The Impact of Horse Adoption on Native American Nations*

David Cuberes, Rob Gillezeau, Motohiro Kumagai, and Sadia Mansoor

March 2024

Abstract

In the 17^{th} and 18^{th} centuries, the horse spread across North America, becoming a crucial part in Indigenous societies, militaries, and economies. In this work, we study both the determinants of the spread of the horse among Native American nations and how the animal's spread shaped the population and other socioeconomic aspects of these nations from the 17^{th} century through to the present. In order to address the possible endogeneity of horse adoption, we use a least cost path strategy that accounts for geographic constraints impacting the natural spread of the animal. Our results indicate that the horse was adopted earlier in nations that had land less suitable for agriculture. Earlier access to the horse resulted in substantially greater population growth in the short and medium run, but not in higher income or height in the long run. Nations with earlier access to the horse depended more on hunting and ceded land to colonizing nations at later dates. We interpret these results through the lens of a Malthusian model with resource depletion to reflect the disastrous effect that the near extinction of the bison had on Native American nations.

*Cuberes: Department of Economics, Clark University, E-mail: dcuberes@clarku.edu. Gillezeau: Department of Economics, University of Toronto, E-mail: rob.gillezeau@rotman,utoronto.ca. Kumagai: Department of Economics, Hitotsubashi University, E-mail: motohiro.kumagai@r.hit − u.ac.jp . Mansoor: Department of Economics, Clark University, E-mail: smansoor@clarku.edu We thank Fernando Arteaga, Javier Birchenall, John Brown, Ann Carlos, Metin Cosgel, Pedro Dalbó, Klaus Desmet, Anita Fabos, Oded Galor, Donn Feir, Jesús Fernández-Villaverde, Kris Inwood, Fabio Mariani, Melissa Miller, Stelios Michalopoulos, Ömer Özak, Louis Putterman, Laura Robinson, Assaf Sharif, Tianyi Wang, David Weil, Junfu Zhang, and seminar participants at Newcastle University Business School, University of Birmingham, Universidad Pablo Olavide, Brown University, University of Toronto, and Clark University. We also received very valuable comments at the Canadian Economics Association Meetings (Winnipeg), the Canadian Network for Economic History Conference (Toronto), the Kyoto Summer Workshop on Applied Economics, and the Southern Economic Association Meetings (New Orleans).

Keywords: Native Americans, horses, colonization, Malthusian model, conflicts and wars

JEL Codes: J11, J15, N11, N31, N51, O13, O51, Q32, Q55

PRELIMINARY, COMMENTS WELCOME. PLEASE, DO NOT CIRCULATE

1 Introduction

Throughout the United States and the Americas more broadly, Indigenous peoples and nations have worse socioeconomic outcomes than non-Indigenous communities (Akee et al., 2019; Feir et al., 2023, Feir et al. 2024,). This has been linked to a variety of institutional features, such as forced coexistence and land fractionation, and resource depletion, such as the extermination of the bison. At the time of European arrival, Indigenous economies were in many cases complex, wealthy, and heterogeneous (Carlos 2023). However, the path of Indigenous nations and their economies was shaped by the rapid spread of European diseases followed by increasing European encroachment into the interior of the continent Importantly, the speed and extent of recovery from early waves of disease spread likely shaped nations' interactions with European colonizers during this period of encroachment.

In this work, we seek to understand how a potentially mitigating factor may have shifted the economic and demographic recovery of these nations. In particular, we consider the diffusion of the horse across North America after the re-introduction of the animal by Spanish conquistadors in Central America. Starting in the late 16th century through to the late 18th century, the horse spread across the central and western parts of the United States, transforming Indigenous economies and militaries. What suggests that there might have been a meaningful impact of horses on Native American nations? There are at least two reasons to suspect that the horse was known to be valuable to these peoples. First, horses transformed the nature of the bison hunt (Feir et al., 2023) and trade (Featherling, 2019). As noted by Carlos (2023), it was a key feature of sophisticated and specialized Indigenous economies engaged in substantial trade relationships.[1] Horses additionally granted significant military advantages in combat on the open plains and over wide distances (Kumagai, 2022).

We are able to investigate the role of the horse empirically by digitizing historical mappings of the spread of the horse in North America using Wissler (1914) and Haines (1938), building a new dataset of historical population at the Indigenous nation level between the

[1]Eivdence from heights suggest that Indigenous peoples in the Great Plains were arguably the "tallest people in the world" (Prince and Steckel, 2001; 2003) in the 18^{th} century, suggesting high levels of well-being and income.

16^{th} and 20^{th} centuries. Using this mapping, we are able to predict the primary determinants of horse adoption. We then employ a range of empirical methods to determine the causal impact of access to the horse on Indigenous population levels, heights, methods of food production, and land cessions to the United States government.In particular, we are able to take advantage of least cost pathing for the natural diffusion of the animal to help achieve causation in some of these approaches.

We find that the horse was selectively adopted earlier in regions with less suitable land for agricultural activity, pointing to continued specialization or re-specialization towards hunting activities for food production. With respect to societal outcomes, we find that earlier access to the horse led to large increases in the population of Indigenous nations. Consistent with this result, we see that these same nations are able to resist colonial pressures for longer, ceding their lands to the United States government at a later date. However, these gains do not necessarily appear to be long lived as we see no impact of earlier access to the horse on either heights in the 19th century or later incomes. Thus, it appears that while the horse provided an important short to medium run boost to economic and military power, these gains are largest exhausted by the latter half of the 19th century.

The paper is organized as follows. Section 2 provides some brief historical background about the reintroduction of the horse to North America. In Section 3, we present a simple Malthusian model with resource depletion to help interpret our results. Section 4 discusses the data constructed and used in the study. The empirical methodology is developed in Section 5 and we present our results in Section 6. Finally, Section 7 concludes the paper.

2 Historical background

The horse originated in the Americas, migrating from North America to Eurasia via land bridges, traversing areas now submerged beneath the Bering Strait. However, the horse in North America itself was driven to extinction thousands of years ago. While Eurasia boasted expansive grasslands conducive to their survival, North America underwent significant environmental changes, with diminishing grasslands and increased human activity. This, coupled with predation pressure and habitat alterations, posed challenges to horse populations, contributing to their extinction on the North American continent.[2] In contrast, the vast grasslands of Asia and Europe provided ideal conditions for their flourishing and eventual domestication by human societies.

[2]Evidence suggests that this early iteration of the horse was primarily hunted by humans right than domesticated (Waters et al 2015).

The horse was reintroduced to the Americas by European colonizers in the late 15th and early 16th centuries and spread rapidly (Delsol et al 2022). Work by Taylor et al (2023) finds that North American horses show genetic commonalities primarily with Iberian and to a lesser extent British sources. There is no genetic evidence of horse diffusion occurring through earlier Viking travel to North America.

After their initial re-introduction, the vast majority of the horses were of Iberian origin[3]. While the traditional view was that the horse arrived to the Indigenous nations of the Great Plains in the 18th century, a rapidly growing body of literature suggestions that the horse arrived well in advance of this period and that this estimate was likely grounded in the European observational bias.[4] Starting early in the 16th century, it appears that the animal rapidly began to spread primarily through Indigenous trading routes across the Americas rather than through trade with Europeans (Taylor et al., 2023).

The horse rapidly became a central element of life in many Indigenous nations taking on practical roles in combat and hunting, travel and as a method of wealth accumulation(Fleharty, E. D., 1995). In the Great Plains in particular, it had a dramatic impact as a transportation technology, although this certainly applied more broadly as well (Hämäläinen, P., 2003). Simultaneously the animal played an important cultural and spiritual role in many of these same societies. Importantly, the earlier arrival of the horse than previously thought to the Great Plains meant that these societies could have fully incorporated the animal by the time pressure from American westward expansion reached their borders. While many traditional histories emphasized the transmission of horse related human capital from Europeans to Indigenous peoples, this is, in fact, inconsistent with the archaeological record with these skills being developed internally or passed between Indigenous nations.

This early spread of animal and the presence of many well-established Indigenous, horse-dependent nations by the time of European encroachment into the Great Plains alleviates a possible threat to our identification strategy as it points to the exogeneity of horse access relative to the path of colonization and European trade networks.

[3]Of particular relevance to our work, the Spanish horse (Iberian mustang) is well-suited to the arid and semiarid plains and mesas of the American West.

[4]The past emphasis placed on the Pueblo Revolt and the resulting "Great Horse Disperal" is consistent with this point.In the American wilderness tale, the initial dissemination of Spanish horses was constrained, with Indigenous peoples lacking the expertise or inclination for their domestication. Against a backdrop of scarcity and the formidable cost of taming the animal in the 16th century, Native Americans, grappled with the challenges of horse integration in daily life (McKnight, T. L., 1959).

3 A Malthusian Model with Resource Depletion

To give a sense of how the introduction of the horse may have altered the economic path of Indigenous nations, we build a model that extends prior work from Brander and Taylor (1998). In the model there is a renewable natural resource (the bison or other hunted animals in our case) and there is a population (Indigenous peoples in our case) that exploits this resource. A key distinction between our model and the Brander-Taylor model is that we add the a technological shock, namely the adoption of the horse. This new technology makes it much easier for Native Americans to exploit natural resources – hunting the bison in our context.

3.1 Renewable Resource Dynamics

The resource stock at time t is $S(t)$. It evolves according to:

$$\frac{dS}{dt} = G(S) - H \tag{1}$$

where $G(S)$ is the natural growth rate of the bison and H is the "harvest" of bison. We assume a logistic function for $G(S)$:

$$G(S) = rS\left(1 - \frac{S}{K}\right) \tag{2}$$

where r is the intrinsic growth of the natural resource and K is the carrying capacity, i.e. the maximum possible size of the natural resource stock. When $S = K$, further growth in the natural resource cannot occur.

The economy has two goods: H and M. H is the harvest of the bison (the natural resource) and M is a composite good (mostly agriculture in our case). M is the numeraire and so its price is set to one. The other factor of production is labor, L. The good M is produced under constant returns to scale so that 1 unit of labor produces one unit of M. Therefore the wage must be equal to one in the agricultural sector.

Hunting is carried out according to the Schaefer harvesting production function:

$$H^p = \alpha S L_H \tag{3}$$

where L_H is labor used in hunting. This implies that the unit labor requirement in hunting is:

$$a_{LH}(S) = \frac{L^H}{H^P} = \frac{1}{\alpha S} \tag{4}$$

There is free entry in the hunting sector. Open access implies that the rental cost for using S is the one that sets price equal to its marginal cost:

$$p = wa_{LH} = \frac{w}{\alpha S} \tag{5}$$

Consumers have the following preferences:

$$u = h^{\beta} m^{1-\beta} \tag{6}$$

where h and m are individual consumption of the resource good and agriculture, and $\beta \in (0, 1)$. Individuals maximize utility subject to their budget constraint $ph + m = w$. Since total domestic demand is L times individual demand we get:

$$H^D = \frac{w\beta L}{p} \tag{7}$$

$$M^D = w(1-\beta)L \tag{8}$$

At a given point in time, the resource stock and population are fixed so:

$$H^P a_{LH}(S) + M = L \tag{9}$$

This linear structure is known as a Ricardian production structure. Combining (5) and (7) we get the temporary equilibrium resource harvest, H:

$$H = \alpha\beta LS \tag{10}$$

The Ricardian temporary equilibrium output of M is $M = (1-\beta)L$, implying that agriculture will always be produced and so $w = 1$.

At a Ricardian temporary equilibrium, the harvest will not necessarily equal the underlying biological growth rate of the resource. If, for example, temporary equilibrium harvest, H, exceeds biological growth G, then the stock diminishes. Substituting (2) and (10) into (1) yields the following expression for the evolution of the resource stock:

$$\frac{dS}{dt} = rS(1 - \frac{S}{K}) - \alpha\beta LS \tag{11}$$

This equation tells us that if the resource stock falls, then labor productivity in the re-

source sector falls, the Ricardian production possibility frontier shifts in, and this establishes a new temporary equilibrium with a lower harvest.

Figure 1 illustrates a typical steady state (i.e., where $\frac{dS}{dt} = 0$) using the harvest function (10) and the resource growth function (2). As shown in the figure, stock S^* implies that harvest $H = \alpha\beta LS^*$ just matches resource growth $G(S)$ at $S = S^*$. Thus, S^* is a steady state if the actual population level is L.

3.2 Malthusian Population Dynamics

Our discussion of Figure 1 so far implicitly treats population as fixed at size L, allowing us to focus purely on resource dynamics. We now consider population dynamics. We assume an underlying proportional birth rate, b, and an underlying proportional death rate, d. Thus, the base rate of population increase is $(b - d)$. We assume $b - d < 0$, implying that without any bison stock or arable soil the Native American population would eventually disappear.

$$\frac{dL}{dt} = L[b - d + F] \tag{12}$$

where $F = \phi\frac{H}{L}$ is the fertility function and ψ is a positive constant. Thus, higher per capita consumption of the resource good leads to higher population growth. It is in this sense that population dynamics are "Malthusian." Noting from (10) that $\frac{H}{L} = \alpha\beta S$ equation (12) can be rewritten as:

$$\frac{dL}{dt} = L(b - d + \phi\alpha\beta S) \tag{13}$$

3.3 Population and Resource Interactions

Equations (11) and (13) form a two-equation system of differential equations characterizing the evolution of the Ricardo-Malthus model. Human population, L, is the "predator" and the resource stock, S, is the "prey." We concentrate first on steady-state analysis, then turn to the system's dynamic behavior.

3.3.1 Steady-State Analysis

The system given by (11) and (13) will have a steady state if $\frac{dL}{dt}$ and $\frac{dS}{dt}$ are simultaneously equal to zero. There are three solutions, as expressed in Proposition 1.

PROPOSITION 1: The Ricardo-Malthus model exhibits three steady states. Steady state l is a corner solution at $(L = 0, S = 0)$. Steady state 2 is a corner solution at

$(L = 0, S = K)$. Steady state 3 is an interior solution at

$$L = [\frac{r}{\alpha\beta}][1 - \frac{(d-b)}{(\phi\alpha\beta K}] \tag{14}$$

and

$$S = \frac{d-b}{\phi\alpha\beta}$$

PROPOSITION 2: The steady-state resource stock:

i. rises if the mortality rate rises, the birth rate falls, or fertility responsiveness falls;

ii. falls if there is technological improvement in hunting; and

iii. is unaffected by changes in the intrinsic resource regeneration rate, r, or carrying capacity, K.

The adoption of the horse can be interpreted as permanent technological improvement in hunting. The parameter that captures this productivity parameter in the model is α. Figure 2 displays the effect of horse adoption on the steady-state stock of horses. A second effect of the technological shock is that, given the labor market clearing condition, now Indigenous nations would, all else equal, spend more time hunting than previously. This is consistent with one of our empirical results to be discussed in the following section. We would get the same qualitative results if r rises: the growth curve will shift up as shown by the dashed line in Figure 1. At stock S^*, the harvest is then less than resource growth and the resource rebuilds. Per capita consumption of the resource good rises, causing population growth, and the harvest function rotates upwards (as L rises). Since nothing has altered the demographic steady state in Panel B, the resource stock must return to S^*, but with a higher population and unchanged per capita real income.

The model also generates a hump-shaped relationship between horse adoption and hunting. In the time allocation constraint (9), the time allocated to hunting is $H^P a_L H(S)$. Because horse adoption reduces S (the stock of bison), this term decreases as fewer bison are available. This implies that, as time passes after horse adoption, the economy spends relatively less time in hunting as a result.[5]

[5]The second exogenous shock in our context is the arrival of colonizers. While these colonizers were initially uninterested in the animal. This shifted after innovations in tanning technology, driving interest

The next proposition shows the effects of changes in different parameters on the steady state level of population.

PROPOSITION 3: The steady-state population level:

i. rises equiproportionately with an increase in the intrinsic rate of resource growth, r;

ii. falls when the hunting technology improves if $S < \frac{K}{2}$ and rises if $S > \frac{K}{2}$;

iii. falls when the taste for the resource good rises if $S < \frac{K}{2}$ and rises if $S > \frac{K}{2}$

iv. rises if the carrying capacity of the environment rises.

It follows from (10) and (14) that per capita steady-state resource consumption, h, is $\frac{d-b}{\phi}$. Per capita consumption of the other good is $1-\beta$. Thus steady-state resource consumption is determined by demographics, and rises if the birth rate falls, mortality increases, or if fertility responsiveness falls.

4 Data

While there exist increasingly rich datasets covering Indigenous nations and peoples starting in the 19th century (Dippel 2014, Feir et al 2023, Feir at al 2024), there remain extremely limited pre-existing data sources for these nations in earlier centuries. To address this data gap, we hand collect population data for 150 Indigenous nations within the present-day boundaries of the United States. Due to limited data on historical population of native nations, We are able to build a particularly rich, long duration panel dataset encompassing 82 Native American nations only, drawing on records from the Smithsonian Handbook of the North American Indian (Volume 3-15), the Handbook of the Indians of California (Kroeber 1925), and the Statistical Record of Native North Americans (Reddy, M.A. 1995).A majority of our observations are sourced from volumes of the Handbook of North American Indians. An important characteristic of the population data is that it comes at very irregular time intervals. For example, for the Mandan nation, the first five observations are 1600, 1645, 1660, 1679, and 1681. Another relevant characteristic of these data is that, in some cases, there are wide fluctuations, pointing to significant measurement error. This is unsurprising

from Europeans in the bison mostly for their fur. A simple way to model this is therefore via an exogenous increase in consumer's preferences for consumption of bison, i.e increase in the parameter β in the consumer's preferences. A second, and more realistic, way to do that is to think of the arrival of the colonizers as a shock to the population size via an increase in the immigration term i. Both ways of introducing the arrival of the colonizers would give the same qualitative results.

given the limited and non-systematic approach to the documentation of population by European observers. Our approach is to include all years, with a very limited number of cases when we drop observations that are clearly extreme outliers.[6]

The interaction with European colonization introduced further hurdles to population data gathering. Diseases such as smallpox, measles, and influenza devastated Native American populations, often spreading faster than European expeditions due to intra-tribal interactions and trade. Forced displacement, warfare, and enslavement also contributed to population decline and demographic shifts, challenging accurate population estimations before and after such events. Moreover, colonial records frequently underrepresented or inaccurately counted Native American populations due to cultural and linguistic differences, focusing more on taxation, labor, or military purposes than comprehensive demographic data. Cultural disparities in counting methods, loss of traditional knowledge, and limited archaeological evidence further complicate efforts to reconstruct accurate population data. Native American societies often employed different population counting methods based on kinship or community ties, diverging from European approaches. The disruption caused by colonization and forced assimilation efforts led to the loss of traditional practices, including population record-keeping, exacerbating the challenge. While archaeology provides insights into past populations, its limitations in providing a complete demographic picture add another layer of complexity. Overall, the scarcity of historical data on North Native American populations in the 1500s and 1600s stems from a combination of factors, including limited contact with Europeans, sparse written records, cultural differences, and the devastating impacts of colonization, collectively posing significant challenges to reconstructing accurate population data. Figure 3 displays the average population of Native Americans over time. The figure displays a clear downward trend but a significant recovery at the beginning of the 18^{th} century. A sharp decline is also apparent after around the mid-19^{th} century.

Spatial and timing data on the adoption of the horse are drawn from Wissler (1914), Haines (1938), Jacobsen and Eighmy (1980). Wissler (1914) who provides the approximate years when 25 nations from the Plains adopted horses. His inference relies on the broad studies in anthropology and ethnography and he reports the earliest date the tribes possessed horses. Haines (1938) provides a map of how horses spread in western America after about one hundred horses were abandoned by the Spanish Conquistadors De Soto and Coronado near modern-day Mexico. Haines uses ethnographical and anthropological data to trace the approximate lines of distribution and dates when the various nations acquired horses. We

[6]In general, our results are robust to the inclusion of these observations.

digitize Haines' map in Figure 6.[7] We also gather information on horse adoption for some nations from the Smithsonian Encyclopedia of Native Americans. For the nations for which we do not observe their adoption data, we use the Inverse Distance Weighted (IDW) spatial interpolation tool from ArcGIS to interpolate these years when possible. The above sources identify nations and points with documented timings of horse adoption. Subsequently, we determine the centroid of these nations' territories using the map in Figure 5. We ascertain the interpolated timing of horse adoption across $1^\circ \times 1^\circ$ grid cell, ranging from 54.8 degrees north latitude to 22.7 degrees north latitude and from 124.2 degrees west longitude to 87.6 degrees west longitude. We then aggregate these interpolated values within each nation's territory, yielding both the earliest and average year of horse adoption for each nation. For nations with available data, we reassign the original values. [8] The histogram of horse adoption years is shown in Figure 6.

Our map of homelands was drawn in the year 1600. It is possible that some of these nations moved to other locations, invaded territories or merged of other nations, or even disappeared. Unfortunately, to our knowledge, homeland maps for other years are not available. However, we can still use the original map to calculate the physical size of each nation's homeland and, for a subset of nations for which we have data, we can use this information to calculate population density in 1600.

The second part of empirical analysis is focused on cross-sectional examination to assess the impact of horse adoption on Native American societies' modes of food production. Utilizing data from the Ethnographic Atlas by George P. Murdock (1967), we delve into the transition in reliance on various food production methods post-horse adoption. This comprehensive dataset offers insights into hunting, fishing, agriculture, and animal husbandry practices across diverse Native American communities. The Ethnographic Atlas's meticulous data collection process involved extensive fieldwork by anthropologists, encompassing observations of customs, social structures, economic activities, and cultural traits. Specifically focused on North American societies, researchers interacted closely with indigenous communities across different regions, employing methodologies such as participant observation, interviews, surveys, and analysis of existing ethnographic literature. The resulting dataset provides a nuanced understanding of how horse adoption influenced the economic and dietary landscapes of Native American societies.

[7] It is important to stress the difference between the first year in which a horse was seen in the a nation's homeland and the first year in which horses were known to have been adopted by the nation (Haines, 1938).

[8] See Brown and Cuberes (2022) for a similar interpolation. Figure 4 shows the original map of the horse spread by Haines (1938). Figure 5 shows the interpolated timing of horse adoption and the points where the raw data is available.

Finally, we gather and assemble data on epidemics (Ewers, 1973; Smithsonian Handbook of the North American Indian), heights (Jantz, 1995), and a range of spatial and colonial controls (Feir et al., 2022) from existing work in the literature.

Tables 1 and 2 display summary statistics for the variables described above. Table 1 refers to the cross-sectional data and Table 2 refers to the panel dataset.

5 Empirical Strategy

The study's empirical analysis is divided into two parts. Firstly, it aims to assess the impact of horse adoption on population growth among Native American nations.

5.1 Horse Adoption and Population-A panel analysis

5.1.1 Two-Way Fixed Effects (TWFE)

The two-way fixed effects method is particularly advantageous when the explanatory variable, such as horse adoption, is represented as a dummy variable in panel analyses due to its ability to effectively capture time-invariant unobserved characteristics specific to each nation. In the context of horse adoption, a binary representation (i.e., whether a nation has adopted horses or not) is common, as the adoption of horses is typically an irreversible event for each nation within the panel dataset. By incorporating fixed effects for both nations and time periods, this method accounts for such time-invariant unobserved factors, such as historical cultural practices or geographical constraints, which may influence both the decision to adopt horses and subsequent population growth. This approach helps overcome issues of heterogeneity by allowing for the modeling of nation-specific effects and time-specific trends, effectively controlling for confounding variables and reducing the potential for omitted variable bias.

$$ln(P_{it}) = \alpha + \beta(Horse_{it}) + \mu_i + \lambda_t + \gamma Epidemics_{it} + \epsilon_{it} \tag{15}$$

where i denotes a nation, t is time, $Horse_{it}$ is a dummy variable that takes a value of one when horse is adopted and zero otherwise, η_i is a nation fixed effect, and λ_t are year dummies. *Epidemics* is a dummy variable that takes a value of one if a nation is subject to an epidemic and zero otherwise. The error term is captured by ε.

Because of the irregularity in the time intervals at which we observe the population data we calculate five-year moving averages to smooth big jumps in population caused by these uneven time intervals.

Table 1 presents descriptive statistics for Native American nations, including population and horse adoption data for panel analysis. This study utilizes population data for 82 North American Indian nations, spanning from 1540 to 1980. Notably, this dataset exhibits an unbalanced panel structure, with population data available for varying years across nations, reflecting the challenges of historical data collection during the 1600s. Some nations are represented by as few as 19 individuals due to these challenges, underscoring the meticulous sourcing of data used in this study. The average population across all nations is approximately 3200 residents, while the average horse adoption year stands at 1740, signifying the commencement of equine utilization for daily needs and domestication within these communities. Land cession data documents Native American tribes transferring land to external entities, reflecting the impacts of colonization. Minimum land cession dates back to 1740, spanning to 2010, revealing centuries of territorial changes and dispossession.

5.1.2 Event Study

We next compare the differences in population in treated nations to differences in population in their counterparts in untreated nations over time. In doing so we make the standard assumptions of sharp design, no anticipation and conditional parallel trends. In our context the conditional parallel trends assume that the population of nations that adopted the horse evolved in a parallel to the population of those who did not adopted the horse.[9]

As discussed in Roth (2024), event-study plots allow one to perform "visual inference" to determine whether the treated group's trends appear to have deviated from the comparison group's trends right around the time of treatment. Rambachan and Roth (2023) point out that an event-study plot is generally considered more convincing if there is a discontinuity or kink in the coefficients close to the treatment date, so that the pattern observed cannot simply be explained by a linear violation of parallel trends.[10]

We first define the variable *TimeToTreat* as the difference between the actual year and the year of horse adoption for every nation. We estimate a traditional dynamic TWFE specification as in Roth (2024):

$$Y_{it} = \alpha_i + \lambda_t + \sum_{r \neq -1} \beta_r \times 1[D_i = 1] \times 1[t = r + 1] + \varepsilon_{it} \tag{16}$$

where D_i is an indicator for treatment (horse adoption) and α_i and λ_t are unit and period fixed effects. ε is the error term. The event-study coefficients are given by:

[9] We carry out our analysis using the Stata command *eventdd*.

[10] An event model like the one used here can also be considered a "dynamic" version of the TWFE model.

$$\hat{\beta}_r = \Big(\hat{E}[Y_{i,r+1}|D_i = 1] - \hat{E}[Y_{i,r+1}|D_i = 0]\Big) - \Big(\hat{E}[Y_{i0}|D_i = 1] - \hat{E}[Y_{i0}|D_i = 0]\Big) \tag{17}$$

where $\hat{E}$ denotes the sample average. In other words, $\hat{\beta}_r$ represents a 2-group, 2-period DiD estimate comparing the $D_i = 1$ and $D_i = 0$ groups between periods $r + 1$ and 0. This formula is symmetric between pre-treatment and post-treatment periods, i.e. it is the same regardless of whether r is greater than or less than zero.

5.1.3 Difference-in-Differences

Several recent papers (Borusyak and Jaravel, 2017; de Chaisemartin and D'Haultfuille, 2020; Sun and Abraham, 2021; Goodman-Bacon, 2021; Roth et al., 2023) argue that TWFE do not always perform well with staggered and heterogeneous treatment effects[11], even under a parallel trends assumption. In particular, Sun and Abraham (2021) demonstrate that the β's in (16)-(17) cannot be rigorously interpreted as reliable measures of "dynamic treatment effects". Using the language of Rios-Avila et al. (2021), the main problem with these TWFE specifications is that OLS is "variational hungry". This is the case because this estimator attempts to compare all cohorts with each other, as long as there is variation in treatment status in that given time-window. This is a problem since OLS does not appropriately account for the treatment and comparison groups; it simply minimizes the mean-squared error of the problem.

We first use the Two-Stage Difference-in-Differences proposed by Butts and Gardner (2021). We use as control groups nations that, at a given point in time, had not yet adopted the horse. The model we estimate is as follows:

$$ln(P_{it}) = \alpha + \beta Horse_{it} + \gamma Horse_{it} * Post_t + \eta_i + \delta t + \varepsilon_{it} \tag{18}$$

where $Horse_i$ is a treatment dummy variable that takes a value of one for nations that have ever adopted the horse and zero for those that have never done so in the time interval considered in the regression.

Callaway and Sant'Anna (2021) propose a different way to proceed. We estimate their Difference in Difference with Multiple periods estimator using this stata command *csdid*. [12]

[11] Treatment effects may be heterogeneous over time, across units, or both.

[12] This is a generalization of *drdid*. Note that *drdid* cannot be estimated in our context we don't have a

Their procedure allows for various estimators. More specifically, their technique estimates every feasible 2x2 DiD design available in the selected sample. It is important to note that a "never-treated" group is not required for this estimator to work. This is important because in our sample and if we use the entire period 1600-1860 all nations are eventually treated i.e all nations end up adopting horses at some point. In one of our robustness exercises we study shorter periods of time in order to have some nations that are never treated. As in *drdid*, the underlying assumption is that all covariates are time constant. When using panel data, even if covariates are time-varying, only the base-period (earlier-period) values are used for the estimation.

5.2 Horse Adoption and Subsistence Modes- A cross section analysis

After having investigated the impact of horse adoption on Native Americans population using panel data, in this subsection we use cross-sectional data to analyze whether the number of years a nation has had horses has an effect on their subsistence mode.

5.2.1 OLS

We first estimate a simple OLS regression:

$$Hunting = \alpha + \beta_1 YWH + \Gamma' Geography + \varepsilon \tag{19}$$

The variable YWH is the number of years a nation has had horses. This is calculated as the difference between the year at which nations were first observed by the *Ethnographic Atlas* (Murdock, 1967) and the year of horse adoption. The atlas represents 1,267 societies from around the world, and all nations are observed before industrialization and European contact. Among these societies, we use 115 nations in North America. The database includes information on cultural, institutional, and economic characteristics. Figure 27 in the appendix provides the nation's name and the nation's observation year used in the paper. *Geography* is a vector that includes several geographic controls, namely latitude, longitude, their product, terrain ruggedness, mean elevation, area, and caloric suitability pre-1500 CE. Table 2 presents descriptive statistics for various variables in the cross-sectional analysis.

5.2.2 Two-Step Procedure

We next proceed in two steps to analyze the effect of the number of years with horses on a nation's subsistence method. We first predict the geographical determinants of horse

2*2 design.

adoption. Next we use this predicted value as an independent variable in the regression that analyzes the subsistence methods.

Step 1: Predicting the Number of Years with Horses

We start by predicting the year at which nations adopted the horse. To do so we use information on these years from different sources. The first one is Wissler (1914) who provides information on the first time different explorers mentioned the presence of horses near the homeland of 25 nations. Figure 6 displays this map. The map shows that the horses abandoned by the Spanish explorers De Soto and Coronado around 1580 the Southwest U.S., near today's Santa Fé. As stated in Haines (1938) "...the center from where horses spread to the Indian tribes was the stock-raising area around Santa Fé, and from here they spread very slowly at first but later more rapidly until the entire plains and plateau country had been supplied." (Haines, p.429). From there, the so called "Great Horse Dispersal" moved northwards. The second one is Haines (1938) who documents the spread of horses among the Plains Indians. His study provides specific dates for several Plains Nations as well as a map of the diffusion of horses in North America in the 17th and 18th century. Our third source of data on horse adoption is Brooke Jacobsen and Eighmy (1957).

In terms of explanatory variables, a natural one we consider is land suitability for agriculture. Under the assumption that horses in that period were mostly used for hunting and warfare, one would expect that, everything else equal, nations with better land suitability would adopt horses later on. We measure suitability of land for agriculture in two different ways. The first measure, developed by Galor and Ozak (2016) and labeled *caloric suitability* calculates the maximum calories obtainable from a specific crop. This measure is available both before and after the Columbian Exchange (as 'Caloric Suitability pre 1500 CE' and 'post 1500 CE') in the slides). Caloric suitability is available pre and post-1500 CE, and we examine the explanatory power of both variables.

We also control for he latitude, longitude of the centroid of the nation's homeland as well as its average terrain ruggedness. The regression we estimate is:

$$Y_{ij} = \alpha + \beta_1 L_{ij} + \beta_2 Dist_{ij} + \lambda_j + \Gamma' X_j \varepsilon_{ij} \tag{20}$$

where the subscripts i and j respectively index nations and regions, Y denotes the year of horse adoption, L is a measure of land suitability for agriculture, $Dist$ is distance from the location where the De Soto's horses were abandoned. λ_j are regional fixed effects and the matrix X contains the latitude, longitude of the centroid of the nation's homeland, their

product, terrain ruggedness, area, mean elevation, and caloric suitability.

Step 2: Estimating the Effect of Horse Adoption on Subsistence Modes

We estimate

$$y_i = \alpha + \beta(Years\ Since\ Adoption_i) + X_i'\Gamma + \epsilon_i \quad (21)$$

where y is a categorical variable that indicates the percentage range of the dependence on hunting for nation i. We then estimate similar regressions for agriculture, fishing, and gathering. *Years since adoption* are calculated as the difference between a nation's observational time (from Murdoch, 1967) and the year of horse adoption. The vector X includes as controls latitude, longitude, their product, terrain ruggedness, area, mean elevation, and caloric suitability.

5.2.3 Instrumental Variables: 2SLS

A natural concern in our empirical exercise is the endogeneity of horse adoption. More populated nations may have the incentives and means to adopt horses more easily. Moreover, several omitted variables may drive changes in population or in our outcome variables. Examples of such variables may be trade routes with the colonizers, access to guns and rifles, etc.

Reverse causality in the context of horse adoption and population growth can be framed within the Malthusian theory, wherein the potential mechanism is that regions experiencing population growth may engender an increased proclivity to adopt horses. According to Malthusian dynamics, as articulated in the Malthusian trap, population growth tends to outpace the ability of resources to sustain it. In response to the burgeoning populace, societies may seek technological advancements, such as adopting horses, to enhance productivity and alleviate the strain on resources. The Malthusian argument posits that a growing population, encountering the limitations of traditional agricultural practices, may opt for the adoption of horses to augment agricultural output, transportation capabilities, and overall productivity. In this narrative, population growth becomes a driving force behind the decision to adopt horses, reflecting the intrinsic connection between demographic patterns and technological choices in a Malthusian framework.

In this scenario, the exogeneity of horse diffusion is compromised, introducing potential endogeneity concerns in our analysis of adoption rates across nations. To address this identifi-

cation challenge, we have shifted our analytical approach towards employing an instrumental variable. A suitable instrumental variable should exert an impact on horse diffusion while maintaining independence from population growth influences.

Existing literature posits that horses exhibit a reluctance to traverse large water bodies compared to humans. This observation opens the possibility of utilizing water bodies as a variable influencing horse diffusion within and between nations in the North American region. Additionally, the presence of water bodies may exhibit a limited correlation with population growth.The instrumental variable strategy employed parallels the concept of the least-cost path analysis in diffusion studies. Recognizing that water bodies alone may not adequately capture the least-cost path, we amalgamated the human mobility index (HMI) with water bodies to create a comprehensive instrumental variable.

Our instrumental variable is based on the Human Mobility Index (HMI henceforth) constructed by Özak (2010, 2018). This index represents the least cost for human movement between grid cells without major transport interventions such as trains. To make the index relevant for our problem we then interact it with large water bodies. Teh rationale for doing this is as follows. As the HMI is not explicitly constructed to encapsulate the travel cost for horses, we enhanced its utility by incorporating water bodies. Given that horses, even when trained, encounter significant challenges in crossing large rivers and lakes, we applied a multiplier to the HMI values for grid cells that traverse such water bodies. For instance, if a grid cell (i) has a HMI value of five and traverses water bodies, we reassigned the value to 10 (i.e., 5 multiplied by 2) to reflect the heightened cost associated with crossing rivers and lakes.

When assessing the strength of the horse adoption argument using the instrumental variable (IV) based on the interaction between the human mobility index (HMI) and water bodies, considering variation over time becomes crucial. The effectiveness of this instrument in capturing the dynamics of horse diffusion relies on temporal fluctuations and changes in the factors influencing both HMI and water bodies.

In a cross-sectional analysis, the instrument is designed to exploit differences across nations at a specific point in time. The variation in HMI values, coupled with the presence of water bodies, helps discern the impact of these factors on horse adoption rates across different countries. However, for panel analysis where observations are tracked over time, the instrumental variable's strength may be compromised.

The reason lies in the assumption of time-invariant effects of the instrumental variable. If the interaction between HMI and water bodies exhibits substantial variation over time, the

instrument may not adequately capture these temporal dynamics. Panel data necessitates a stable and consistent instrument across time to ensure the validity of the identification strategy. Changes in the relationship between HMI, water bodies, and horse adoption over time could introduce endogeneity concerns, undermining the instrument's suitability for panel analyses. Therefore, while the instrumental variable based on HMI and water bodies holds promise in cross-sectional settings by leveraging spatial differences, its effectiveness in panel analyses may be limited due to the requirement for stable temporal relationships between the instrument and the endogenous variable.

In this subsection, we employ the interpolated timing of horse adoption directly as an independent variable. This approach differs from the two-step procedure where the independent variable is constructed as a predicted value from exogenous geographical variables. Using the interpolated timing of horse adoption directly could potentially introduce endogeneity issues. To alleviate this concern, we take the 2SLS approach that the interpolated timing is instrumented by the least cost path that considers water bodies. Figure 8 displays the first-stage of our instrument. It is apparent that for nations with a more costly path of horse adoption the horse was adopted later.

5.3 Horse Adoption and Other Outcomes

5.3.1 Income and Heights

This study examines the long-term welfare effects on the wellbeing of Native American nations by analyzing the impact of horse adoption on their heights. Existing evidence underscores that changes in height reflect long-term phenomena, often taking several decades to reveal the impact of available food resources. This study has employed height data from Jantz (1995), motivated by its comprehensive examination of health and disease patterns among late prehistoric and early historic populations of the Mississippi River Valley. Through meticulous skeletal analysis, particularly focusing on long bone lengths, researchers were able to estimate the stature of ancient individuals, thereby offering valuable insights into nutrition, health, and living conditions. This research significantly contributes to our understanding of the physical characteristics and temporal changes among Native American populations, thereby assisting archaeologists, anthropologists, and historians in comprehending their experiences before and during European contact.

5.3.2 Land Cessions

TO BE ADDED

6 Empirical Results

6.1 Horse Adoption and Population

6.1.1 Two-Way Fixed Effects (TWFE)

Table 3 shows the estimates of the TWFE models. Column (1) indicates that the horse dummy has a positive and significant effect. The same is true when we control for other covariates (columns (2)-(4)). Epidemics have a negative and significant impact on population in most specifications. Weals ad in column (3) two dummy variables for significant events like the U.S. independence and the Trail of Tears of 1830 were the so-called Civilized Nations-were forced to move to the territory that constitutes Oklahoma today. In column (4) we add nation-specific time trends and the coefficient on horses remains significant at standard levels.

6.1.2 Event Study

Figure 7 displays the results of estimating our event study model. The figure indicates that around the year of horse adoption population tends to increase. Population then increases rapidly and it experiences a sharp decline after about twenty years of the shock. Moreover, the point estimates do not seem to detect any significant pre-trend.

6.1.3 Difference-in-Differences

Table 4 displays the average treatment effects of horse adoption on NA population overall (column (4)) as well as for different groups (column (5)) and over time (column (6)). Using the actual data on horse adoption by different nations we observe that the ATT is positive and significant for the first two cases. However, the p-value of the null hypothesis of no pre-trend is very low, indicating that one can reject the common trend assumption. However, slightly modifying the year of horse adoption by subtracting one year from the actual dates (see the row *Horse year-1)* allows us to not reject the common pre-trend assumption. We need to further study this, but since measurement error is likely to be quite important when predicting the exact year of horse adoption, it seems reasonable to consider more than one year around the estimated date. This table indicates that the pre-trend disappears of one considers a one-year lag in horse adoption.

6.2 Horse Adoption and Subsistence Modes

6.2.1 Two-Step Procedure

Step 1: Predicting the Number of Years with Horses

Table 5 displays the OLS results of estimating equation (20). The suitability of land for agriculture its caloric suitability pre-1500 CE exhibits a negative sign in column (1) but the coefficient becomes positive and significant once more controls are added (columns (8)-(10)). This positive sign is somewhat confirmed when we use caloric suitability post 1500 CE, although in this case, it is statistically significant only in column (1). The same happens when we use *land suitability*, which measures the fraction of land that is suitable for agriculture. These results provide strong evidence that horses were adopted later in areas more suitable for agriculture. One interpretation of this result is that horses were mostly used for hunting and not for agriculture and therefore nations that had a comparative advantage in agriculture had less incentives to adopt the horse.

Latitude is very strongly associated with horse adoption. This makes sense since horses originated in the South of the U/S. and then moved towards the North. Horses also moved from the West to the East (see Figures 4 and 5) and so it is not surprising that the coefficient on longitude is negative. Indeed, the distance from the origin of horses is enters with a positive and significant sign too. We also include the product of latitude and longitude to account for potential spatial correlation in the error terms (Kelly, 2021). This variable is highly significant, suggesting that it is an important variable to control for.

This exercise also provides indirect evidence that horse adoption may have led some nations to switch from agriculture to hunting as their main mode of subsistence. Our argument goes as follows. First, as shown above, we find that nations with better land for agriculture (at observation time) were likely to have agriculture and less hunting as their main subsistence methods. Second, nations with better land for agriculture (at observation time) adopted the horse later. This was indeed the case as shown in Figure 9. Third, assuming that the suitability of land for agriculture did not change much between the year of horse adoption and the observation year, the two previous observations imply that if one finds nations that did mostly hunting (at observation time) in spite of having good land for agriculture one can probably state that these were nations that used to do agriculture but then switched to hunting when they adopted the horse.

Step 2: Estimating the Effect of Horse Adoption on Subsistence Modes

Table 6 shows the results of estimating equation (21), i.e., it uses the predicted number of years a nation had horses as an explanatory variable for the nation's dependence on hunting. Our explanatory variable is constructed from exogenous geography, and thus it is also exogenous. Therefore, we examine the relation of interest simply via the OLS. Column (1) shows an unconditional estimate that is not statistically significant. However, by adding basic geographical characteristics in columns (2) and (3), a hump-shaped relationship appears. Column (4) controls for caloric suitability and this does not alter the estimated coefficient. In all cases, since these two regressors are estimated, we bootstrap the standard errors associated with their estimates. More years with horses leads to more dependence on hunting once we control for several geographical variables (columns (2)-(4)). Moreover, in columns (3) and (4), the relationship between horse adoption and hunting is clearly concave.

Figure 10 displays the hump-shaped relationship between years since horse adoption and hunting. One interpretation of this finding is that, initially horse adoption led nations to increase their hunting. However, as bison become scarce over time, a nation's dependence on hunting decreased.

Table 7 explores the relationship between horse adoption and subsistence modes other than hunting. Our estimates show no relationship between these two variables for any of the alternative subsistence modes: gathering, fishing, animal husbandry, or agriculture.

6.2.2 Instrumental Variables: 2SLS

In this subsection, we shift our approach to directly utilizing the interpolated timing of horse adoption as an independent variable, as opposed to the timing derived from the two-stage procedure. The interpolated year of horse adoption is generated using the Inverse Distance Weighted (IDW) interpolation method within GIS. Drawing upon sources such as Wissler (1914), Haines (1938), and the Smithsonian Handbook, we initially identify nations with documented timings of horse adoption. Subsequently, we determine the centroid of these nations' territories using the map in Figure 5. By integrating this data with the IDW interpolation, we ascertain the interpolated timing of horse adoption across $1^\circ \times 1^\circ$ grid cells, ranging from 54.8 degrees north latitude to 22.7 degrees north latitude and from 124.2 degrees west longitude to 87.6 degrees west longitude. We then aggregate these interpolated values within each nation's territory, yielding both the earliest and average years of horse adoption for each nation.[13] For nations with available data, we reassign the original values.

[13]Both the earliest timing and the average have pros and cons. Conceptually, we are interested in the impacts of horse adoption captured by the earliest adoption timing. Moreover, the earliest timing produces more variation than the average. Thus, we use it below. Given the measurement error related to the maps

Finally, we get the interpolated time elapsed since horse adoption (ITSHA) as the observed year minus the interpolated year of adoption, where the observed year is taken from Murdock (1964).

The interpolated timing might not be exogenous, unlike the horse adoption timing derived from the two-stage procedure. To alleviate the endogeneity concerns, we instrument ITSHA with the HMI augmented by water bodies. In particular, our empirical specification is the following:

$$Hunting_i = \alpha_0 + \alpha_1 \widehat{ITSHA}_i + \alpha_2 \widehat{ITSHA}_i^{\,2} + \lambda Geo + \epsilon_i, \tag{22}$$

where $\widehat{ITSHA}_i$ and $\widehat{ITSHA}_i^{\,2}$ are predictions from the HMI augmented by water bodies and its square terms, *Geo* is a vector of geographical characteristics. We hypothesize that $\alpha_1 > 0$ and $\alpha_2 < 0$.

Table 8 shows a hump-shaped relationship between ITSHA and dependence on hunting for nations in northern America. Unconditional regression in column (1) shows a statistically highly significant association. Columns (2) and (3) add potentially confounding geography such as latitude, longitude, their product, terrain ruggedness, and mean elevation. Reassuringly, the estimated coefficients are very stable to these controls. Column (4) includes caloric suitability pre-1500 CE. Although it reduces the coefficient sizes, the estimates are highly significant. Column (5) provides the 2SLS estimates, further supporting the hump-shaped association between ITSHA and dependence on hunting, figure 10 depicts the estimated association from the 2SLS. The optimal number of years elapsed since horse adoption is 138, closest to Quileute on Northwest Coast.

7 Conclusions

TO BE COMPLETED

Horse adoption is strongly associated with a rapid re-population of Indigenous nations that gain access to the animal. We also find evidence of partial re-specialization into hunting activities and away from other forms of food production. We do not observe effects on long-run heights of horse adopters

of nations' territory, using the earliest timing might increase mismeasurement. However, using the average produces similar results.

We hypothesize that this economic boom, and the associated positive military shock, likely delayed the time of conquest for these nations and increased the likelihood that these peoples persist into the present, although this work remains in progress

References

[1] Akee, R., Jones, M. R., and Porter, S. R. (2019). "Race matters: Income shares, income inequality, and income mobility for all US races". Demography, 56(3), 999-1021.

[2] Brander, J. A. and Taylor, S. M. (1998), The Simple Economics of Easter Island: A Ricardo-Malthus Model of Renewable Resource. The American Economic Review, Vol. 88, No. 1 March, pp. 119-138.

[3] Borusyak, K., and X. Jaravel (2017),"Revisiting Event Study Designs." SSRN Scholarly Paper ID 2826228. Social Science Research Network. August.

[4] Brooke Jacobsen, R., and J. L. Eighmy (1980), Plains Anthropologist , NOVEMBER 1980, Vol. 25, No. 90, November, pp. 333-341.

[5] Brown, J. and D. Cuberes (2024) "The Birth and Persistence of Cities: Evidence from the Oklahoma's First Fifty Years of Urban Growth".

[6] Butts, K., and Gardner, J. (2021). Two-stage difference-in-differences. arXiv preprint arXiv: 2109.05913.

[7] Callaway, B., and Sant'Anna, P. H. C. (2021), "Difference-in-Differences with multiple time periods." 225(2):200-230. Journal of Econometrics.

[8] de Chaisemartin, C., and X. D'Haultfuille (2020), "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." American Economic Review, 110(9), pp. 2964-2996.

[9] De La Croix, D., and D. Dottori (2008), "Easter Island Collapse: a Tale of Population Race." Journal of Economic Growth, 13, 27-55.

[10] Delsol, N., Stucky, B. J., Oswald, J. A., Reitz, E. J., Emery, K. F., and Guralnick, R. (2022). "Analysis of the earliest complete mtDNA genome of a Caribbean colonial horse (Equus caballus) from 16th-century Haiti". PloS one, 17(7), e0270600

[11] Diamond, J. M. (2005). Guns, germs, and steel : the fates of human societies. New York: Norton.

[12] Ewers, J. C. (1973). "The influence of epidemics on the Indian populations and cultures of Texas". Plains Anthropologist, 18(60), 104-115.

[13] Fleharty, E. D. (1995). Wild animals and settlers on the Great Plains. University of Oklahoma Press.

[14] Featherling, J. (2019). "Horses, Culture, and Trade: The Impact of the Horse on Southeastern Native Nations, 1650-1830".

[15] Feir, D. L., Gillezeau, R., and Jones, M. E. (2022). "The slaughter of the bison and reversal of fortunes on the great plains" (No. w30368). National Bureau of Economic Research

[16] Freyaldenhoven, S., C. Hansen, and J. M. Shapiro (2019). "Pre-event Trends in the Panel Event-Study Design." American Economic Review, 109 (9): 3307-38

[17] Gallo, K. and E. Wood (2015) Historical Drought Events of the Great Plains Recorded by Native Americans Great Plains Research, Vol. 25, No. 2, Fall, pp. 151-158, Published by: University of Nebraska Press.

[18] Galor, O., and Özak, Ö. (2016). "The agricultural origins of time preference". American Economic Review, 106(10), 3064-3103.

[19] Goodman-Bacon, A. (2021), "Difference-in-Differences with Variation in Treatment Timing." Journal of Econometrics.

[20] Hämäläinen, P. (2003), "The Rise and Fall of Plains Indian Horse Cultures", The Journal of American History, December.

[21] Haines, F. (1938), "The Northward Spread of Horses among the Plains Indians", American Anthropologist , Jul. - Sep., New Series, Vol. 40, No. 3, Jul. Sep, pp. 429-437.

[22] Heizer, R. F., and Cook, S. F. (1956), "Some aspects of the quantitative approach in archaeology". Southwestern Journal of Anthropology, 12(3), 229-248.

[23] Henderson, V., Tim Squires, and Adam Storeygard, D. Weil (2018), "The Global Spatial Distribution of Economic Activity: Nature, History, and the Role of Trade," Quarterly Journal of Economics, 133:1, February 2018, Pages 357–406.

[24] Jacobsen, R. B., and Eighmy, J. L. (1980). A mathematical theory of horse adoption on the North American plains. Plains Anthropologist, 25(90), 333-342.

[25] Jantz, R. L. (1995). "Franz boas and native american biological variability". Human Biology, 345-353.

[26] Kelly, M. (2021), "Persistence, Randomization, and Spatial Noise." WP21/25 University College Dublin.

[27] Kroeber, A. L. (1925), "Handbook of the Indians of California"(Vol. 78). US Government Printing Office

[28] Kumagai, M. (2023). "The Horse, Battles, and the State: Military Origins of Autocracy". mimeo.

[29] McKnight, T. L. (1959). The Feral Horse in Anglo-America. Geographical Review, 49(4), 506–525. https://doi.org/10.2307/212210.

[30] Michalopoulos (2012), "The Origins of Ethnolinguistic Diversity." American Economic Review, 2012, 102(4): 1508-1539.

[31] Murdock, G. P. (1967). Ethnographic atlas: a summary. Ethnology, 6(2), 109-236.

[32] Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V., Underwood, E. C., ... and Kassem, K. R. (2001). "Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity". BioScience, 51(11), 933-938.

[33] Özak, Ö. (2018). Distance to the pre-industrial technological frontier and economic development. Journal of Economic Growth, 23(2), 175-221.

[34] Perdue, T., and Green, M. D. (2007). "The Cherokee nation and the trail of tears". Penguin.

[35] Ramankutty, N., Foley, J. A., Norman, J., and McSweeney, K. (2002). The global distribution of cultivable lands: current patterns and sensitivity to possible climate change. Global Ecology and biogeography, 11(5), 377-392. 17.

[36] Rambachan A., and J. Roth (2023), "A More Credible Approach to Parallel Trends." The Review of Economic Studies, October, 90(5), pp. 305-322.

[37] Reddy, M. A. (1995). Statistical record of native North Americans. Information/Reference Group, An International Thomson Publishing Co., PO Box 71515, Chicago, IL 60694-1515.

[38] Rios-Avila, F., Callaway, B., and Sant'Anna, P. H. (2021, August). csdid: Difference-in-differences with multiple time periods in stata. In Stata Conference (p. 47).

[39] Roth, J. (2024). "Interpreting Event-Studies from Recent Difference-in-Differences Methods". arXiv preprint arXiv:2401.12309.

[40] Roth, J., P. H. C. SantAnna, A. Bilinski, and J. Poe (2023), "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature." Journal of Econometrics, August, 235(2), pp. 2218-2244.

[41] Stevens, Forrest R., Andrea E. Gaughan, Catherine Linard, and Andrew J. Tatem, "Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data," PLoS One, 10 (2015), e0107042.

[42] Sun, L., and S. Abraham (2021), "Estimating Dynamic Treatment Effects in Event-Studies with Heterogeneous Treatment Effects." Journal of Econometrics, 225(2), pp175-199.

[43] Taylor, W. T. T., Librado, P., Hunska Tašunke Icu, M., Shield Chief Gover, C., Arterberry, J., Luta Win, A.,and Orlando, L. (2023). "Early dispersal of domestic horses into the Great Plains and northern Rockies". Science, 379(6639), 1316-1323.

[44] Ubelaker, D.H. (1988), North American Indian Population Size, A.D. 1500to 1985, American Journal of Physical Anthropoloy 77289-294.

[45] Waters, M. R., Stafford Jr, T. W., Kooyman, B., Hills, L. V. (2015). Late Pleistocene horse and camel hunting at the southern margin of the ice-free corridor: Reassessing the age of Wally's Beach, Canada. Proceedings of the National Academy of Sciences, 112(14), 4263-4267.

[46] Wissler, C. (1914) The Influeence of the Horse in the Development of Plains Culture", American Anthropologist, Vol. 16, pp. 1-25.

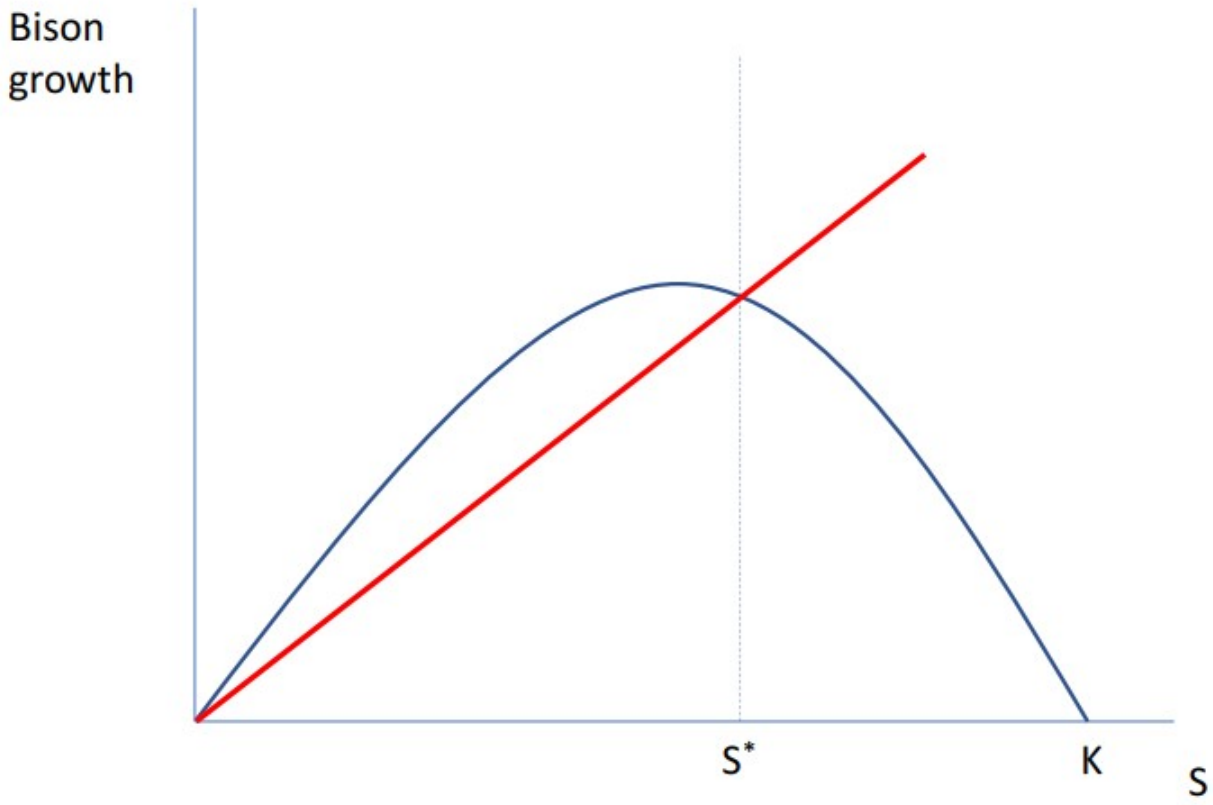


Figure 1: The steady state number of bison

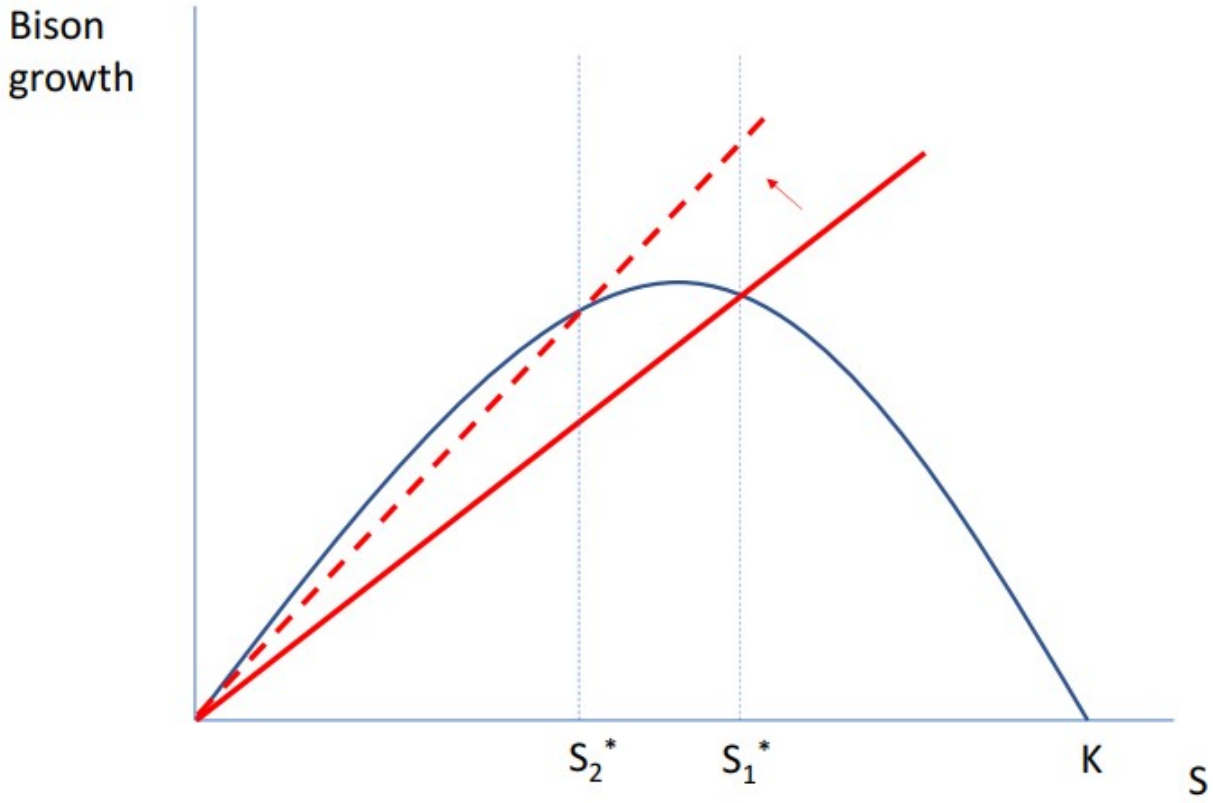


Figure 2: The effect of horse adoption on the stock of bison

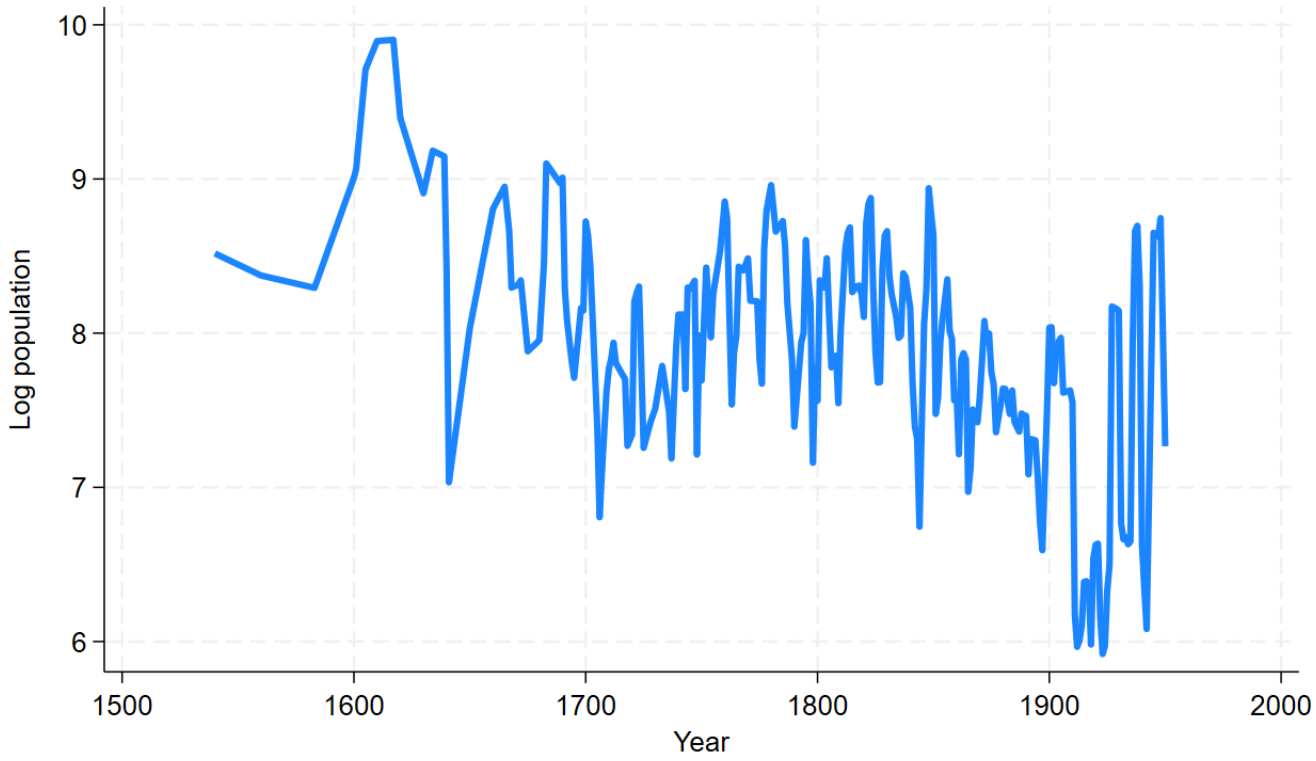


Figure 3: Historical trend of Native American population

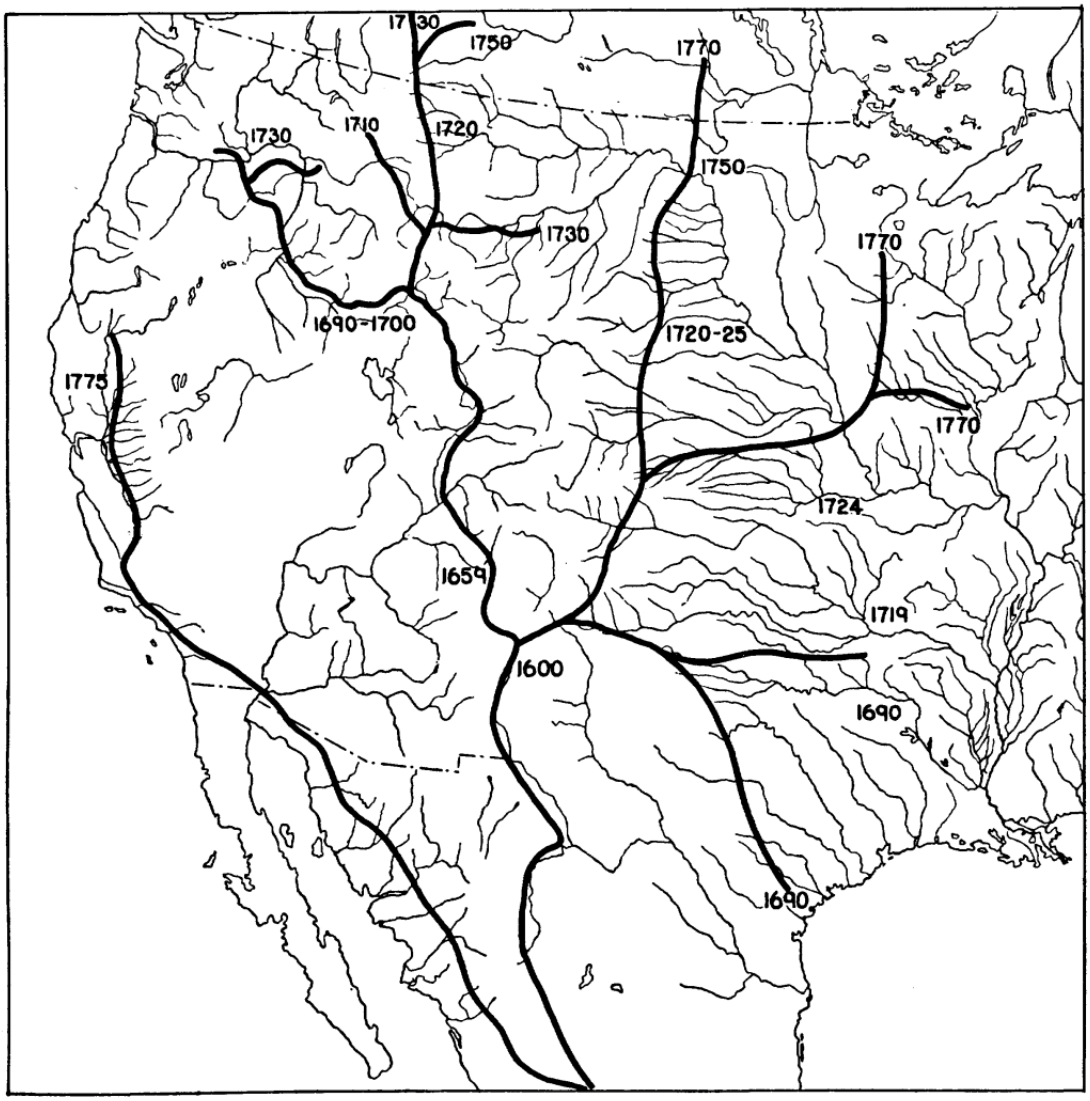


Figure 4: The Diffusion of the Horse (Haines 1938)

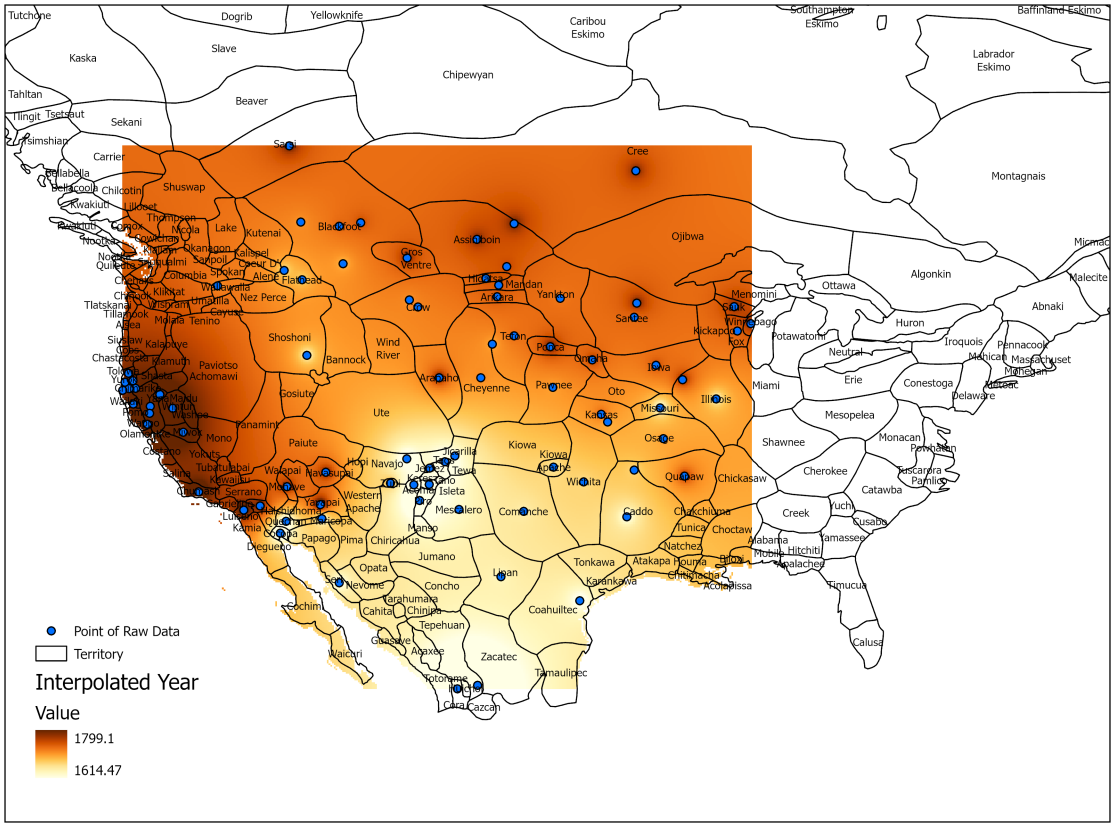


Figure 5: Interpolated Timing of Horse Adoption

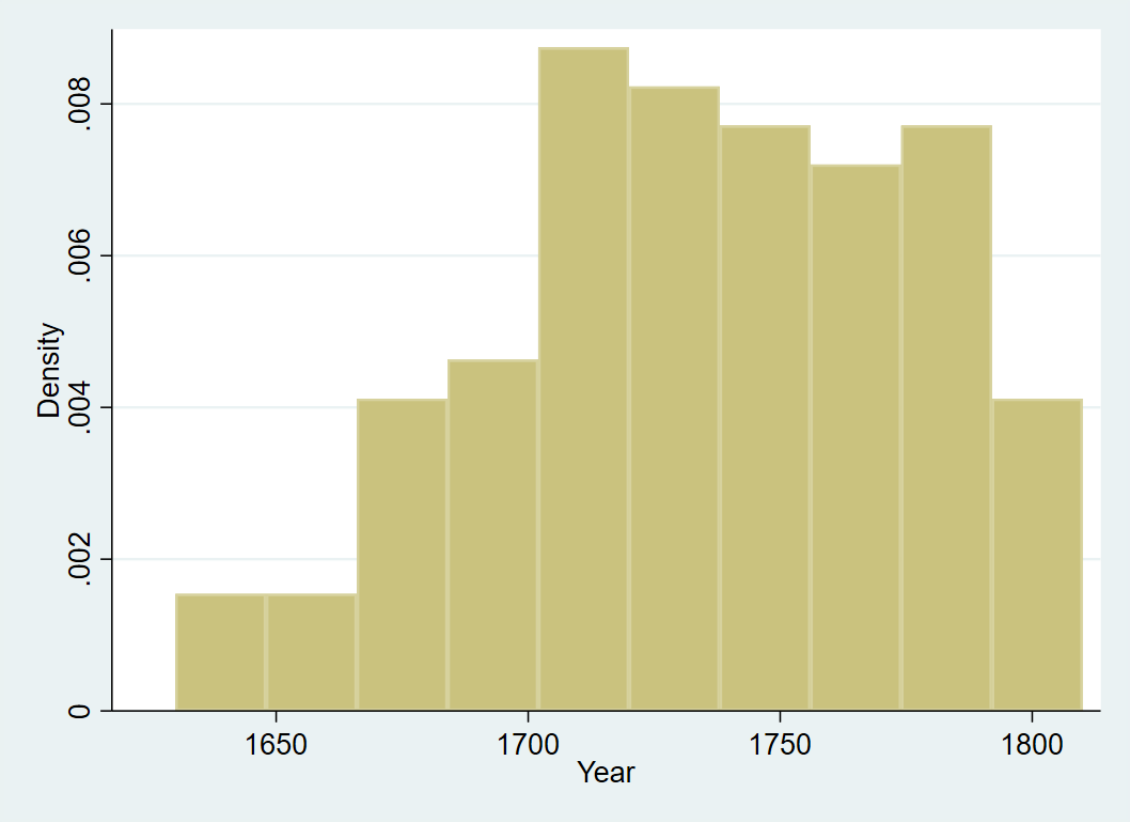


Figure 6: Histogram of Horse Adoption Years

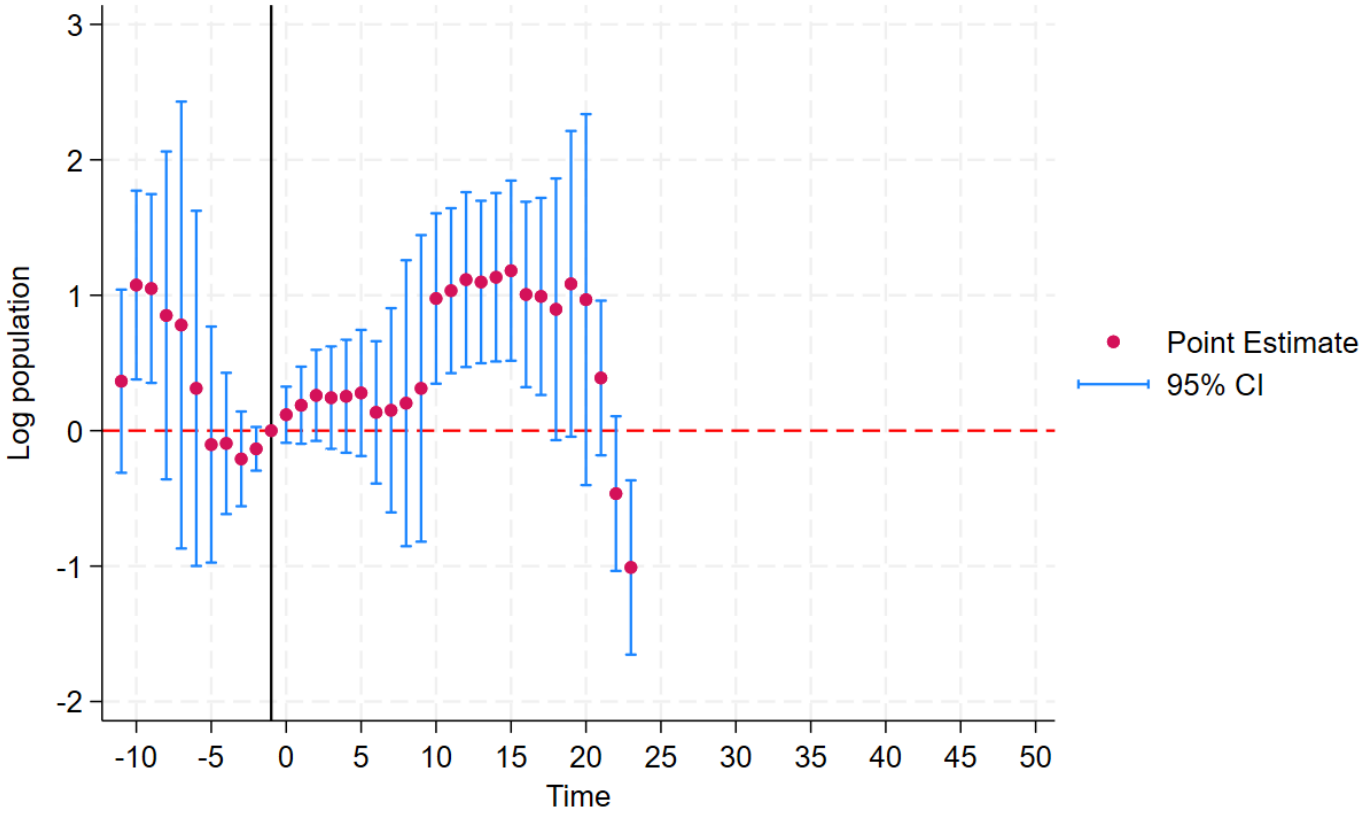


Figure 7: Event Study (1540-1950)

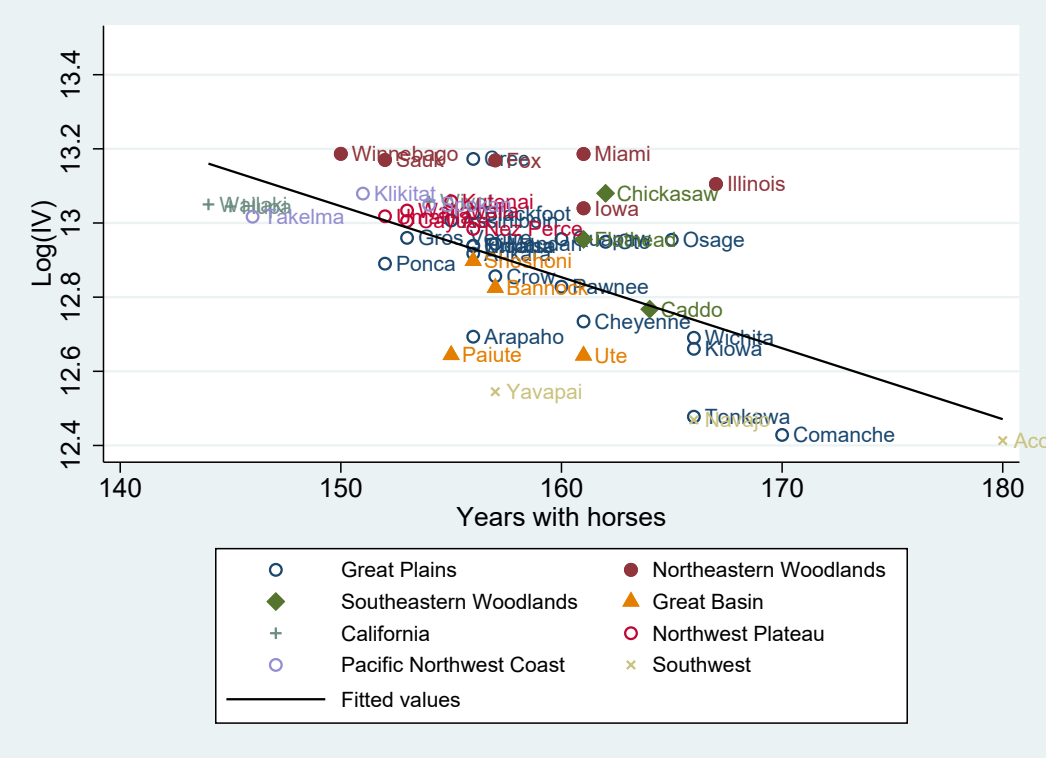


Figure 8: An instrument

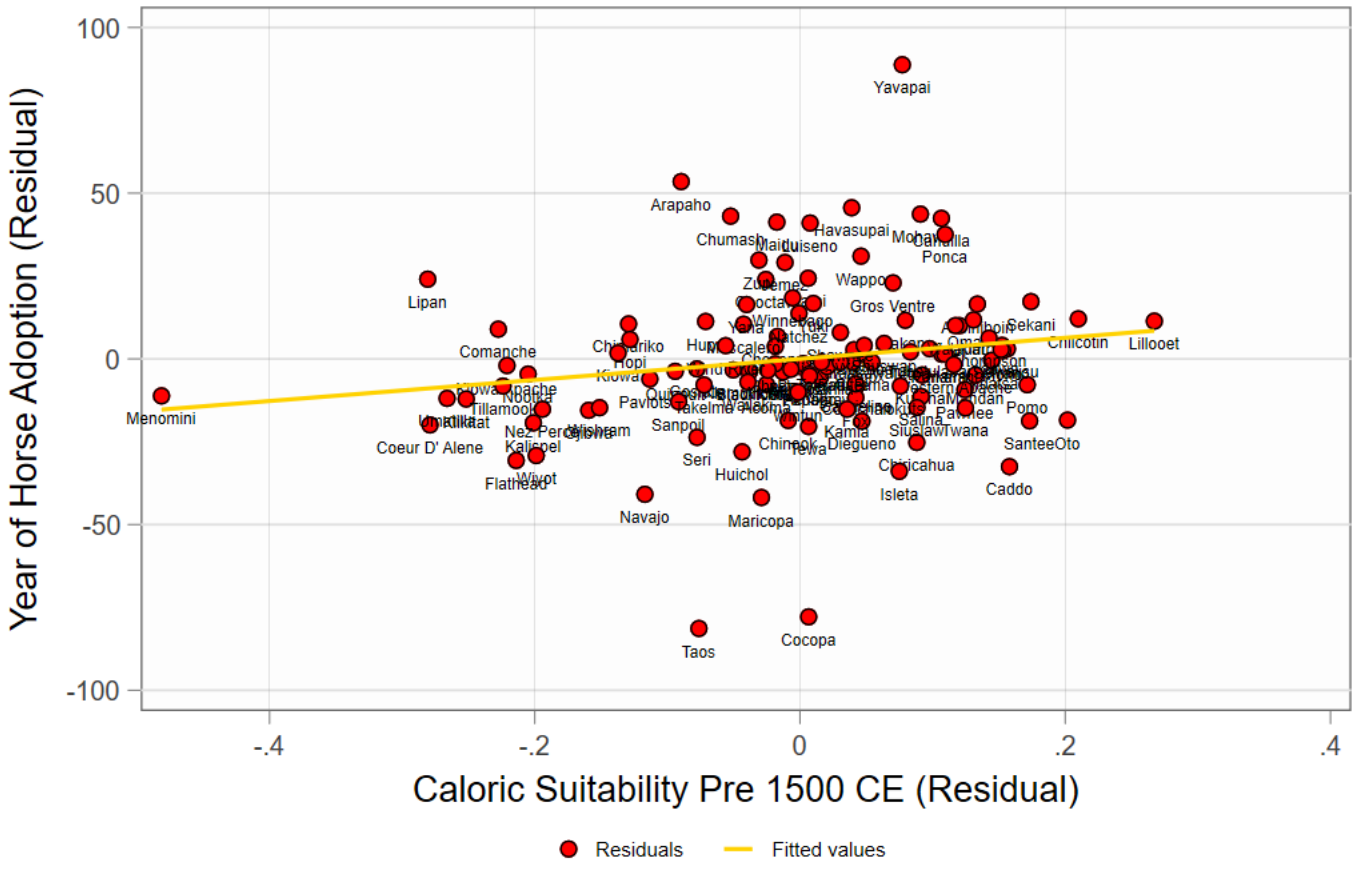


Figure 9: Land Caloric Suitability and Year of Horse Adoption

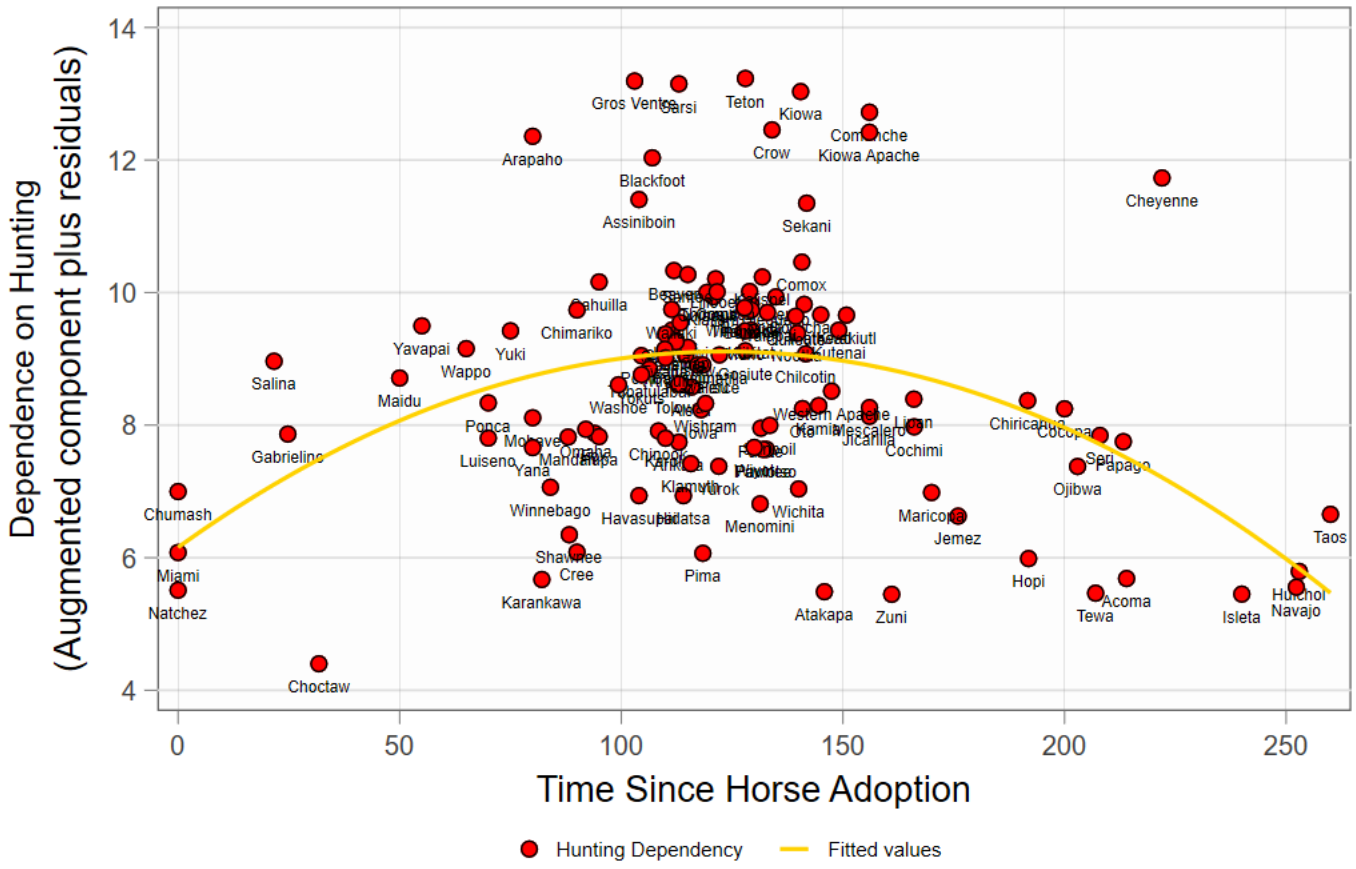


Figure 10: Horse Adoption and Hunting Dependence

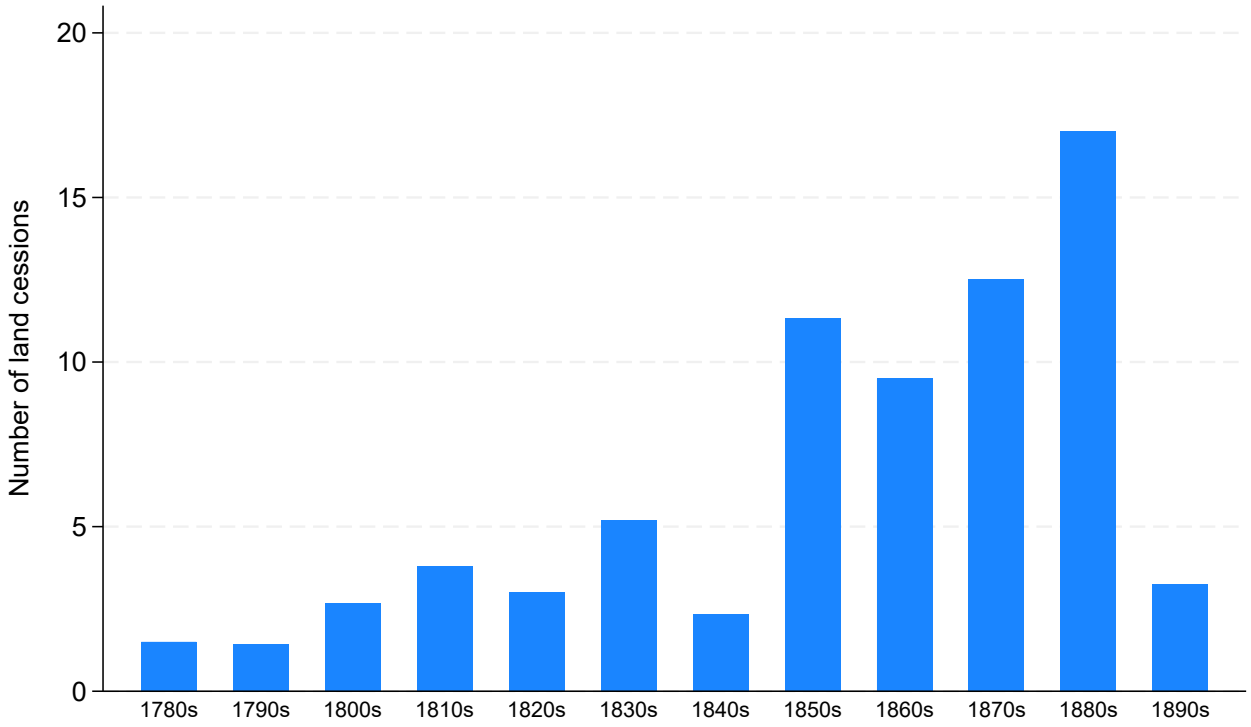


Figure 11: Land Cessions over Time

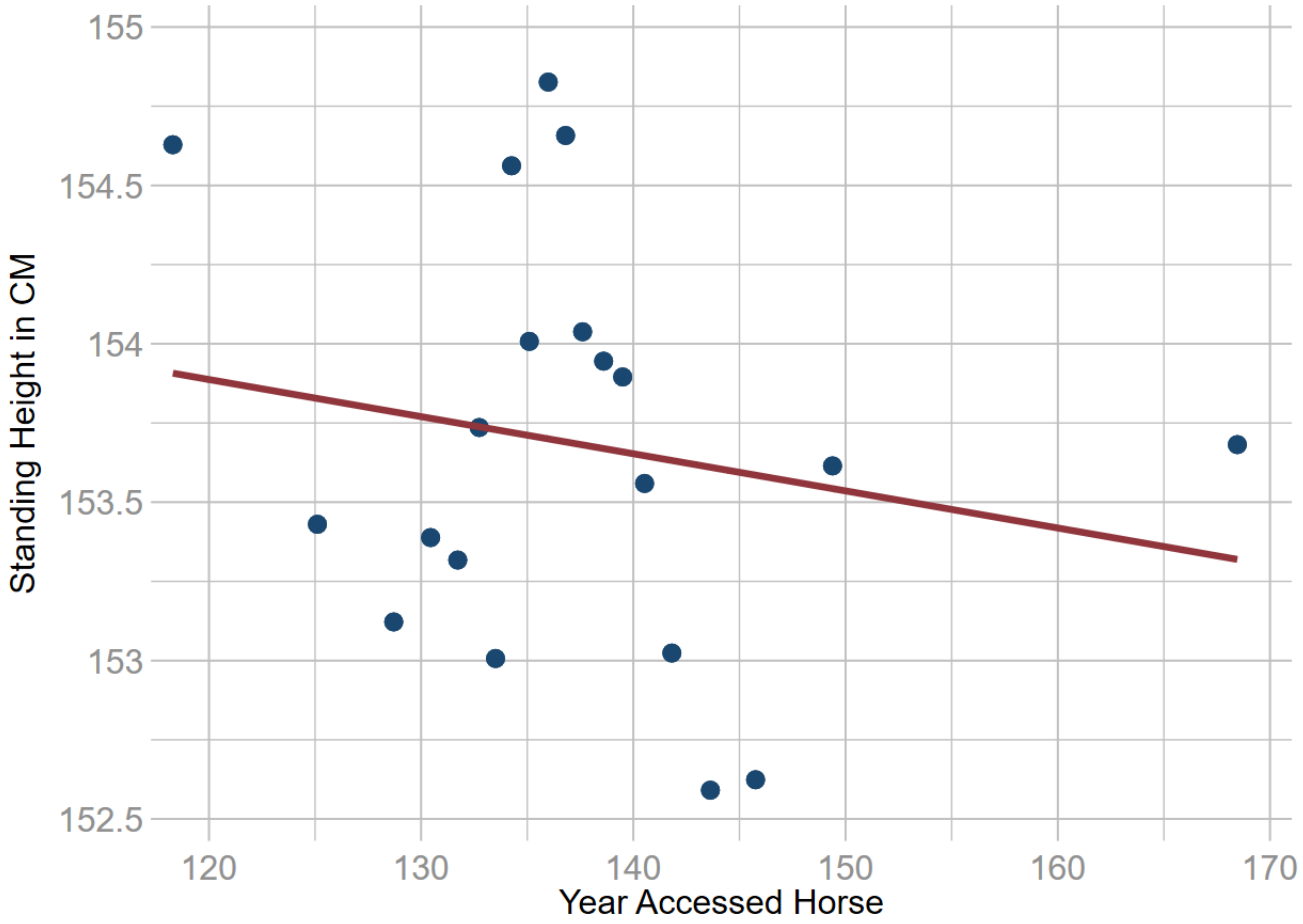


Figure 12: Height Correlations: All Controls

Table 1: Descriptive Statistics: Panel Analysis

	Obs	Means	Stand Dev.	Minimum	Maximum
Nations	82	–	–	–	–
Year	1089	1823	85.94	1540	1980
Horse Adoption Year	1089	1742	0.5713	1740	1743
Population	1089	3204	5413	19	62000
Land Cession Year	839	1897	67.34	1784	2010

Table 2: Descriptive Statistics: Cross-Section Analysis

	Obs	Means	Stand Dev.	Minimum	Maximum
Nations	115	–	–	–	–
Horse Adoption Year	161	1734	18.67	1676	1771
Gathering	161	23	16	0	60
Hunting	161	33	17	1	90
Fishing	161	26	20	0	80
Animal Husbandry	161	6	3	0	30
Agriculture	161	15	23	0	80
Homeland area in 1600	254	75,717	133,497	737.5	1,246,482
Pop density in 1600	72	0.35	0.41	0	65

Table 3: TWFE

	(1)	(2)	(3)	(4)
Horse	0.211**	0.221***	0.210***	0.138**
	(0.08)	(0.076)	(0.078)	(0.053)
Epidemic		-0.130*	-0.124*	-0.009
		(0.069)	(0.068)	(0.038)
Dummy 1776			0.208**	0.154**
			(0.085)	(0.065)
Dummy 1830			-0.179	0.086
			(0.158)	(0.053)
Constant	7.820***	7.852***	6.987***	17.325***
	(0.04)	(0.036)	(0.226)	(5.783)
Observations	448	448	448	448
R-squared	0.105	0.12	0.124	0.84
Number of nations	67	67	67	67
Nation-specific trends	No	No	No	Yes
R-square between	0.0028	0.0106	0.00361	0.0161
R-square within	0.105	0.12	0.124	0.84
R-square overall	0.0156	0.0297	0.0244	0.0118

Clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: ATT with cdid and different leads and lags

	pre	post	ATT_avg	ATT_avg_group	ATT_avg_time	p-value
Horse year	0.009	0.000	0.21*	0.21*	0.160	0.000
	(0.051)	(0.13)	(0.12)	(0.11)	(0.13)	
Horse year-1	-0.07	0.51***	0.54***	0.56***	0.52***	0.947
	(0.08)	(0.15)	(0.10)	(0.10)	(0.12)	
Horse year+1	-0.58***	0.48**	0.49*	0.58***	0.57**	0.000
	(0.14)	(0.23)	(0.25)	(0.16)	(0.26)	
Horse year-2	0.02	0.63***	0.63***	0.63***	0.63***	0.000
	(0.005)	(0.24)	(0.24)	(0.24)	(0.24)	
Horse year+2	0.008	-0.28***	-0.28***	-0.28***	-0.3	0.000
	(0.05)	(0.09)	(0.09)	(0.09)	(0.23)	

Clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: The Determinants of Horse Adoption

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Caloric Suitability Pre 1500 CE	-36.39**						
	-13.83						
Caloric Suitability Post 1500 CE	31.20**						
	-13.8						
Land Suitability	8.10*						
	-4.25						
Latitude		146.53***					
		-38.16					
Longitude		-95.66***					
		-25.09					
Latitude × Longitude		183.39***					
		-49.58					
Terrain Ruggedness			3.84				
			-3.29				
Elevation				-5.91			
				-4.69			
Area					0.78		
					-2.38		
Dist from Origin of Horse Disp						16.11***	
						-4.2	
Plains Dummy							1.3
							-2.44
Adjusted R^2	0.1	0.19	0.01	0.04	-0.01	0.15	-0.01
Observations	116	116	116	116	116	116	116

The unit of analysis is ethnic group reported by the Ethnographic Atlas. Robust standard errors are clustered at the language level. *** p<0.01, ** p<0.05, * p<0.1

Table 5: The Determinants of Horse Adoption (cont.)

	(8)	(9)	(10)
Caloric Suitability Pre 1500 CE	27.52*	31.88***	25.83*
	(14.94)	(11.09)	(12.42)
Caloric Suitability Post 1500 CE	-6.04	-10.69	-7.86
	(12.54)	(10.06)	(9.18)
Land Suitability	-0.96	-0.18	-1.36
	(2.65)	(3.77)	(4.08)
Latitude	245.21***	226.96***	188.00**
	(38.26)	(68.43)	(78.61)
Longitude	-179.20***	-174.11***	-155.52***
	(29.23)	(42.49)	(46.72)
Latitude × Longitude	315.30***	304.83***	273.25***
	(51.37)	(72.52)	(79.78)
Terrain Ruggedness		-1.40	-1.96
		(2.58)	(2.66)
Elevation		2.68	3.33
		(4.32)	(4.15)
Area		1.35	2.78
		(2.52)	(2.13)
Dist. from Origin of Horse Disp		12.13	27.99
		(20.87)	(24.79)
Plains Dummy			3.80
			(3.93)
Adjusted R^2	0.33	0.31	0.31
Observations	116	116	116

The unit of analysis is ethnic group reported by the Ethnographic Atlas. Robust standard errors are clustered at the language level. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Hunting and Horse Adoption: Two-Stage Procedure

	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Time since Horse Adoption	0.03	0.03*	0.03*	0.03*
	(0.02)	(0.02)	(0.02)	(0.02)
Time since Horse Adoption Sqr.	-0.00	-0.00	-0.00**	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)
Latitude		0.91	-3.71	-3.89
		(2.86)	(4.24)	(4.66)
Longitude		1.17	3.30	3.45
		(1.74)	(2.47)	(3.14)
Latitude × Longitude		0.19	-5.91	-6.14
		(3.87)	(5.87)	(6.26)
Terrain Ruggedness			-0.64	-0.64
			(0.48)	(0.49)
Elevation (Avg.)			0.34	0.33
			(0.27)	(0.35)
Area			0.58	0.57
			(1.09)	(1.08)
Caloric Suitability Pre 1500 CE (Avg.)				-0.04
				(0.57)
Optimal Years since Adoption	128.01	131.62	112.25	111.26
Adjusted R^2	0.01	0.21	0.29	0.29
Observations	120	120	120	120

The unit of analysis is ethnic group reported by the Ethnographic Atlas. Robust standard errors are clustered at the language level. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Other Subsistence Modes and Horse Adoption: Two-Stage Procedure

	(1) Gathering	(2) Fishing	(3) Animal Husb.	(4) Agriculture
Time since Horse Adoption	0.02	-0.01	-0.01	-0.03
	(0.02)	(0.01)	(0.00)	(0.02)
Time since Horse Adoption Sqr.	-0.00	0.00	0.00	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)
Latitude	1.03	-1.57	0.73	3.70
	(4.85)	(5.80)	(1.17)	(7.05)
Longitude	-2.20	0.11	-0.57	-0.78
	(3.21)	(3.56)	(0.80)	(4.54)
Latitude × Longitude	3.12	-3.09	0.94	5.16
	(6.32)	(7.81)	(1.48)	(8.77)
Terrain Ruggedness	0.18	0.68*	-0.03	-0.19
	(0.19)	(0.39)	(0.03)	(0.21)
Elevation (Avg.)	0.14	-0.81***	0.10***	0.24
	(0.27)	(0.28)	(0.04)	(0.35)
Area	0.34	0.05	-0.03	-0.94
	(0.66)	(0.77)	(0.06)	(0.90)
Caloric Suitability Pre 1500 CE (Avg.)	-0.33	0.33	0.11	-0.08
	(0.60)	(0.36)	(0.10)	(0.73)
Optimal Years since Adoption	69.68	120.30	79.09	86.26
Adjusted R^2	0.35	0.62	0.31	0.41
Observations	120	120	120	120

Notes: The unit of analysis is ethnic group reported by the Ethnographic Atlas. Time since horse adoption is a predicted value constructed from the regression of the year of horse adoption on geographical variables. Robust standard errors are clustered at the language level and calculated by Bootstrapp. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Hunting and Horse Adoption: Interpolated Timing of Adoption

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) 2SLS
Time since Horse Adoption	0.03**	0.03***	0.03***	0.01***	0.14***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.05)
Time since Horse Adoption Sqr.	-0.00***	-0.00***	-0.00***	-0.00***	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Latitude		0.77	-3.49	-4.09	-1.16
		(1.99)	(2.10)	(2.64)	(6.23)
Longitude		1.20	3.13**	3.66**	0.94
		(1.36)	(1.27)	(1.96)	(4.48)
Latitude × Longitude		0.15	-5.50*	-6.31*	-1.28
		(2.80)	(3.01)	(3.57)	(8.13)
Terrain Ruggedness			-0.62*	-0.61	-0.46
			(0.35)	(0.37)	(0.35)
Elevation (Avg.)			0.41*	0.37	0.48
			(0.23)	(0.30)	(0.38)
Area			0.56*	0.54*	0.48
			(0.29)	(0.27)	(0.38)
Caloric Suitability Pre 1500 CE (Avg.)				-0.16	0.82
				(0.45)	(0.87)
Optimal Years since Adoption	120.00	123.24	106.32	103.31	137.96
Adjusted R^2	0.04	0.23	0.32	0.31	
First-Stage F-Stat (First Endo. Var.)					8.54
First-Stage F-Stat (Second Endo. Var.)					18.00
Observations	120	120	120	120	120

The unit of analysis is ethnic group reported by the Ethnographic Atlas. Robust standard errors are clustered at the language level. Time since Horse Adoption is the observed year minus the interpolated year of horse adoption. First-stage F-statistics are calculated for each endogenous variables: time since horse adoption and its square. *** p<0.01, ** p<0.05, * p<0.1

Appendices

Name	Observed year	Name	Observed year	Name	Observed year	Name	Observed year	Name	Observed year	Name	Observed year
Achomawi	1860	Coeur D' Alene	1860	Jicarilla	1870	Nabesna	1930	Shasta	1860	Wiyot	1860
Acoma	1920	Comanche	1870	Kalispel	1860	Natchez	1700	Shawnee	1820	Yana	1860
Aleut	1830	Comox	1880	Kamia	1860	Navajo	1930	Shuswap	1850	Yavapai	1870
Alsea	1860	Coos	1860	Karankawa	1800	Netsilk Eskimo	1920	Siuslaw	1860	Yokuts	1850
Arapaho	1860	Copper Eskimo	1920	Karok	1860	Nez Perce	1850	Slave	1860	Yuchi	1750
Arikara	1860	Cowichan	1880	Kaska	1920	Nootka	1880	Tahltan	1870	Yuki	1850
Assiniboin	1870	Cree	1850	Kawaiisu	1860	Ojibwa	1940	Takelma	1860	Yurok	1850
Atakapa	1870	Creek	1750	Kickapoo	1870	Omaha	1850	Taos	1890	Zuni	1910
Baffinland Eskimo	1880	Crow	1870	Kiowa	1860	Opata	1550	Tepehuan	1920		
Bannock	1860	Delaware	1700	Kiowa Apache	1870	Oto	1870	Teton	1870		
Beaver	1850	Diegueno	1850	Klallam	1860	Ottawa	1650	Tewa	1900		
Bellabella	1880	Dogrib	1860	Klamuth	1860	Paiute	1860	Thompson	1860		
Bellacoola	1880	Flathead	1860	Klikitat	1870	Panamint	1850	Tillamook	1870		
Blackfoot	1850	Fox	1830	Kutchen	1880	Papago	1930	Timucua	1560		
Cahuilla	1870	Gabrielino	1770	Kutenai	1880	Paviotso	1870	Tlingit	1880		
Caribou Eskimo	1900	Gosiute	1860	Kwakiutl	1890	Pawnee	1860	Tolowa	1870		
Catawba	1920	Gros Ventre	1870	Labrador Eskimo	1890	Pima	1840	Tsimshian	1880		
Cherokee	1750	Haida	1890	Lillooet	1860	Pomo	1860	Tubatulabai	1850		
Cheyenne	1960	Havasupai	1880	Lipan	1880	Ponca	1850	Twana	1850		
Chilcotin	1880	Hidatsa	1860	Luiseno	1860	Potawatomi	1760	Umatilla	1860		
Chimariko	1860	Hopi	1920	Maidu	1850	Quileute	1880	Wailaki	1860		
Chinook	1850	Huichol	1920	Mandan	1830	Quinault	1860	Walapai	1870		
Chipewyan	1880	Hupa	1860	Maricopa	1850	Salina	1770	Wappo	1860		
Chiricahua	1880	Huron	1640	Menomini	1870	Sanpoil	1870	Washoe	1850		
Chitimacha	1750	Iglulik Eskimo	1920	Mescalero	1870	Santee	1850	Western Apache	1870		
Choctaw	1760	Ingalik	1880	Miami	1720	Sarsi	1880	Wichita	1860		
Chumash	1800	Iowa	1870	Micmac	1700	Satudene	1860	Wind River	1860		
Coahuiltec	1760	Iroquois	1750	Mohave	1850	Sekani	1880	Winnebago	1850		
Cochimi	1890	Isleta	1920	Mono	1870	Seri	1900	Wintun	1860		
Cocopa	1850	Jemez	1920	Montagnais	1880	Serrano	1870	Wishram	1860		

Figure 13: Nations' Name and Observed Year (source: Murdoch, 1967)