.492***	-0.027	-0.469
	(0.217)	(0.181)	(0.091)	(0.102)	(0.571)	(0.593)
Precipitation (Avg.)	-0.062	-0.081	-0.157***	-0.165**	0.267	0.261
	(0.112)	(0.093)	(0.056)	(0.065)	(0.222)	(0.239)
Temperature (Std)	0.095	-0.085	0.042	-0.040	-0.648	-1.021**
	(0.249)	(0.194)	(0.154)	(0.140)	(0.469)	(0.452)
Precipitation (Std)	-0.056	-0.013	0.145***	0.166***	-0.192	-0.094
	(0.076)	(0.062)	(0.047)	(0.041)	(0.184)	(0.170)
Agricultural Intensity	✓	✓	✓	✓	✓	✓
Continent FE	✓	✓	✓	✓	✓	✓
First Stage F-Statistics		19.758		21.063		18.112
J-Test (p-value)		0.444		0.483		0.968
Adjusted R^2	0.310		0.183		0.461	
Observations	595	595	550	550	278	278

Note: OLS regressions with robust standard errors clustered at the language family level. The unit of analysis is an ethnic group, restricted on observations in the Old World because the data on horse-riding is not available in the New World. Agricultural intensity ("v28") is a categorical variable taking six categories. Hence, the estimated coefficient and associated standard error are excluded from the table to save space. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C18: The Association between Proxies of the State and Horse-Riding: Spread of Iron

	Centralization		Social Stratification		Community Size	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	2SLS	OLS	2SLS	OLS	2SLS
Time since Cavalry Emergence	0.042	0.358	0.030	0.200**	0.328	1.243**
	(0.106)	(0.212)	(0.105)	(0.085)	(0.340)	(0.516)
Time since Iron Emergence	0.059	-0.076	0.087	0.013	0.175	-0.284
	(0.061)	(0.101)	(0.062)	(0.075)	(0.177)	(0.287)
Log Dist. to Agricultural Frontier	-0.017	-0.028	-0.026	-0.031	-0.317**	-0.299*
	(0.069)	(0.065)	(0.035)	(0.032)	(0.150)	(0.153)
Latitude	0.181	-0.140	-0.261	-0.429	1.367**	0.630
	(0.404)	(0.515)	(0.239)	(0.312)	(0.619)	(0.481)
Longitude	0.254	-0.079	0.037	-0.139	-0.279	-1.133
	(0.527)	(0.631)	(0.196)	(0.245)	(0.789)	(0.816)
Latitude × Longitude	-1.604***	-1.168**	-0.840***	-0.623***	-1.095	0.142
	(0.396)	(0.535)	(0.155)	(0.216)	(0.788)	(1.082)
Elevation (Avg.)	-0.046	-0.134	-0.019	-0.066	0.226	0.013
	(0.091)	(0.101)	(0.060)	(0.063)	(0.358)	(0.327)
Land Productivity (Avg.)	0.181**	0.181***	0.050	0.049	0.667***	0.626***
	(0.068)	(0.058)	(0.042)	(0.038)	(0.172)	(0.155)
Elevation (Std)	-0.135*	-0.170**	0.004	-0.017	-0.324**	-0.448***
	(0.067)	(0.076)	(0.030)	(0.029)	(0.143)	(0.154)
Land Productivity (Std)	-0.126**	-0.140**	-0.060	-0.067	0.179	0.149
	(0.056)	(0.061)	(0.086)	(0.087)	(0.134)	(0.147)
Island Dummy	-0.069	-0.272	-1.019***	-1.143***	0.000	0.000
	(0.218)	(0.237)	(0.198)	(0.161)	(.)	(.)
Temperature (Avg.)	-0.437	-0.645**	-0.275**	-0.389**	0.534	0.042
	(0.292)	(0.250)	(0.115)	(0.144)	(0.812)	(0.633)
Precipitation (Avg.)	-0.061	-0.080	-0.161**	-0.164**	0.282	0.154
	(0.121)	(0.099)	(0.067)	(0.077)	(0.368)	(0.369)
Temperature (Std)	0.355	0.103	0.274*	0.153	-0.145	-0.903
	(0.369)	(0.285)	(0.158)	(0.151)	(0.609)	(0.648)
Precipitation (Std)	-0.014	0.019	0.175***	0.194***	-0.063	0.112
	(0.091)	(0.084)	(0.043)	(0.042)	(0.218)	(0.180)
Continent FE	✓	✓	✓	✓	✓	✓
First Stage F-Statistics		14.539		15.918		14.981
J-Test (p-value)		0.176		0.222		0.508
Adjusted R^2	0.210		0.122		0.243	
Observations	583	583	538	538	267	267

Note: OLS regressions with robust standard errors clustered at the language family level. The unit of analysis is an ethnic group, restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C19: The Association between Proxies of the State and Horse-Riding: Cereal Advantage

	Centralization		Social Stratification		Community Size	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
Time since Cavalry Emergence	0.240**	0.277*	0.144*	0.199***	0.335	0.779**
	(0.115)	(0.158)	(0.082)	(0.050)	(0.272)	(0.329)
Cereal Advantage	-0.079	-0.080	-0.181	-0.183	0.087	0.076
	(0.151)	(0.151)	(0.113)	(0.112)	(0.301)	(0.296)
Log Dist. to Agricultural Frontier	-0.023	-0.026	-0.017	-0.021	-0.318**	-0.332*
	(0.070)	(0.068)	(0.041)	(0.040)	(0.153)	(0.161)
Latitude	-0.340	-0.383	-0.573*	-0.635*	1.362*	0.955*
	(0.516)	(0.618)	(0.312)	(0.366)	(0.681)	(0.538)
Longitude	-0.504	-0.513	-0.537*	-0.543*	-0.332	-0.301
	(0.558)	(0.572)	(0.279)	(0.283)	(0.722)	(0.738)
Latitude × Longitude	-0.505	-0.477	-0.095	-0.059	-0.969	-0.670
	(0.358)	(0.423)	(0.203)	(0.228)	(0.692)	(0.663)
Elevation (Avg.)	-0.082	-0.096	-0.056	-0.077	0.162	0.019
	(0.079)	(0.091)	(0.077)	(0.079)	(0.290)	(0.243)
Land Productivity (Avg.)	0.211***	0.212***	0.122**	0.124**	0.512*	0.511*
	(0.057)	(0.057)	(0.058)	(0.058)	(0.260)	(0.266)
Elevation (Std)	-0.098*	-0.102*	0.012	0.005	-0.342**	-0.404***
	(0.051)	(0.052)	(0.041)	(0.041)	(0.135)	(0.134)
Land Productivity (Std)	-0.104*	-0.107*	-0.039	-0.043	0.192	0.168
	(0.058)	(0.055)	(0.090)	(0.086)	(0.153)	(0.154)
Island Dummy	-0.256	-0.259	-1.399***	-1.407***	0.000	0.000
	(0.187)	(0.179)	(0.170)	(0.153)	(.)	(.)
Temperature (Avg.)	-0.473*	-0.510**	-0.313**	-0.370**	0.351	-0.081
	(0.241)	(0.220)	(0.132)	(0.149)	(0.706)	(0.508)
Precipitation (Avg.)	-0.221**	-0.222**	-0.303**	-0.304**	0.003	0.032
	(0.096)	(0.093)	(0.110)	(0.115)	(0.257)	(0.261)
Temperature (Std)	0.047	0.016	0.060	0.015	-0.312	-0.729
	(0.264)	(0.211)	(0.131)	(0.132)	(0.569)	(0.597)
Precipitation (Std)	0.067	0.073	0.265***	0.275***	0.164	0.230
	(0.085)	(0.078)	(0.048)	(0.049)	(0.225)	(0.193)
Continent FE	✓	✓	✓	✓	✓	✓
First Stage F-Statistics		17.382		17.869		13.592
J-Test (p-value)		0.085		0.088		0.367
Adjusted R^2	0.136		0.137		0.249	
Observations	565	565	520	520	253	253

Note: OLS regressions with robust standard errors clustered at the language family level. The unit of analysis is an ethnic group, restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C20: Difference-In-Difference Analysis: Archaeological Sites, Horses and Cereals

	Ancient Site (Dummy)		Ancient Site (Count)		Ancient Settlement (Dummy)		Ancient Settlement (Count)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After Agr. × Log Dist. to Tell el-Ajjul	-0.035**	-0.095***	-0.146**	-0.439***	-0.030*	-0.083***	-0.110*	-0.313***
	(0.015)	(0.029)	(0.056)	(0.115)	(0.016)	(0.030)	(0.056)	(0.107)
After Agr. × Historical Horse Index	0.025**	0.051**	0.096*	0.222***	0.022*	0.045**	0.072*	0.161**
	(0.013)	(0.020)	(0.050)	(0.079)	(0.013)	(0.020)	(0.042)	(0.065)
After Agr. × Wild Relatives of Cereals (Dummy)	0.095***	0.062***	0.271***	0.163***	0.110***	0.077***	0.259***	0.171***
	(0.024)	(0.020)	(0.080)	(0.062)	(0.024)	(0.019)	(0.069)	(0.055)
After Agr. × Wild Relatives of Roots and Tubers (Dummy)	0.035*	0.009	0.128**	0.024	0.040**	0.014	0.114***	0.035
	(0.020)	(0.019)	(0.050)	(0.047)	(0.018)	(0.018)	(0.043)	(0.043)
After Agr. × Wild Relatives of Cereals, Roots and Tubers (Dummy)	0.015	-0.006	0.024	-0.062	0.026	0.005	0.038	-0.025
	(0.019)	(0.019)	(0.048)	(0.053)	(0.017)	(0.017)	(0.044)	(0.053)
After Agr. × Log Dist. to the Neolithic Origin		-0.007		-0.004		-0.006		-0.011
		(0.006)		(0.020)		(0.007)		(0.020)
After Agr. × Log Dist. to Trade Routes		0.001		0.023		-0.002		0.012
		(0.007)		(0.024)		(0.008)		(0.023)
After Agr. × Log Dist. to Volga-Don		0.098***		0.428***		0.096***		0.308***
		(0.030)		(0.116)		(0.029)		(0.103)
Cell FE	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.253	0.261	0.234	0.250	0.217	0.226	0.182	0.194
Observations	25552	25552	25552	25552	25552	25552	25552	25552

Note: The table reports difference-in-differences OLS regression. The unit of observation is the 1×1 virtual country (cell) before or after the Neolithic transition. Robust standard errors are clustered at the country level. All the independent variables except for the distances are standardized. *** p<0.01, ** p<0.05, * p<0.10.

Table C21: Difference-In-Difference Analysis: Archaeological Sites, Horses and Transport Mammals

	Ancient Site (Dummy)		Ancient Site (Count)		Ancient Settlement (Dummy)		Ancient Settlement (Count)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After Agr. × Log Dist. to Tell el-Ajjul	-0.043**	-0.110***	-0.173**	-0.474***	-0.035*	-0.103***	-0.131**	-0.354***
	(0.018)	(0.029)	(0.067)	(0.114)	(0.019)	(0.032)	(0.065)	(0.108)
After Agr. × Historical Horse Index	0.037**	0.064***	0.134**	0.257***	0.034**	0.060***	0.106**	0.196***
	(0.015)	(0.020)	(0.055)	(0.082)	(0.014)	(0.020)	(0.046)	(0.067)
After Agr. × Transport Mammal Dummy	-0.010	-0.025	-0.056	-0.066	-0.002	-0.021	-0.035	-0.060
	(0.025)	(0.017)	(0.081)	(0.057)	(0.025)	(0.017)	(0.068)	(0.052)
After Agr. × Log Dist. to the Neolithic Origin		-0.010		-0.012		-0.010		-0.020
		(0.007)		(0.023)		(0.008)		(0.023)
After Agr. × Log Dist. to Trade Routes		-0.006		0.006		-0.009		-0.004
		(0.008)		(0.028)		(0.010)		(0.027)
After Agr. × Log Dist. to Volga-Don		0.122***		0.485***		0.127***		0.375***
		(0.028)		(0.118)		(0.027)		(0.101)
Cell FE	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.240	0.258	0.222	0.246	0.194	0.219	0.166	0.189
Observations	25538	25538	25538	25538	25538	25538	25538	25538

Note: The table reports difference-in-differences OLS regression. The unit of observation is the 1×1 virtual country (cell) before or after the Neolithic transition. Robust standard errors are clustered at the country level. All the independent variables except for the distances and transport mamaml dummy are standardized. *** p<0.01, ** p<0.05, * p<0.10.

Table C22: The Association between Historical Battles and Horse-Riding: Cereals

	Log Dist. to the Closest Battle					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	OLS	OLS	2SLS
Time since Cavalry Emergence (900 CE)	-1.003***	-0.995***	-0.742***	-0.849***	-0.598***	-0.790**
	(0.128)	(0.129)	(0.128)	(0.166)	(0.116)	(0.364)
Wild Relatives of Cereals (Dummy)	-0.664**	-0.823***	-0.307	-0.296	-0.195	-0.205
	(0.274)	(0.286)	(0.244)	(0.193)	(0.136)	(0.138)
Log Dist. to Agricultural Frontier			0.198**	0.234***	0.195**	0.188**
			(0.095)	(0.085)	(0.092)	(0.092)
Latitude			-0.130	-0.073	-1.479	-1.413
			(0.469)	(0.441)	(1.110)	(1.215)
Longitude			1.046	0.509	-2.496*	-2.598*
			(0.680)	(0.675)	(1.315)	(1.319)
Latitude × Longitude			0.186	0.262	2.489**	2.565**
			(0.578)	(0.581)	(1.237)	(1.273)
Elevation (Avg.)			0.066	0.374**	0.338**	0.426
			(0.045)	(0.148)	(0.166)	(0.263)
Land Productivity (Avg.)			-0.243	-0.263	-0.115	-0.088
			(0.183)	(0.179)	(0.148)	(0.168)
Elevation (Std)			-0.286***	-0.277***	-0.155***	-0.136***
			(0.082)	(0.071)	(0.035)	(0.050)
Land Productivity (Std)			0.095*	0.073	0.020	0.027
			(0.057)	(0.050)	(0.028)	(0.025)
Island Dummy			0.934**	0.811	-0.070	-0.016
			(0.435)	(0.569)	(0.309)	(0.364)
Temperature (Avg.)				1.122**	0.607	0.947
				(0.484)	(0.585)	(0.945)
Precipitation (Avg.)				0.452***	0.379	0.370
				(0.155)	(0.253)	(0.266)
Temperature (Std)				1.271***	0.342	0.521
				(0.463)	(0.382)	(0.552)
Precipitation (Std)				0.067	-0.111	-0.130
				(0.121)	(0.128)	(0.131)
Continent FE		✓	✓	✓		
Country FE					✓	✓
First-F						13.863
J-Test (p-value)						0.155
Adjusted R^2	0.312	0.390	0.493	0.526	0.730	
Observations	7406	7406	7406	7406	7406	7406

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances, island dummy and dummy of the wild relatives of cereals are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C23: The Association between Historical Battles and Horse-Riding: Transport Mammals

	Log Dist. to the Closest Battle					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	OLS	OLS	2SLS
Time since Cavalry Emergence (900 CE)	-0.935***	-0.864***	-0.720***	-0.853***	-0.539***	-0.769*
	(0.137)	(0.143)	(0.142)	(0.165)	(0.108)	(0.403)
Transport Mammal Dummy	-0.190	-0.385*	-0.024	0.047	-0.285	-0.150
	(0.222)	(0.198)	(0.218)	(0.249)	(0.273)	(0.391)
Log Dist. to Agricultural Frontier			0.200**	0.237***	0.193**	0.187**
			(0.090)	(0.080)	(0.088)	(0.086)
Latitude			-0.051	-0.089	-1.577	-1.449
			(0.467)	(0.421)	(1.149)	(1.300)
Longitude			1.180*	0.588	-2.697**	-2.703**
			(0.663)	(0.619)	(1.202)	(1.270)
Latitude × Longitude			0.102	0.205	2.633**	2.641**
			(0.566)	(0.544)	(1.121)	(1.203)
Elevation (Avg.)			0.067	0.364**	0.362**	0.444*
			(0.046)	(0.171)	(0.160)	(0.242)
Land Productivity (Avg.)			-0.259	-0.273*	-0.103	-0.084
			(0.170)	(0.163)	(0.128)	(0.141)
Elevation (Std)			-0.293***	-0.278***	-0.158***	-0.136**
			(0.087)	(0.071)	(0.035)	(0.053)
Land Productivity (Std)			0.084	0.059	0.015	0.022
			(0.059)	(0.055)	(0.031)	(0.028)
Island Dummy			0.881*	0.779	-0.088	-0.030
			(0.458)	(0.569)	(0.276)	(0.344)
Temperature (Avg.)				1.084*	0.677	1.016
				(0.561)	(0.578)	(0.906)
Precipitation (Avg.)				0.479***	0.352*	0.365*
				(0.144)	(0.203)	(0.209)
Temperature (Std)				1.303***	0.380	0.561
				(0.464)	(0.399)	(0.551)
Precipitation (Std)				0.032	-0.113	-0.138
				(0.136)	(0.135)	(0.140)
Continent FE		✓	✓	✓		
Country FE					✓	✓
First-F						15.732
J-Test (p-value)						0.175
Adjusted R^2	0.296	0.369	0.489	0.522	0.730	
Observations	7395	7395	7395	7395	7395	7395

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances, island dummy and transport mammal dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C24: The Association between Historical Battles and Horse-Riding: Trade Routes

	Log Dist. to the Closest Battle							
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) OLS	(8) 2SLS
Time since Cavalry Emergence (900 CE)	-0.655***	-0.719***	-0.552***	-0.705***	-0.529***	-0.615	-0.581***	-0.985**
	(0.116)	(0.120)	(0.117)	(0.149)	(0.099)	(0.371)	(0.089)	(0.397)
Log Dist. to Trade Routes	0.307***	0.299***	0.217***	0.194***	0.127***	0.123***	0.132***	0.113***
	(0.045)	(0.043)	(0.030)	(0.030)	(0.027)	(0.039)	(0.027)	(0.037)
Log Dist. to Agricultural Frontier			0.202**	0.237***	0.196**	0.193**	0.201**	0.190**
			(0.087)	(0.078)	(0.091)	(0.095)	(0.089)	(0.083)
Latitude			-0.175	-0.124	-1.374	-1.347	-1.165	-0.964
			(0.460)	(0.452)	(1.058)	(1.131)	(1.016)	(1.210)
Longitude			1.038*	0.687	-2.605**	-2.643**	-2.008*	-1.940
			(0.564)	(0.583)	(1.226)	(1.204)	(1.140)	(1.175)
Latitude × Longitude			0.058	0.019	2.556**	2.585**	2.259**	2.271**
			(0.485)	(0.486)	(1.156)	(1.160)	(1.017)	(1.051)
Elevation (Avg.)			0.040	0.328**	0.346**	0.384	0.346**	0.516*
			(0.041)	(0.139)	(0.160)	(0.260)	(0.157)	(0.271)
Land Productivity (Avg.)			-0.203	-0.216	-0.109	-0.098	-0.126	-0.084
			(0.161)	(0.166)	(0.139)	(0.154)	(0.143)	(0.170)
Elevation (Std)			-0.237***	-0.225***	-0.123***	-0.116**	-0.117***	-0.083**
			(0.087)	(0.073)	(0.035)	(0.047)	(0.034)	(0.041)
Land Productivity (Std)			0.082	0.071	0.022	0.025	0.014	0.024
			(0.055)	(0.050)	(0.025)	(0.022)	(0.024)	(0.024)
Island Dummy			0.845	0.784	-0.046	-0.024	-0.011	0.099
			(0.516)	(0.600)	(0.280)	(0.331)	(0.280)	(0.378)
Temperature (Avg.)				1.073**	0.830	0.971	0.822	1.447
				(0.466)	(0.525)	(0.887)	(0.518)	(0.911)
Precipitation (Avg.)				0.317**	0.355	0.353	0.342	0.328
				(0.139)	(0.226)	(0.233)	(0.228)	(0.252)
Temperature (Std)				1.180***	0.340	0.420	0.331	0.680
				(0.441)	(0.345)	(0.516)	(0.316)	(0.511)
Precipitation (Std)				0.078	-0.110	-0.119	-0.083	-0.112
				(0.128)	(0.127)	(0.125)	(0.123)	(0.138)
Log Dist. to the Lower Volga-Don							-0.169***	-0.237**
							(0.047)	(0.093)
Continent FE		✓	✓	✓				
Country FE					✓	✓	✓	✓
First-F						13.572		20.530
J-Test (p-value)						0.170		0.199
Adjusted R^2	0.378	0.437	0.523	0.548	0.738		0.740	
Observations	7406	7406	7406	7406	7406	7406	7406	7406

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C25: The Association between Historical Battles and Horse-Riding: Time since Agricultural Transition

	Log Dist. to the Closest Battle							
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) OLS	(8) 2SLS
Time since Cavalry Emergence (900 CE)	-0.999***	-0.713***	-0.603***	-0.871***	-0.567***	-1.043	-0.604***	-1.326*
	(0.113)	(0.164)	(0.176)	(0.153)	(0.145)	(0.732)	(0.126)	(0.742)
yst900ce (standardized)	-0.703***	-0.847***	-0.770***	-0.693***	-0.448**	-0.360	-0.531**	-0.420
	(0.167)	(0.137)	(0.165)	(0.147)	(0.225)	(0.288)	(0.236)	(0.273)
Log Dist. to Agricultural Frontier			0.127**	0.171***	0.192***	0.171**	0.186***	0.154**
			(0.057)	(0.040)	(0.055)	(0.068)	(0.051)	(0.059)
Latitude			0.474	0.805	-0.169	-0.027	-0.069	0.168
			(0.570)	(0.578)	(1.019)	(1.284)	(0.983)	(1.297)
Longitude			1.277**	1.246***	1.259	0.714	1.700	0.991
			(0.494)	(0.465)	(1.189)	(1.526)	(1.292)	(1.475)
Latitude × Longitude			-1.542	-1.946**	-2.383*	-2.471*	-2.189*	-2.273
			(1.047)	(0.869)	(1.256)	(1.375)	(1.183)	(1.443)
Elevation (Avg.)			0.080	0.815***	0.763***	0.935***	0.747***	1.002***
			(0.053)	(0.158)	(0.142)	(0.340)	(0.148)	(0.355)
Land Productivity (Avg.)			-0.197**	-0.187**	-0.149*	-0.112	-0.157*	-0.104
			(0.089)	(0.084)	(0.078)	(0.114)	(0.080)	(0.115)
Elevation (Std)			-0.153**	-0.190***	-0.121***	-0.101**	-0.118***	-0.087**
			(0.060)	(0.045)	(0.032)	(0.045)	(0.032)	(0.042)
Land Productivity (Std)			0.049	0.085*	0.021	0.031	0.014	0.026
			(0.050)	(0.045)	(0.036)	(0.034)	(0.035)	(0.037)
Island Dummy			0.385	0.202	0.079	0.014	0.094	-0.000
			(0.305)	(0.199)	(0.261)	(0.273)	(0.256)	(0.294)
Temperature (Avg.)				2.562***	1.989***	2.609**	1.985***	2.914**
				(0.578)	(0.499)	(1.144)	(0.498)	(1.188)
Precipitation (Avg.)				0.550***	0.466**	0.492**	0.449**	0.483**
				(0.116)	(0.199)	(0.204)	(0.198)	(0.216)
Temperature (Std)				2.329***	1.705***	2.256***	1.497***	2.273***
				(0.467)	(0.431)	(0.764)	(0.398)	(0.756)
Precipitation (Std)				0.098	0.014	-0.004	0.016	-0.010
				(0.084)	(0.116)	(0.131)	(0.115)	(0.141)
Log Dist. to the Lower Volga-Don							-0.263***	-0.328***
							(0.085)	(0.113)
Continent FE		✓	✓	✓				
Country FE					✓	✓	✓	✓
First-F						17.922		17.789
J-Test (p-value)						0.572		0.403
Adjusted R^2	0.465	0.575	0.606	0.648	0.767		0.769	
Observations	3680	3680	3680	3680	3680	3680	3680	3680

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C26: The Association between Historical Battles and Horse-Riding: Spread of Iron

	Log Dist. to the Closest Battle					
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS
Time since Cavalry Emergence (900 CE)	-0.766***	-0.728***	-0.554***	-0.772***	-0.683***	-1.171***
	(0.141)	(0.116)	(0.146)	(0.122)	(0.082)	(0.417)
Time since Iron Emergence (900 CE)	-0.326**	-0.302**	-0.250**	-0.308***	-0.102*	0.036
	(0.139)	(0.133)	(0.110)	(0.097)	(0.056)	(0.122)
Log Dist. to Agricultural Frontier			0.204**	0.244***	0.187**	0.174**
			(0.094)	(0.075)	(0.090)	(0.085)
Latitude			-0.533	-0.509	-2.024*	-1.811
			(0.441)	(0.478)	(1.121)	(1.357)
Longitude			0.504	0.653	-1.722*	-1.159
			(0.700)	(0.589)	(1.029)	(0.995)
Latitude × Longitude			0.893	-1.081	0.064	-1.111
			(0.631)	(0.846)	(1.352)	(1.264)
Elevation (Avg.)			0.046	0.527***	0.413**	0.567**
			(0.049)	(0.138)	(0.185)	(0.273)
Land Productivity (Avg.)			-0.246	-0.228	-0.113	-0.082
			(0.172)	(0.160)	(0.127)	(0.147)
Elevation (Std)			-0.300***	-0.296***	-0.148***	-0.093*
			(0.086)	(0.068)	(0.027)	(0.047)
Land Productivity (Std)			0.077	0.078	0.016	0.027
			(0.057)	(0.049)	(0.030)	(0.031)
Island Dummy			0.998**	0.925*	-0.042	0.049
			(0.423)	(0.520)	(0.313)	(0.400)
Temperature (Avg.)				1.614***	0.799	1.354
				(0.464)	(0.684)	(0.956)
Precipitation (Avg.)				0.368***	0.443**	0.444**
				(0.134)	(0.205)	(0.214)
Temperature (Std)				2.359***	1.600***	2.249***
				(0.440)	(0.342)	(0.462)
Precipitation (Std)				0.239**	-0.043	-0.072
				(0.105)	(0.109)	(0.120)
Continent FE		✓	✓	✓		
Country FE					✓	✓
First-F						14.455
J-Test (p-value)						0.233
Adjusted R^2	0.284	0.368	0.476	0.530	0.730	
Observations	6886	6886	6886	6886	6886	6886

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country (cell), restricted on observations in the Old World because the data on horse-riding is not available in the New World. Distances are taken logarithm. Independent variables except for distances and island dummy are standardized. Continent fixed effects are Africa, Asia, and Europe. *** p<0.01, ** p<0.05, * p<0.10.

Table C27: The Number of Historical Battles and the Historical Horse Index: The Natural Experiment Associated with the Columbian Exchange

	Number of Historical Battles							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Historical Horse Index	0.0114***	0.0114***	0.0103**	0.0198***	0.0212**	0.0216**	0.0216**	0.0219***
	(0.0041)	(0.0041)	(0.0039)	(0.0070)	(0.0085)	(0.0083)	(0.0083)	(0.0077)
Land Productivity		0.0003	0.0004	0.0006**	0.0007***	0.0007***	0.0007***	0.0006**
		(0.0015)	(0.0004)	(0.0003)	(0.0002)	(0.0003)	(0.0003)	(0.0002)
Cell FE			✓	✓	✓	✓	✓	✓
Time FE				✓	✓	✓	✓	✓
Time × Geolocation					✓	✓	✓	✓
Time × Elevation						✓	✓	✓
Time × Island							✓	✓
Time × Climate								✓
Adjusted R^2	0.007	0.007	0.000	0.012	0.012	0.012	0.012	0.013
Observations	96529	96529	96529	96529	96529	96529	96529	96529

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country in the Americas every 50 years, spanning since 1000 CE to 2000 CE. Geolocation are latitude, longitude and latitude × longitude. Elevation are average of elevation and standard deviation of elevation. Island is a dummy variable. Climate are average temperature, average precipitation, standard deviation of temperature and standard deviation of precipitation. All the variables except for the dependent variable are standardized. *** p<0.01, ** p<0.05, * p<0.10.

Table C28: The Number of Historical Battles and the Historical Horse Index: Pre-Trend

	Presence of Historical Battles							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Change HHI × Year	0.0003	0.0003	0.0006	0.0022	-0.0008	-0.0011	-0.0011	-0.0007
	(0.0003)	(0.0003)	(0.0010)	(0.0021)	(0.0030)	(0.0029)	(0.0029)	(0.0039)
Change Land Suit. × Year		0.0001	0.0007	0.0009	0.0005	0.0005	0.0005	0.0004
		(0.0001)	(0.0010)	(0.0012)	(0.0012)	(0.0013)	(0.0013)	(0.0013)
Cell FE			✓	✓	✓	✓	✓	✓
Time FE				✓	✓	✓	✓	✓
Time × Geolocation					✓	✓	✓	✓
Time × Elevation						✓	✓	✓
Time × Island							✓	✓
Time × Climate								✓
Observations	39600	39600	39600	39600	39600	39600	39600	39600

Note: OLS regressions with robust standard errors clustered at the country level. The unit of analysis is a 1×1 virtual country in the Americas every 50 years, spanning since 1000 CE to 2000 CE. Geolocation are latitude, longitude and latitude × longitude. Elevation are average of elevation and standard deviation of elevation. Island is a dummy variable. Climate are average temperature, average precipitation, standard deviation of temperature and standard deviation of precipitation. All the variables except for the dependent variable are standardized. *** p<0.01, ** p<0.05, * p<0.10.

Appendix D. Variable Definitions

D.1. Outcome Variables

- **Ancient City**: The presence, number, and distance from the closest ancient city are calculated by using the data on the location of ancient cities reported by Degroff (2009) and Reba et al. (2016). For the latter source, 500 BCE and 450 CE are used as reference points.

- **State Index**: This variable is a flow index reflecting three dimensions of state presence. These dimensions are hierarchy, autonomy and territory. The variable is taken from Borcan et al. (2018).

- **State History**: This variable is an accumulative score of *State Index*, discounting the distant past at the 1% discount factor. The variable is taken from Borcan et al. (2018).

- **Jurisdictional Hierarchy Beyond Local Community**: This variable is "v33" in the *Ethnographic Atlas*.

- **Social Stratification**: This variable is based on "v66" in the *Ethnographic Atlas*. The social stratification variable is grouped into the following categories. It takes on the value 0 when the original variable indicates "Absence among freemen," takes 1 when it indicates "Wealth distinctions" or "Elite," and takes 2 when it indicates "Dual" or "Complex."

- **Mean Size of Local Communities**: This variable is "v31" in the *Ethnographic Atlas*.

- **Radiocarbon-Dated Prehistorical archaeological sites**: These variables are taken from Mayshar et al. (2022), who georeference the number of pre-Neolithic and post-Neolithic sites from Whitehouse and Whitehouse (1975).

- **Historical Battle**: The occurrence, number, and distance from the closest battle are calculated by using information on geolocation and year reporeted by the *World Historical Battles Database* (Kitamura, 2021). These variables are calculated, referring to 2000 BCE, 500 BCE, 0 CE, 500 CE, 1000 CE, and 1500 CE.

- **Population Count**: This variable is average of population count within a target unit in 2000 BCE. The original raster file is taken from the HYDE 3.2.1 (Klein Goldewijk et al., 2017).

- **Population Density**: This variable is average of population density within a target unit in 2000 BCE. The original raster file is taken from the HYDE 3.2.1 (Klein Goldewijk et al., 2017).
- **Urbanization**: This variable is average of urban population count within a target unit. The original raster file is taken from the HYDE 3.2.1 (Klein Goldewijk et al., 2017).

D.2. Independent Variables

- **Time Since Cavalry Emergence**: This variable is the years elapsed since cavalry emergence. The original map is from Turchin et al. (2021), which indicates the year when cavalry emerged. After translating the original values into years since cavalry emergence, these values are aggregated at the arbitrary scale.
- **Distance from Tell el-Ajjul**: The distance from tell el-Ajjul is calculated as average across 1×1 raster grid cells within a target observation. The geolocation of Tell el-Ajjul is taken from the *Ancient Location*.
- **Historical Horse Index**: This variable is a product of an indicator of the presence of native horses and climatic suitability index for wild-living horse populations. The dummy variable is created by using a present natural map of *Equus ferus* reported by the PHYLACINE (Faurby et al., 2018). Climatic suitability index is taken from Naundrup and Svenning (2015). Value 0 is assigned in cells in the Americas.

D.3. Control Variables

- **Latitude/Longitude**: For a country, it is the absolute value of the latitude of that country's approximate geodesic centroid, as reported by the CIA's World Factbook. For an ethnic group, it is the value of the latitude, as reported by the *Ethnographic Atlas* or Binford (2019).
- **Elevation**: Average and standard deviation of elevation within an area. The data is taken from the Atlas of Bioshpere.
- **Land Productivity for Agriculture**: Average and standard deviation of land suitability for agriculture within an area. The data is taken from Ramankutty et al. (2002). For panel analyses, change in caloric suitability before and after 1500 CE is calculated using the data from Galor and Özak (2016).

- **Island Dummy**: It is an indicator for whether or not a country shares a land border with any other country, as reported by the CIA's World Factbook online.

- **Temperature**: Average and standard deviation of temperature within an area over the period 1901-2012 are calculated based on the Climate Research Unit (CRU).

- **Precipitation**: Average and standard deviation of precipitation within an area over the period 1901-2012 are calculated based on the Climate Research Unit (CRU).

- **Time Elapsed since the Neolithic Transition**: The variable at the country level is the number of years elapsed as of the year 2,000 since the majority of the population residing within a country's modern national borders began practicing sedentary agriculture as the primary mode of subsistence. The data is taken from Borcan et al. (2018). For the virtual country level analysis, it is taken from Currie et al. (2020).

- **Cereal Advantage**: It is a difference of the optimal yield from cereal crops and the optimal yield from roots and tubers. Firstly, a raster file is created and then values are aggregated at the arbitrary level. The data is taken from the FAO GAEZ version 4.

- **Wild Relatives of Cereals**: It is the number of wild plants that are genetically related to cultivated cereal crops. The original data is taken from the Crop Wild Relatives Project (CWRP, 2021) and the measure is calculated by using ArcGIS Pro.

- **Dummy of Transport Mammals**: It is a dummy variable that takes 1 if there is a transport mammal, and 0 otherwise. Transport mammals are *Bos gaurus*, *Bos javanicus*, *Bos mutus*, *Bos primigenius*, *Bubalus arnee*, *Camelus dromedarius*, *Camelus ferus*, *Equus africanus*, *Equus ferus*, and *Lama guanicoe*, which are identified by Link (2021). Destributions of these mammals except for *Equus ferus* are taken from the PHYLACINE (Faurby et al., 2018). For *Equus ferus*, it is taken from Naundrup and Svenning (2015).

- **Distance from the Lower Volga-Don Region**: The distance from the lower-Volga Don region is calculated as average across 1×1 raster grid cells within a target observation. The map of the lower Volga-Don region is taken from Librado et al. (2021).

- **Distances from Other Major Cities**: The distances from Eridu, Susa, Knossos and Erlingang are calculated as average across 1×1 raster grid cells within a target observation. The geolocation of these places are taken from the *Ancient Location*.

- **Distance from the Closest Agricultural Origin**: The distance from the closest agriculutural origin is calculated as average across 1×1 raster grid cells within a target observation. The map of seven independent agricultural centres is taken from Purugganan and Fuller (2009).

- **Distances from Other Major Cities**: The distances from Eridu, Susa, Knossos and Erlingang are calculated as average across 1×1 raster grid cells within a target observation. The geolocation of these places are taken from the *Ancient Location*.
- **Dependence on Agriculture**: This variable is "v5" in the *Ethnographic Atlas*.
- **Intensity of Agriculture**: This variable is "v28" in the *Ethnographic Atlas*.
- **Motifs Related to Trade**: This variable is taken from Michalopoulos and Xue (2021). The variable is taken log after it is divided by the total number of motifs and added 0.01.