A Theoretical Framework: Further Insights

Here, we complement the theoretical framework presented in the paper by discussing special cases in Section A.1, and providing a rationale for Theorem 1 in Section A.2.

A.1 Special Cases

The main paper lays out the emperor's utility maximisation problem as

$$\max_{N_c,N_p,locations} U\Big[\underbrace{\sum_{c=1}^{N_c}\sum_{i=1}^{N_{ic}}(1-\tau D_{ci})\lambda P_i}_{+}, \underbrace{\sum_{p=1}^{N_p}\sum_{c=1}^{N_{cp}} D_{pc}}_{-}, \underbrace{\sum_{p=1}^{N_p}\sum_{i=1}^{N_{ip}}(D_{pi}M_i)^m}_{-}\Big] \tag{(A-1)}$$

subject to the county seat constraints

$$G_c + E_c \leq \sum_{i=1}^{N_{ic}}(1-\tau D_{ci})\lambda P_i \quad \text{for} \quad 1 \leq c \leq N_c \tag{(A-2)}$$

$$E_c \geq 0 \quad \text{for} \quad 1 \leq c \leq N_c \tag{(A-3)}$$

and prefecture seat constraints

$$G_p \leq \sum_{c=1}^{N_{pc}}(E_c - \tau \mathcal{I}_{E_{pq}>0} D_{pc}E_c) + \sum_{q\neq p}^{N_p}(E_{pq} - \tau D_{pq}E_{pq}) \quad \text{for} \quad 1 \leq p \leq N_p \tag{(A-4)}$$

No closed-form solution for the number of locations of the county and prefecture seats is possible. But it is insightful to consider special cases which allow us to further simplify the expressions involved and highlight various features of the framework.

Assumption 4. *All county seats carry the same cost G_c and they all export the same amount E_c to the prefecture seats*

$$G_c = \bar{G} \quad \textit{for} \quad 1 \leq c \leq N_c \tag{(A-5)}$$

$$E_c = \bar{E} \quad \textit{for} \quad 1 \leq c \leq N_c \tag{(A-6)}$$

Imposing this additional assumption allows us to write the aggregate-level resources constraint

$$\sum_{c=1}^{N_c}(G_c + E_c) \leq \sum_{c=1}^{N_c}\sum_{i=1}^{N_{ic}}(1-\tau D_{ci})\lambda P_i \tag{(A-7)}$$

as

$$
\begin{aligned}
N_c \cdot (\bar{G} + \bar{E}) &\leq \sum_{i=1}^{N_i} (1 - \tau D_{ci}) \lambda P_i && \text{((A-8))} \\
N_c &\leq \frac{\sum_{i=1}^{N_i} (1 - \tau D_{ci}) \lambda P_i}{\bar{G} + \bar{E}} && \text{((A-9))}
\end{aligned}
$$

Note that this is not a closed-form solution for N_c as the distance D_{ci} between cell i and prefecture seat c depends on it. Yet, it is independent from considerations at the prefecture seat level and illustrates the trade-off: A higher population P_i will make it worthwhile to have more county seats, by boosting tax revenues, while higher expenses in the form of $\bar{G}$ and $\bar{E}$ will lead to fewer county seats.

At the prefecture seat level, Assumption 4 allows to simplify

$$
\sum_{p=1}^{N_p} G_p \leq \sum_{p=1}^{N_p} \sum_{c=1}^{N_{pc}} (E_c - \tau D_{pc} E_c) - \sum_{q<p}^{N_p} \tau D_{pq} |E_{pq}| \qquad \text{((A-10))}
$$

to

$$
\begin{aligned}
N_p \bar{G} &\leq N_p \bar{E} \sum_{i=c}^{N_{pc}} (1 - \tau D_{pc}) - \sum_{q<p}^{N_p} \tau D_{pq} |E_{pq}| && \text{((A-11))} \\
\bar{G} &\leq \bar{E} \sum_{i=c}^{N_{pc}} (1 - \tau D_{pc}) - \frac{1}{N_p} \sum_{q<p}^{N_p} \tau D_{pq} |E_{pq}| && \text{((A-12))}
\end{aligned}
$$

The number N_p of prefecture seats can be determined implicitly from this equation. It balances the costs of each prefecture seat with the receipts, net of all transport costs.

A.2 Rationale for Theorem 1

We can show the effects described in Theorem 1 based on the equations laid out before.

(a) In areas with favorable geography, there are more county seats.

Rationale: By Assumption 1, favorable geography A_i in cell i increases population P_i in this cell. The maximisation problem eq. (8) depends negatively on the population-weighted travel time distance $D_{ci} P_i$ between cell i and the nearest county seat. Decreasing the distance to populous cells has the strongest effect on utility, leading to more county seats in areas with favorable geography.

(b) The location of prefecture seats is determined to a lesser extent and only indirectly by the presence of geographical features.

Rationale: In contrast to the direct geographical effect of geography on county seat density via population, the prefecture seat locations are chosen primarily on military considerations. While favorable geography A_i increases P_i according to Assumption 1, the latter term in eq. (8) does not include P_i. There is only an indirect channel through which geography affects the location of prefecture seats: County seat density is determined by favorable geography according to (a), and according to eq. (13), it is advantageous for prefecture seats to locate close to county seats to minimise the transport costs of the transfers E_c.

(c) There are more prefecture seats per county seat in regions that are prone to military invasion.

Rationale: Military threats affect the location of prefecture seats directly and that of county seats only indirectly. A military threat M_i in cell i has a negative effect on utility in eq. (8), which is reinforced by the non-linearity of $m > 1$. It can only be compensated by decreasing the distance D_{ip} to the nearest prefecture seat. This leads to an increase in the number of prefecture seats, while the number of county seats will either stay the same or might decrease through an indirect effect: According to Assumption 1, the total threat level increases in response to the increase in M_i, decreasing population P_i. Following the mechanism discussed in (a), a decreased population leads to fewer county seats, reinforcing the result that there are more prefecture seats per county seat in regions that are prone to military invasion.

(d) If the population grows at a higher rate than the costs of maintaining county seats, the optimal number of county seats will increase.

Rationale: According to eq. (8), the optimal number N_c of county seats will always be N_i, but it is constrained by the financing restrictions eq. (9) and eq. (10) for each county seat. If P_i increases at a higher rate than G_c, resources for establishing additional county seats are available, as specified in the aggregate budget constraint eq. (11).

B Additional Information on Geographical Data

In this part of the appendix, we present additional information on the data used in our empirical analysis. It complements Section 3.1. Table B-1 provides the detailed data sources of the geographical variables. Table B-2, Table B-3 and Table B-4 contain the summary statistics of, respectively, the dominant soil type, landform, and lithology.

Table B-1: Geography Data Sources

Variable	Source	Original Format	Derivations
Distance from equator	Own computations	Raster (5 arc minutes)	Distance calculation between pixel centroid and equator using latitude
Distance from coast	Wessel and Smith (1996, 2017)	Polygon	Distance calculation between land pixel centroid and centroid of nearest ocean pixel
Distance from river	Natural Earth (2019)	Spatial lines	Distance calculation between pixel centroid and nearest river's spatial line
Ruggedness	Nunn and Puga (2012)	Raster (30 arc seconds)	Grid cells containing index measuring elevation difference between grid cells in mm aggregated to 5 arc minute level using averages
Temperature	Matsuura and Willmott (2018a)	Raster (30 arc minutes)	Averages of monthly temperature data from 1900 - 1950 disaggregated to 5 arc minute level using bilinear interpolation
Precipitation	Matsuura and Willmott (2018b)	Raster (30 arc minutes)	Averages of annual precipitation in mm from 1900 - 1950 disaggregated to 5 arc minute level using bilinear interpolation
Elevation	Danielson and Gesch (2011)	Raster (7.5 arc seconds)	Grid cells aggregated to 5 arc minute level using averages
Dominant soil type	Dijkshoorn et al. (2008)	Polygon	Polygons converted to grid cells
Landform	Dijkshoorn et al. (2008)	Polygon	Polygons converted to grid cells
Lithology	Dijkshoorn et al. (2008)	Polygon	Polygons converted to grid cells

Notes: Our regressions use this data converted to an equal-area Mollweide projection. The average monthly temperature does not entail more information on the time dimension than average annual precipitation does. Both variables sum up observations from 612 months between 1900 and 1951. In the case of temperature, we then divide that sum by 612 and in the case of precipitation by 51. Cumulating precipitation measured in mm over a year makes sense, while cumulating temperature measured in degrees Celsius does not.

Table B-2: Dominant Soil Summary Statistics

Abbreviation	Full Name	Frequency	Abbreviation	Full Name	Frequency
ACf	ferric acrisols	296	HSs	terric histosols	5
ACh	haplic acrisols	596	KSh	hapic kastanozems	108
ACp	plinthic acrisols	10	KSk	caleic kastanozems	25
ACu	humic acrisols	398	KSl	luvic kastanozems	96
ALf	ferric alisols	24	LP	leptosols	4
ALh	haplic alisols	367	LPd	dystric leptosols	8
ALp	plinthic alisols	12	LPe	eutric leptosols	200
ANh	haplic andosols	3	LPi	gelic leptosols	409
ARb	cambic arenosols	86	LPk	rendzic leptosols	79
ARc	calcaric arenosols	87	LPm	mollic leptosols	204
ARh	haplic arenosols	139	LVa	albic luvisols	73
ATa	aric anthrosols	8	LVg	gleyic luvisols	118
ATc	cumulic anthrosols	628	LVh	haplic luvisols	1360
ATf	fimic anthrosols	7	LVj	stagnic luvisols	2
CHg	gleyic chernozems	21	LVk	calcic luvisols	117
CHh	haplic chernozems	53	LVx	chromic luvisols	94
CHk	calcic chernozems	82	LXa	albic lixisols	1
CHl	luvic chernozems	19	LXf	ferric lisols	17
CLh	haplic calcisols	54	NTu	humic nitosols	13
CLl	luvic calcisols	50	PDd	dystric podzoluvisols	12
CLp	petric calcisols	8	PHc	calcaric phaeozems	68
CMc	calcaric cambisols	627	PHg	gleyic phaeozems	110
CMd	dystric cambisols	377	PHh	haplic phaeozems	281
CMe	eutric cambisols	142	PHj	stagnic phaeozems	8
CMg	gleyic cambisols	56	PLd	dystric planosols	7
CMi	gelic cambisols	23	PLe	eutric planosols	62
CMo	ferralic cambisols	48	RGc	calcaric regosols	161
CMu	humic cambisols	1	RGd	dystric regosols	73
CMx	chromic cambisols	66	RGe	eutric regosols	85
FLc	calcaric fluvisols	435	RK	rock, rock outcrop	18
FLe	eutric fluvisols	43	SC	solonchaks	1
FLs	salic fluvisols	80	SCg	gleyic solonchaks	6
FRh	heplc ferralsols	32	SCh	haplic solonchaks	24
FRx	xanthic ferralsols	17	SCk	calcic solonchaks	12
GG	glaciers, ice	17	SCm	mollic solonchaks	30
GLe	eutric gleysols	15	SCn	sodic solonchaks	1
GLi	geic gleysols	2	SCy	gypsic solonchaks	10
GLk	caleic gleysols	24	SNg	gleyic solonetz	10
GLm	molic gleysols	155	SNh	haplic solonetz	1
GLt	thionic gleysols	1	ST	salt flats	35
GRh	hapic greyzems	73	UR	urban areas	4
GYh	heplc gypsisols	5	VRd	distric vertsols	6
GYk	calcic gypsisols	93	VRe	eutric vertsols	35
GYl	luvic gypsisols	84	VRk	calcic vertsols	1
GYp	petric gypsisols	154	WR	inland water, lakes	78

Notes: The frequency refers to the number of pixels in the baseline setting, which are 21.99 x 28.53 km in size.

Table B-3: Landform Summary Statistics

Abbreviation	Full Name	Frequency	Abbreviation	Full Name	Frequency
LP	plain	4,160	SP	dissected plain	64
LP wet	plain wet	10	TH	high-gradient hill	330
SH	medium-gradient hill	3,153	TM	high-gradient mountain	1,657
SM	medium-gradient mountain	138	WR	inland water, lakes	78

Notes: The frequency refers to the number of pixels in the baseline setting, which are 21.99 x 28.53 km in size.

Table B-4: Lithology Summary Statistics

Abbr.	Full Name	Freq.	Abbr.	Full Name	Freq.
GG	glaciers, ice	17	SC4	shale	697
IA	acid igneous rock	3	SC5	ironstone	35
IA1	granite	1,505	SC7		17
IA2	grano-diorite	5	SO1	limestone, other carbonate rocks	697
IA4	rhyolite	4	SO2	marl and other mixtures	71
IB2	basalt	96	SO3	coals, bitumen and related rocks	35
IB3	dolerite	1	ST	salt flats	35
II1	andesite, trachyte, phonolite	95	UA1	redeposited natural material	1
II2	diorite-syenite	3	UE1	loess	1,445
IP2	volcanic scoria/ breccia	13	UE1/UR1	loess/ bauxite, laterite	3
IP4	ignimbrite	166	UE2	sand	291
MA	acid metamorphic rock	2	UF	fluvial	1,667
MA1	quartzite	32	UF/UL	fluvial/ lacustrine unconsolidated rock	60
MA1/SC2	quartzite/ sandstone, greywacke, arkose	12	UF/UM	fluvial/ marine unconsolidated rock	15
MA2	gneiss, migmatite	158	UF1	sand and gravel	14
MA3/MB1	slate, phyllite (pelitic rocks)	208	UF2	clay, silt and loam	20
MA4/MB2	schist	74	UG	glacial	23
MB1	slate, phyllite (pelitic rocks)	201	UL	lacustrine	186
MB1/MB2	slate, phyllite (pelitic rocks)/ schist	222	UL/UM	lacustrine unconsolidated rock/	
RK	rock outcrop	18		marine unconsolidated rock	1
SC	clastic sedimentary rock	96	UL2	silt and clay	1
SC1	conglomerate, breccia	89	UM	marine unconsolidated rock	22
SC16	glacial sedimentary environments	49	UO	organic	1
SC2	sandstone, greywacke, arkose	712	UR	urban areas	4
SC2/SC4	sandstone, greywacke, arkose/ shale	27	UR1	bauxite, laterite	203
SC3	siltstone, mudstone, claystone	160	WR	lakes, permanent water	78

Notes: The frequency refers to the number of pixels in the baseline setting, which are 21.99 x 28.53 km in size.

C Supplementary Empirical Results

Here we provide additional results and robustness checks to supplement the empirical analysis in Section 4 and Section 5 of the paper.

C.1 Supplementary OLS Results on Soil Fixed Effects

Table 2 and Table 3, the baseline local geography regression tables, do not report coefficient estimates on the categorical soil variables. The soil variables are included as fixed effects and thus accounted for via demeaning. In this section, we repeat the estimations with the soil variables included as indicators. Such modification, of course, leaves the coefficient estimates displayed in the baseline tables unchanged and simply adds further rows to the output. Table C-1 corresponds to Table 2 and Table C-2 to Table 3. Unsurprisingly, the effects vary considerably across soil types. Random forests are able to explore these categories in much larger complexity than the linear regressions do.

Table C-1: Local Geography County Seat Regressions with Soil Coefficients

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	−0.23	−0.37	−0.37	−0.70***	−0.93***	−1.22***	−0.80***	−0.76***	−0.85***	−0.99***	−0.87***
	(0.18)	(0.25)	(0.24)	(0.25)	(0.29)	(0.32)	(0.28)	(0.26)	(0.25)	(0.27)	(0.26)
Dist. Coast	−0.14	−0.32	0.07	0.45***	0.80***	0.86***	0.89***	0.75***	0.30*	0.49***	0.31*
	(0.16)	(0.21)	(0.18)	(0.17)	(0.22)	(0.24)	(0.23)	(0.22)	(0.18)	(0.19)	(0.18)
Dist. River	0.07	−0.13	−0.17	−0.66	−1.49***	−0.79	−0.25	−0.10	−0.36	−0.58	−0.26
	(0.37)	(0.44)	(0.41)	(0.42)	(0.47)	(0.55)	(0.55)	(0.54)	(0.52)	(0.53)	(0.49)
Ruggedness	−0.50**	−0.57*	−0.77***	−1.28***	−0.66*	−1.21***	−1.35***	−1.33***	−1.32***	−1.42***	−1.42***
	(0.20)	(0.31)	(0.29)	(0.27)	(0.35)	(0.39)	(0.39)	(0.37)	(0.36)	(0.37)	(0.35)
Temperature	0.58***	0.81***	0.61**	0.27	0.12	−0.48	0.00	0.10	0.07	0.07	0.13
	(0.19)	(0.27)	(0.25)	(0.23)	(0.29)	(0.31)	(0.27)	(0.25)	(0.24)	(0.26)	(0.26)
Temperature^2	−0.25	0.35	−0.14	−1.40	2.42**	3.96***	2.18*	1.36	0.94	0.42	0.20
	(0.79)	(1.10)	(1.02)	(1.00)	(1.16)	(1.38)	(1.25)	(1.22)	(1.21)	(1.26)	(1.21)
Precipitation	−1.15***	−1.86***	−0.84**	−0.16	−1.31**	−0.15	0.55	0.59	0.14	0.40	0.45
	(0.36)	(0.42)	(0.36)	(0.38)	(0.52)	(0.56)	(0.52)	(0.48)	(0.43)	(0.46)	(0.44)
Precipitation^2	1.86***	3.07***	1.15	0.10	2.04*	−0.32	−1.58	−1.69	−0.92	−1.55	−1.55*
	(0.70)	(0.82)	(0.73)	(0.78)	(1.06)	(1.16)	(1.13)	(1.06)	(0.90)	(0.97)	(0.91)
Elevation	−1.95**	−1.33	−2.01*	−4.12***	−6.42***	−7.31***	−6.05***	−5.39***	−4.23***	−5.17***	−4.56***
	(0.87)	(1.14)	(1.07)	(1.10)	(1.28)	(1.38)	(1.23)	(1.17)	(1.15)	(1.25)	(1.19)
Dominant Soil Type											
ACf	0.17*	0.29**	0.22	0.45***	0.59***	0.65***	0.30*	0.30*	0.40***	0.51***	0.48***
	(0.10)	(0.14)	(0.13)	(0.14)	(0.18)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.15)
ACh	0.22**	0.35**	0.29**	0.50***	0.52***	0.59***	0.39**	0.38**	0.52***	0.60***	0.55***
	(0.10)	(0.14)	(0.13)	(0.14)	(0.17)	(0.19)	(0.17)	(0.15)	(0.15)	(0.16)	(0.15)
ACp	0.63***	0.70***	0.62***	0.67***	0.70***	0.77***	0.51**	0.50**	0.54***	0.69***	0.62***
	(0.18)	(0.20)	(0.20)	(0.20)	(0.21)	(0.23)	(0.21)	(0.20)	(0.18)	(0.22)	(0.22)
ACu	0.24**	0.37***	0.30**	0.47***	0.53***	0.60***	0.33**	0.32**	0.48***	0.56***	0.51***
	(0.10)	(0.14)	(0.13)	(0.14)	(0.17)	(0.19)	(0.16)	(0.15)	(0.15)	(0.15)	(0.15)
ALf	0.26**	0.38**	0.26*	0.47***	0.74***	0.72***	0.55***	0.54***	0.67***	0.71***	0.71***
	(0.12)	(0.16)	(0.14)	(0.16)	(0.19)	(0.20)	(0.19)	(0.18)	(0.17)	(0.18)	(0.17)
ALh	0.23**	0.35**	0.25*	0.41***	0.56***	0.58***	0.28*	0.27*	0.44***	0.56***	0.52***
	(0.10)	(0.14)	(0.13)	(0.14)	(0.17)	(0.19)	(0.16)	(0.15)	(0.15)	(0.16)	(0.16)

Table C-1 (Continued)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
ALp	0.17	0.24	0.18	0.51***	0.44**	0.53**	0.35*	0.35*	0.47***	0.52***	0.47**
	(0.14)	(0.17)	(0.16)	(0.18)	(0.18)	(0.21)	(0.19)	(0.19)	(0.18)	(0.19)	(0.19)
ANh	0.19*	0.29**	0.20	0.35**	0.46**	0.48**	0.22	0.21	0.28*	0.34**	0.32**
	(0.10)	(0.15)	(0.14)	(0.15)	(0.18)	(0.20)	(0.17)	(0.16)	(0.16)	(0.17)	(0.16)
ARb	0.23**	0.33**	0.26*	0.41***	0.54***	0.62***	0.35**	0.34**	0.44***	0.51***	0.44***
	(0.11)	(0.15)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
ARc	0.22**	0.32**	0.26*	0.42***	0.57***	0.63***	0.36**	0.34**	0.44***	0.51***	0.45***
	(0.11)	(0.15)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
ARh	0.21**	0.31**	0.24*	0.40***	0.50***	0.56***	0.30*	0.29*	0.41***	0.48***	0.45***
	(0.10)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.15)	(0.15)	(0.16)	(0.16)
ATa	0.12	0.20	0.14	0.31**	0.40**	0.50***	0.24	0.23	0.35**	0.40***	0.36**
	(0.10)	(0.14)	(0.13)	(0.14)	(0.17)	(0.19)	(0.16)	(0.15)	(0.14)	(0.15)	(0.15)
ATc	0.26***	0.40***	0.32**	0.50***	0.61***	0.68***	0.41**	0.41***	0.50***	0.59***	0.56***
	(0.10)	(0.14)	(0.13)	(0.14)	(0.17)	(0.19)	(0.16)	(0.15)	(0.15)	(0.16)	(0.16)
ATf	0.54***	0.62***	0.41**	0.57***	1.08***	1.12***	0.86***	0.86***	1.00***	1.06***	1.02***
	(0.13)	(0.16)	(0.17)	(0.20)	(0.24)	(0.21)	(0.19)	(0.18)	(0.16)	(0.17)	(0.17)
CHg	0.16	0.28**	0.21	0.36**	0.51***	0.57***	0.28*	0.27*	0.40***	0.47***	0.42***
	(0.10)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.15)	(0.15)	(0.16)	(0.16)
CHh	0.22**	0.34**	0.26*	0.43***	0.59***	0.65***	0.37**	0.35**	0.47***	0.54***	0.47***
	(0.10)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.15)	(0.15)	(0.16)	(0.16)
CHk	0.21*	0.32**	0.27*	0.41***	0.57***	0.60***	0.32**	0.32**	0.45***	0.51***	0.48***
	(0.11)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.16)	(0.15)	(0.15)	(0.16)	(0.16)
CHl	0.29**	0.42***	0.34**	0.50***	0.59***	0.70***	0.43**	0.41**	0.48***	0.56***	0.50***
	(0.12)	(0.15)	(0.15)	(0.15)	(0.17)	(0.20)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
CLh	0.15	0.30**	0.24*	0.39***	0.47***	0.56***	0.29*	0.30**	0.39***	0.45***	0.42***
	(0.10)	(0.15)	(0.14)	(0.14)	(0.17)	(0.19)	(0.16)	(0.15)	(0.14)	(0.15)	(0.15)
CLl	0.15	0.26*	0.20	0.38***	0.49***	0.59***	0.31*	0.30*	0.43***	0.50***	0.45***
	(0.10)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.15)	(0.15)	(0.16)	(0.16)
CLp	0.19*	0.31**	0.24*	0.41***	0.54***	0.61***	0.34**	0.33**	0.45***	0.52***	0.47***
	(0.10)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
CMc	0.32***	0.47***	0.39***	0.52***	0.75***	0.79***	0.53***	0.50***	0.60***	0.66***	0.62***
	(0.11)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
CMd	0.24**	0.37***	0.28**	0.44***	0.61***	0.67***	0.37**	0.33**	0.47***	0.56***	0.51***
	(0.10)	(0.14)	(0.13)	(0.14)	(0.17)	(0.18)	(0.16)	(0.15)	(0.15)	(0.16)	(0.15)
CMe	0.23**	0.39***	0.30**	0.48***	0.62***	0.65***	0.38**	0.37**	0.51***	0.58***	0.52***
	(0.10)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.16)	(0.15)	(0.15)	(0.16)	(0.15)
CMg	0.25**	0.43***	0.35**	0.46***	0.64***	0.65***	0.38**	0.39**	0.54***	0.61***	0.59***
	(0.11)	(0.15)	(0.14)	(0.14)	(0.18)	(0.20)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
CMi	0.32***	0.45***	0.35**	0.52***	0.72***	0.75***	0.48***	0.43***	0.54***	0.63***	0.58***
	(0.11)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
CMo	0.17*	0.25*	0.20	0.38***	0.48***	0.57***	0.28*	0.31**	0.46***	0.55***	0.51***
	(0.10)	(0.14)	(0.13)	(0.14)	(0.17)	(0.19)	(0.17)	(0.15)	(0.15)	(0.16)	(0.16)
CMu	0.17*	0.28**	0.18	1.32***	0.43***	1.44***	0.17	0.15	0.28*	0.35**	0.30**
	(0.10)	(0.14)	(0.13)	(0.13)	(0.16)	(0.18)	(0.16)	(0.15)	(0.14)	(0.15)	(0.15)
CMx	0.20**	0.40***	0.33**	0.48***	0.49***	0.64***	0.37**	0.35**	0.45***	0.51***	0.45***
	(0.10)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.16)	(0.16)	(0.16)
FLc	0.40***	0.57***	0.48***	0.54***	0.74***	0.77***	0.51***	0.50***	0.60***	0.66***	0.61***
	(0.11)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.15)	(0.15)	(0.16)	(0.16)
FLe	0.38***	0.50***	0.47***	0.66***	0.80***	0.76***	0.55***	0.52***	0.60***	0.62***	0.60***
	(0.14)	(0.17)	(0.16)	(0.17)	(0.18)	(0.20)	(0.18)	(0.16)	(0.16)	(0.17)	(0.17)
FLs	0.30***	0.51***	0.39***	0.45***	0.69***	0.80***	0.50***	0.50***	0.61***	0.68***	0.62***
	(0.11)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
FRh	0.12	0.17	0.11	0.32**	0.43**	0.45**	0.17	0.12	0.20	0.29*	0.25
	(0.10)	(0.14)	(0.13)	(0.14)	(0.18)	(0.19)	(0.18)	(0.15)	(0.15)	(0.16)	(0.15)
FRx	0.19*	0.29*	0.21	0.37***	0.41**	0.40**	0.17	0.17	0.36**	0.43***	0.39**
	(0.10)	(0.15)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)

Table C-1 (Continued)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
GG	0.32***	0.45***	0.38***	0.54***	0.75***	0.83***	0.57***	0.53***	0.60***	0.68***	0.62***
	(0.11)	(0.14)	(0.14)	(0.15)	(0.18)	(0.20)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
GLe	0.41***	0.47***	0.38**	0.64***	0.71***	0.75***	0.42**	0.42**	0.57***	0.60***	0.53***
	(0.14)	(0.17)	(0.16)	(0.18)	(0.20)	(0.21)	(0.21)	(0.20)	(0.19)	(0.20)	(0.20)
GLi	0.35***	0.52***	0.39***	0.52***	0.70***	0.63***	0.36**	0.36**	0.49***	0.58***	0.53***
	(0.11)	(0.14)	(0.14)	(0.15)	(0.18)	(0.20)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
GLk	0.40***	0.51***	0.41**	0.51***	0.74***	0.63***	0.36**	0.39**	0.50***	0.56***	0.54***
	(0.13)	(0.16)	(0.16)	(0.16)	(0.19)	(0.20)	(0.18)	(0.17)	(0.16)	(0.17)	(0.18)
GLm	0.23**	0.34**	0.25*	0.39***	0.57***	0.60***	0.32*	0.31*	0.44***	0.50***	0.42***
	(0.11)	(0.15)	(0.14)	(0.15)	(0.18)	(0.20)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
GLt	0.16	0.22	0.14	0.34**	1.37***	1.30***	1.09***	1.07***	0.14	0.20	0.12
	(0.10)	(0.15)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.16)	(0.17)	(0.17)
GRh	0.28***	0.43***	0.35**	0.50***	0.64***	0.70***	0.44***	0.44***	0.56***	0.64***	0.56***
	(0.11)	(0.15)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
GYh	0.12	0.41*	0.36	0.53**	0.62**	0.71***	0.25	0.24	0.37***	0.44***	0.59**
	(0.10)	(0.22)	(0.22)	(0.23)	(0.25)	(0.26)	(0.16)	(0.15)	(0.14)	(0.15)	(0.24)
GYk	0.12	0.21	0.14	0.33**	0.44**	0.52***	0.23	0.22	0.36**	0.42***	0.38**
	(0.11)	(0.14)	(0.14)	(0.14)	(0.18)	(0.20)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
GYl	0.12	0.23	0.17	0.35**	0.44**	0.52***	0.24	0.22	0.36**	0.42***	0.36**
	(0.10)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.15)	(0.15)	(0.16)	(0.16)
GYp	0.10	0.19	0.11	0.29**	0.36**	0.47**	0.19	0.18	0.34**	0.39**	0.35**
	(0.10)	(0.14)	(0.14)	(0.14)	(0.18)	(0.20)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
HSs	0.21*	0.30**	0.21	0.36**	0.55***	0.58***	0.31*	0.29*	0.40***	0.45***	0.39**
	(0.11)	(0.15)	(0.14)	(0.15)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
KSh	0.18*	0.30**	0.23*	0.41***	0.55***	0.63***	0.36**	0.34**	0.44***	0.52***	0.48***
	(0.10)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.16)	(0.15)	(0.15)	(0.16)	(0.16)
KSk	0.17*	0.29**	0.22	0.40***	0.53***	0.61***	0.33**	0.31**	0.44***	0.51***	0.45***
	(0.10)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.15)	(0.15)	(0.16)	(0.16)
KSl	0.20*	0.31**	0.25*	0.41***	0.56***	0.64***	0.36**	0.34**	0.44***	0.53***	0.47***
	(0.10)	(0.15)	(0.14)	(0.14)	(0.17)	(0.19)	(0.16)	(0.15)	(0.15)	(0.16)	(0.16)
LP	0.44*	0.54**	0.45*	0.58**	0.74***	0.98***	0.97***	0.95***	0.83***	0.86***	0.81***
	(0.24)	(0.27)	(0.27)	(0.27)	(0.28)	(0.31)	(0.27)	(0.27)	(0.31)	(0.31)	(0.31)
LPd	0.19*	0.31**	0.22	0.38***	0.54***	0.49**	0.21	0.18	0.46***	0.53***	0.47***
	(0.11)	(0.15)	(0.14)	(0.14)	(0.18)	(0.19)	(0.17)	(0.16)	(0.18)	(0.18)	(0.18)
LPe	0.28***	0.41***	0.31**	0.47***	0.66***	0.67***	0.39**	0.37**	0.49***	0.58***	0.53***
	(0.11)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
LPi	0.34***	0.48***	0.38***	0.54***	0.74***	0.77***	0.50***	0.47***	0.57***	0.65***	0.59***
	(0.11)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
LPk	0.25**	0.39***	0.33**	0.48***	0.68***	0.76***	0.38**	0.37**	0.54***	0.59***	0.52***
	(0.10)	(0.14)	(0.14)	(0.15)	(0.18)	(0.20)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
LPm	0.31***	0.43***	0.33**	0.49***	0.67***	0.72***	0.44***	0.40**	0.51***	0.59***	0.54***
	(0.11)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
LVa	0.22**	0.37**	0.29**	0.43***	0.58***	0.63***	0.35**	0.34**	0.45***	0.53***	0.45***
	(0.11)	(0.15)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
LVg	0.38***	0.57***	0.48***	0.51***	0.81***	0.80***	0.50***	0.49***	0.61***	0.67***	0.60***
	(0.12)	(0.15)	(0.14)	(0.15)	(0.18)	(0.20)	(0.17)	(0.16)	(0.16)	(0.17)	(0.16)
LVh	0.26**	0.41***	0.32**	0.48***	0.64***	0.68***	0.41**	0.39**	0.51***	0.59***	0.53***
	(0.10)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.15)	(0.15)	(0.16)	(0.16)
LVj	0.31***	0.48***	0.38***	0.54***	0.68***	0.72***	0.46***	0.44***	0.57***	0.65***	0.58***
	(0.11)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
LVk	0.23**	0.39***	0.32**	0.51***	0.69***	0.75***	0.46***	0.43***	0.52***	0.60***	0.56***
	(0.10)	(0.14)	(0.13)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
LVx	0.25**	0.40***	0.30**	0.47***	0.64***	0.70***	0.36**	0.35**	0.64***	0.72***	0.53***
	(0.10)	(0.14)	(0.13)	(0.14)	(0.17)	(0.19)	(0.17)	(0.15)	(0.16)	(0.17)	(0.17)
LXa	0.22**	0.33**	0.23*	0.37***	0.54***	0.53***	0.27*	0.25*	0.36***	0.40***	0.34**
	(0.10)	(0.13)	(0.13)	(0.14)	(0.16)	(0.18)	(0.16)	(0.14)	(0.14)	(0.15)	(0.14)

Table C-1 (Continued)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
LXf	0.21**	0.44***	0.40**	0.49***	0.61***	0.70***	0.37**	0.35**	0.47***	0.52***	0.57***
	(0.10)	(0.17)	(0.18)	(0.17)	(0.18)	(0.21)	(0.18)	(0.17)	(0.17)	(0.18)	(0.18)
NTu	0.10	0.22	0.16	0.26*	0.35**	0.36*	0.08	0.08	0.25	0.31*	0.30*
	(0.10)	(0.15)	(0.14)	(0.14)	(0.17)	(0.20)	(0.17)	(0.18)	(0.17)	(0.18)	(0.17)
PDd	0.29***	0.42***	0.33**	0.51***	0.67***	0.72***	0.44**	0.41**	0.52***	0.61***	0.54***
	(0.11)	(0.15)	(0.14)	(0.15)	(0.18)	(0.20)	(0.18)	(0.16)	(0.16)	(0.17)	(0.16)
PHc	0.15	0.25*	0.18	0.35**	0.50***	0.56***	0.28*	0.27*	0.38**	0.44***	0.38**
	(0.10)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
PHg	0.16	0.28*	0.20	0.36**	0.51***	0.57***	0.29*	0.28*	0.40***	0.46***	0.40**
	(0.10)	(0.14)	(0.14)	(0.15)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
PHh	0.16	0.29**	0.21	0.37**	0.52***	0.58***	0.29*	0.28*	0.40***	0.46***	0.40**
	(0.11)	(0.15)	(0.14)	(0.15)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
PHj	0.17*	0.30**	0.22	0.37**	0.52***	0.57***	0.29*	0.28*	0.41***	0.48***	0.42***
	(0.10)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.15)	(0.15)	(0.16)	(0.16)
PLd	0.22	0.29*	0.25	0.26*	0.35**	0.37**	0.11	0.11	0.39**	0.44**	0.40**
	(0.14)	(0.17)	(0.16)	(0.14)	(0.17)	(0.18)	(0.16)	(0.15)	(0.18)	(0.18)	(0.18)
PLe	0.41***	0.51***	0.38***	0.52***	0.69***	0.70***	0.44***	0.41***	0.51***	0.58***	0.52***
	(0.11)	(0.15)	(0.15)	(0.15)	(0.18)	(0.19)	(0.17)	(0.16)	(0.15)	(0.17)	(0.17)
RGc	0.28***	0.41***	0.35**	0.49***	0.82***	0.91***	0.62***	0.57***	0.55***	0.61***	0.59***
	(0.11)	(0.14)	(0.14)	(0.14)	(0.17)	(0.18)	(0.16)	(0.15)	(0.15)	(0.16)	(0.16)
RGd	0.22**	0.32**	0.26**	0.47***	0.57***	0.65***	0.32*	0.30*	0.50***	0.63***	0.58***
	(0.10)	(0.14)	(0.13)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.16)	(0.16)	(0.16)
RGe	0.27**	0.49***	0.34**	0.48***	0.61***	0.65***	0.39**	0.39**	0.55***	0.64***	0.59***
	(0.11)	(0.15)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
RK	0.29***	0.40***	0.33**	0.40***	0.57***	0.61***	0.47***	0.44***	0.54***	0.66***	0.60***
	(0.11)	(0.14)	(0.14)	(0.14)	(0.18)	(0.19)	(0.18)	(0.17)	(0.16)	(0.19)	(0.19)
SC	0.09	0.16	0.11	0.29**	0.40**	0.50***	0.23	0.21	0.32**	0.37**	0.32**
	(0.10)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.16)	(0.15)	(0.14)	(0.15)	(0.15)
SCg	0.19*	0.25*	0.17	0.33**	0.48***	0.51**	0.24	0.23	0.38**	0.44***	0.35**
	(0.11)	(0.15)	(0.14)	(0.15)	(0.18)	(0.20)	(0.17)	(0.16)	(0.15)	(0.16)	(0.16)
SCh	0.12	0.23	0.18	0.34**	0.45***	0.56***	0.29*	0.27	0.34**	0.41**	0.36**
	(0.10)	(0.15)	(0.15)	(0.14)	(0.17)	(0.19)	(0.17)	(0.16)	(0.16)	(0.17)	(0.17)
SCk	0.20*	0.28*	0.18	0.32**	0.51***	0.52**	0.23	0.22	0.41***	0.46***	0.37**
	(0.11)	(0.15)	(0.14)	(0.15)	(0.18)	(0.20)	(0.18)	(0.16)	(0.15)	(0.16)	(0.16)
SCm	0.10	0.19	0.18	0.33**	0.45***	0.53***	0.26	0.24	0.37**	0.46***	0.44***
	(0.10)	(0.14)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.15)	(0.15)	(0.16)	(0.17)
SCn	0.10	0.18	0.13	0.32**	0.47***	0.55***	0.28*	0.26*	0.35**	0.41**	0.34**
	(0.10)	(0.15)	(0.14)	(0.14)	(0.17)	(0.19)	(0.16)	(0.15)	(0.15)	(0.16)	(0.16)
SCy	0.07	0.35*	0.29	0.47**	0.44**	0.54**	0.26	0.15	0.31**	0.36**	0.31**
	(0.10)	(0.18)	(0.18)	(0.19)	(0.20)	(0.22)	(0.20)	(0.16)	(0.15)	(0.16)	(0.16)
SNg	0.32*	0.29*	0.23	0.40**	0.63***	0.67***	0.41*	0.40**	0.52***	0.57***	0.50**
	(0.17)	(0.17)	(0.16)	(0.17)	(0.21)	(0.22)	(0.21)	(0.20)	(0.19)	(0.20)	(0.20)
SNh	0.24**	0.33**	0.24	0.37**	0.56***	0.57***	0.30*	0.29*	0.45***	0.51***	0.41**
	(0.11)	(0.15)	(0.15)	(0.15)	(0.18)	(0.20)	(0.17)	(0.16)	(0.16)	(0.17)	(0.17)
ST	0.11	0.18	0.14	0.29**	0.36**	0.50**	0.25	0.24	0.37**	0.41***	0.38**
	(0.10)	(0.14)	(0.13)	(0.14)	(0.18)	(0.20)	(0.17)	(0.15)	(0.14)	(0.15)	(0.15)
UR	0.12	0.20	0.15	0.54**	0.92***	1.01***	0.76***	0.74***	0.58**	0.63***	0.83***
	(0.10)	(0.14)	(0.13)	(0.24)	(0.28)	(0.28)	(0.27)	(0.27)	(0.24)	(0.24)	(0.28)
VRd	0.27*	0.39**	0.36**	0.54***	0.61**	0.66**	0.41*	0.38	0.70***	0.74***	0.71***
	(0.15)	(0.18)	(0.17)	(0.17)	(0.25)	(0.26)	(0.25)	(0.24)	(0.20)	(0.21)	(0.21)
VRe	0.41***	0.74***	0.69***	0.65***	0.63***	0.59***	0.36**	0.43**	0.58***	0.62***	0.47***
	(0.13)	(0.16)	(0.16)	(0.16)	(0.18)	(0.20)	(0.18)	(0.18)	(0.17)	(0.18)	(0.17)
VRk	0.07	0.14	0.07	0.24*	0.26	0.39**	0.13	0.13	0.28*	0.32**	0.29*
	(0.10)	(0.14)	(0.13)	(0.14)	(0.17)	(0.19)	(0.17)	(0.15)	(0.15)	(0.16)	(0.16)
WR	0.34***	0.46***	0.38***	0.49***	0.54***	0.61***	0.38**	0.37**	0.50***	0.53***	0.48***
	(0.11)	(0.15)	(0.14)	(0.14)	(0.17)	(0.19)	(0.17)	(0.15)	(0.15)	(0.15)	(0.15)

Table C-1 (Continued)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Landform											
LP wet	−0.07	−0.16***	−0.14**	−0.07	−0.12**	−0.19***	−0.17***	−0.18***	−0.18***	−0.20***	−0.12
	(0.05)	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.08)
SH	−0.05***	−0.08***	−0.07***	−0.06***	−0.06***	−0.05***	−0.05***	−0.05***	−0.06***	−0.06***	−0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
SM	−0.04*	−0.07***	−0.04	−0.08**	−0.02	−0.03	−0.05	−0.02	−0.04	−0.08**	−0.09***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)
SP	−0.03	−0.05	−0.07**	−0.06	−0.04	−0.06	−0.08**	−0.09**	−0.06	−0.04	−0.03
	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
TH	−0.06***	−0.12***	−0.11***	−0.09***	−0.13***	−0.09***	−0.06**	−0.07**	−0.11***	−0.12***	−0.09***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
TM	−0.03**	−0.06***	−0.05***	−0.04***	−0.06***	−0.05***	−0.04**	−0.04**	−0.06***	−0.07***	−0.06***
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Lithology											
IA	−0.07**	−0.10**	−0.07*	−0.07*	−0.11***	−0.06	−0.04	−0.03	−0.03	−0.05	−0.04
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)
IA1	−0.03	−0.04	−0.02	−0.04	−0.03	0.01	0.02	0.02	−0.00	−0.02	−0.02
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
IA2	−0.15***	−0.21***	−0.18***	−0.19***	−0.23***	−0.28***	−0.24***	−0.27***	−0.25***	−0.07	−0.10
	(0.04)	(0.06)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.14)	(0.14)
IA4	−0.04*	−0.05*	−0.03	0.18	0.15	−0.09**	−0.04	−0.05	−0.11***	−0.19***	−0.21***
	(0.02)	(0.03)	(0.03)	(0.25)	(0.25)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
IB2	−0.03	−0.04	−0.02	−0.03	−0.03	0.02	0.03	0.03	0.06	0.04	0.00
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.06)	(0.06)	(0.05)
IB3	−0.15***	−0.24***	−0.21***	−0.22***	−0.29***	−0.27***	−0.23***	−0.22***	−0.23***	−0.25***	−0.27***
	(0.03)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
II1	−0.07**	−0.10***	−0.06*	−0.07**	−0.05	−0.01	−0.00	−0.01	−0.04	−0.06	−0.06
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
II2	0.16	0.06	0.15	0.15	−0.21***	−0.15***	−0.12***	−0.12***	−0.17***	0.14	−0.19***
	(0.19)	(0.19)	(0.19)	(0.19)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.23)	(0.04)
IP2	0.01	0.00	0.03	0.01	0.06	0.08**	0.09**	0.08*	0.05	0.05	0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
IP4	−0.03	−0.03	0.00	−0.05	−0.03	−0.02	0.03	0.03	0.00	−0.01	−0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
MA	−0.22***	−0.31***	−0.27***	−0.17***	−0.27***	−0.19**	−0.17**	−0.18**	−0.22***	−0.23***	−0.23***
	(0.07)	(0.08)	(0.07)	(0.05)	(0.08)	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)
MA1	−0.08***	−0.02	−0.01	−0.07	−0.13**	0.02	0.05	0.00	−0.11*	−0.15**	−0.16**
	(0.03)	(0.06)	(0.06)	(0.08)	(0.06)	(0.09)	(0.08)	(0.07)	(0.06)	(0.07)	(0.07)
MA1, SC2	0.27**	0.22**	0.08	−0.09*	−0.03	0.01	0.03	0.02	0.10	0.09	0.06
	(0.11)	(0.11)	(0.09)	(0.05)	(0.09)	(0.09)	(0.09)	(0.09)	(0.10)	(0.10)	(0.09)
MA2	−0.03	−0.05	−0.01	−0.03	−0.00	0.03	0.06	0.06	0.04	0.02	0.02
	(0.02)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.04)
MA3, MB1	−0.00	−0.00	0.02	−0.01	0.02	0.05	0.07*	0.08*	0.05	0.03	0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)
MA4, MB2	0.03	0.00	−0.00	−0.03	−0.01	0.01	0.03	0.03	−0.01	−0.02	−0.01
	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
MB1	−0.04*	−0.05	−0.03	−0.06*	0.01	0.05	0.03	0.01	−0.01	−0.02	−0.03
	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
MB1, MB2	−0.02	−0.05*	−0.03	−0.05	−0.09**	−0.06	−0.06	−0.06	−0.06	−0.06	−0.07
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)
SC	0.01	−0.01	0.02	−0.02	−0.02	0.01	0.04	0.04	0.02	0.01	0.00
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
SC1	−0.02	−0.01	0.00	0.00	−0.04	0.01	0.02	0.04	0.03	0.01	0.01
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
SC16	−0.00	−0.03	−0.03	−0.09**	−0.04	0.00	−0.05	0.03	0.08	0.06	0.02
	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.06)	(0.06)	(0.07)	(0.09)	(0.08)	(0.07)

Table C-1 (Continued)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
SC2	−0.03	−0.04	−0.02	−0.03	−0.04	0.01	0.02	0.03	0.01	−0.02	−0.03
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
SC2, SC4	0.06	0.08	0.14*	0.10	−0.02	0.01	0.11	0.11	0.05	0.00	0.03
	(0.06)	(0.07)	(0.07)	(0.07)	(0.05)	(0.05)	(0.07)	(0.07)	(0.06)	(0.06)	(0.07)
SC3	−0.05*	−0.07**	−0.04	−0.05	−0.04	0.10*	0.12**	0.09*	0.05	0.02	0.02
	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
SC4	−0.04	−0.04	−0.02	−0.04	−0.03	−0.03	−0.00	0.01	0.00	−0.03	−0.02
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
SC5	0.02	0.01	−0.00	0.04	−0.06	0.04	0.10	0.06	0.05	0.03	0.02
	(0.05)	(0.05)	(0.05)	(0.06)	(0.04)	(0.09)	(0.10)	(0.08)	(0.10)	(0.09)	(0.09)
SC7	0.07	0.01	0.04	0.06	0.07	0.04	−0.05	0.02	0.07	0.15*	0.16*
	(0.10)	(0.10)	(0.10)	(0.11)	(0.10)	(0.11)	(0.12)	(0.13)	(0.11)	(0.09)	(0.09)
SO1	−0.03	−0.03	−0.01	−0.04	−0.10***	−0.04	−0.02	−0.01	−0.01	−0.02	−0.01
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
SO2	−0.04	−0.05	−0.02	0.00	−0.11**	0.00	−0.02	−0.02	−0.04	−0.07	0.00
	(0.03)	(0.04)	(0.03)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)
SO3	−0.15***	−0.18***	−0.14***	−0.15***	−0.26***	−0.13**	−0.11*	−0.09*	−0.08	−0.13**	−0.12**
	(0.04)	(0.05)	(0.04)	(0.04)	(0.05)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)
UA1	0.81***	0.73***	0.79***	−0.24***	0.72***	0.76***	0.79***	0.78***	0.78***	0.74***	0.72***
	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)
UE1	0.01	−0.01	0.01	−0.01	0.01	0.06	0.07**	0.06	0.02	0.01	0.01
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
UE1, UR1	−0.13***	−0.17***	−0.14***	−0.13***	−0.23***	−0.16***	−0.15***	−0.14***	−0.14***	−0.16***	−0.17***
	(0.03)	(0.04)	(0.04)	(0.03)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
UE2	−0.08**	−0.10**	−0.07	−0.09**	−0.07*	−0.01	0.00	−0.01	−0.03	−0.04	−0.04
	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)
UF	0.03	0.03	0.05	0.02	0.03	0.07*	0.08**	0.08**	0.06	0.06	0.06
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
UF, UL	−0.02	−0.08	−0.09	−0.12**	−0.08	−0.07	−0.07	−0.06	−0.10	−0.10	−0.12*
	(0.06)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)
UF, UM	−0.13*	−0.15	−0.12	−0.13	−0.18**	−0.22***	−0.11	−0.05	−0.04	−0.08	−0.02
	(0.07)	(0.11)	(0.11)	(0.09)	(0.09)	(0.08)	(0.11)	(0.12)	(0.11)	(0.11)	(0.14)
UF1	0.09	0.10	0.04	0.05	0.07	0.10**	0.10**	0.09*	0.06	0.05	0.03
	(0.07)	(0.07)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)
UF2	−0.01	−0.04	−0.01	−0.03	−0.01	0.03	0.05	0.04	0.00	−0.00	0.04
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.07)
UG	0.01	0.03	0.05	0.03	0.01	0.07	0.09**	0.09**	0.06	0.06	0.08
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
UL	−0.05	−0.03	0.01	0.02	−0.02	0.08	0.09*	0.08	0.02	0.03	0.06
	(0.04)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)
UL, UM	0.86***	0.81***	0.82***	0.77***	0.81***	0.76***	0.77***	0.75***	0.75***	0.71***	0.68***
	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
UL2	−0.09***	−0.11***	−0.07**	−0.07**	−0.13***	−0.04	−0.03	−0.02	−0.01	−0.03	−0.04
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
UM	−0.07**	−0.08	−0.05	−0.11***	−0.09	−0.04	−0.04	−0.03	0.02	−0.00	0.03
	(0.03)	(0.06)	(0.06)	(0.04)	(0.06)	(0.07)	(0.07)	(0.07)	(0.08)	(0.07)	(0.08)
UO	0.03	0.09*	0.11**	0.10**	0.12**	0.15***	0.15***	0.13**	0.12**	0.12**	0.10*
	(0.04)	(0.05)	(0.04)	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)
Adj. R^2	0.17	0.20	0.17	0.12	0.19	0.18	0.18	0.17	0.14	0.15	0.14
F Stat.	13.35	16.63	13.56	9.12	15.82	14.98	14.24	13.35	11.20	11.60	11.27
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the county seats. The dependent variable is an indicator that equals one, if the pixel hosts a county seat in that year, and zero otherwise. Conley standard errors using a 150 km radius and a Bartlett kernel are in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. The table omits six indicator variables due to collinearity. Table B-2, Table B-3, and Table B-4 list the full names behind the soil variables' abbreviations.

Table C-2: Local Geography Prefecture Seat Regressions with Soil Coefficients

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	−0.08*	−0.10*	−0.04	−0.27***	−0.36***	−0.53***	−0.46***	−0.66***	−0.31**	−0.32***	−0.35***
	(0.04)	(0.05)	(0.06)	(0.10)	(0.09)	(0.14)	(0.15)	(0.14)	(0.12)	(0.12)	(0.11)
Dist. Coast	0.02	0.03	0.05	0.23***	0.40***	0.22***	0.22**	0.33***	0.16**	0.16**	0.10
	(0.03)	(0.04)	(0.05)	(0.07)	(0.07)	(0.08)	(0.10)	(0.09)	(0.07)	(0.07)	(0.07)
Dist. River	−0.03	−0.04	−0.07	−0.21	−0.36***	−0.50***	−0.30	−0.39**	−0.15	−0.18	−0.09
	(0.07)	(0.09)	(0.10)	(0.15)	(0.13)	(0.17)	(0.20)	(0.20)	(0.16)	(0.16)	(0.15)
Ruggedness	0.00	−0.05	−0.12*	−0.32***	−0.18	−0.49***	−0.46**	−0.39**	−0.51***	−0.47***	−0.41**
	(0.05)	(0.06)	(0.07)	(0.11)	(0.12)	(0.17)	(0.19)	(0.18)	(0.17)	(0.18)	(0.18)
Temperature	−0.03	0.04	0.16***	−0.03	−0.04	−0.31**	−0.17	−0.29**	−0.06	−0.02	−0.08
	(0.04)	(0.05)	(0.06)	(0.10)	(0.09)	(0.12)	(0.14)	(0.13)	(0.12)	(0.12)	(0.10)
Temperature^2	−0.01	−0.22	−0.35*	−0.10	0.00	1.29***	0.63	0.11	−0.06	−0.01	0.06
	(0.15)	(0.19)	(0.20)	(0.31)	(0.34)	(0.50)	(0.52)	(0.52)	(0.48)	(0.48)	(0.43)
Precipitation	−0.08	−0.22***	−0.22**	−0.27**	−0.15	−0.29	−0.25	−0.31	0.05	−0.15	−0.22
	(0.05)	(0.07)	(0.09)	(0.13)	(0.14)	(0.20)	(0.20)	(0.20)	(0.17)	(0.17)	(0.15)
Precipitation^2	0.13	0.41***	0.39**	0.49*	0.23	0.46	0.35	0.51	−0.24	0.16	0.31
	(0.11)	(0.15)	(0.19)	(0.27)	(0.28)	(0.40)	(0.41)	(0.41)	(0.36)	(0.35)	(0.31)
Elevation	−0.37*	−0.26	0.05	−0.97**	−1.99***	−1.83***	−1.69**	−2.67***	−0.11	−0.10	−0.86*
	(0.19)	(0.24)	(0.25)	(0.42)	(0.45)	(0.58)	(0.68)	(0.68)	(0.60)	(0.61)	(0.52)
Dominant Soil Type											
ACf	0.05**	0.08**	0.04	0.17***	0.17***	0.32***	0.26***	0.35***	0.16**	0.18***	0.20***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.09)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
ACh	0.06**	0.08**	0.05	0.17***	0.18***	0.28***	0.27***	0.37***	0.17**	0.19***	0.23***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
ACp	0.05*	0.06*	0.01	0.13**	0.26**	0.34***	0.50***	0.60***	0.21*	0.22**	0.35**
	(0.03)	(0.03)	(0.04)	(0.06)	(0.11)	(0.13)	(0.14)	(0.14)	(0.11)	(0.11)	(0.14)
ACu	0.06**	0.08**	0.05	0.16***	0.17***	0.28***	0.26***	0.36***	0.17**	0.19***	0.22***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.05)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
ALf	0.05*	0.10**	0.06	0.17**	0.22***	0.31***	0.28***	0.38***	0.20**	0.22**	0.24***
	(0.02)	(0.05)	(0.05)	(0.07)	(0.08)	(0.10)	(0.10)	(0.10)	(0.09)	(0.09)	(0.08)
ALh	0.06**	0.08**	0.04	0.16***	0.18***	0.30***	0.26***	0.36***	0.16**	0.19**	0.22***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
ALp	0.05*	0.06*	0.02	0.13**	0.23***	0.31**	0.27**	0.36***	0.20*	0.21*	0.23**
	(0.03)	(0.03)	(0.04)	(0.06)	(0.08)	(0.12)	(0.12)	(0.12)	(0.11)	(0.11)	(0.11)
ANh	0.06**	0.08**	0.04	0.17***	0.18***	0.28***	0.25***	0.35***	0.14*	0.16**	0.18***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)	(0.07)
ARb	0.05**	0.07**	0.03	0.16**	0.19***	0.28***	0.24***	0.34***	0.16**	0.16**	0.20***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.07)
ARc	0.06**	0.07**	0.04	0.16**	0.20***	0.28***	0.24***	0.33***	0.15**	0.16**	0.22***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.07)
ARh	0.06**	0.07**	0.03	0.15**	0.17***	0.26***	0.23***	0.31***	0.14*	0.14**	0.20***
	(0.03)	(0.03)	(0.04)	(0.07)	(0.06)	(0.08)	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)
ATa	0.05*	0.18	0.13	0.24*	0.26*	0.25***	0.21**	0.29***	0.12*	0.12*	0.15**
	(0.03)	(0.13)	(0.13)	(0.14)	(0.14)	(0.08)	(0.08)	(0.08)	(0.07)	(0.07)	(0.06)
ATc	0.06**	0.08**	0.04	0.18***	0.21***	0.31***	0.30***	0.39***	0.20***	0.21***	0.24***
	(0.02)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
ATf	0.05*	0.06*	0.02	0.26***	0.27**	0.37***	0.48***	0.56***	0.26**	0.27**	0.44***
	(0.03)	(0.03)	(0.04)	(0.09)	(0.12)	(0.13)	(0.12)	(0.12)	(0.13)	(0.13)	(0.11)
CHg	0.05**	0.07**	0.03	0.15**	0.17***	0.28***	0.24***	0.33***	0.15**	0.16**	0.18***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
CHh	0.06**	0.07**	0.04	0.16***	0.19***	0.30***	0.26***	0.35***	0.16**	0.17**	0.20***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
CHk	0.05**	0.07**	0.04	0.16***	0.19***	0.28***	0.24***	0.34***	0.16**	0.17**	0.19***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
CHl	0.06**	0.08**	0.04	0.15**	0.18***	0.34***	0.25***	0.34***	0.15**	0.16**	0.20***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.10)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)

Table C-2 (Continued)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
CLh	0.05**	0.06*	0.02	0.20***	0.15***	0.25***	0.21**	0.30***	0.16**	0.16**	0.16***
	(0.03)	(0.03)	(0.03)	(0.07)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
CLl	0.05**	0.06*	0.02	0.13**	0.16***	0.28***	0.23***	0.32***	0.14*	0.14*	0.18***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
CLp	0.05**	0.07**	0.03	0.15**	0.18***	0.28***	0.24***	0.34***	0.15**	0.16**	0.19***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
CMc	0.06**	0.09***	0.05	0.17***	0.22***	0.31***	0.30***	0.39***	0.16**	0.17**	0.22***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
CMd	0.06**	0.07**	0.04	0.16***	0.18***	0.31***	0.28***	0.38***	0.16**	0.18**	0.22***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
CMe	0.06**	0.07**	0.04	0.17***	0.21***	0.30***	0.27***	0.38***	0.16**	0.18**	0.21***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.05)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
CMg	0.07**	0.10***	0.05	0.16**	0.18***	0.28***	0.26***	0.38***	0.15**	0.18**	0.19***
	(0.03)	(0.04)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.09)	(0.07)	(0.08)	(0.07)
CMi	0.07**	0.08**	0.05	0.17***	0.21***	0.32***	0.28***	0.39***	0.15**	0.16**	0.21***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
CMo	0.05**	0.07**	0.03	0.14**	0.15***	0.27***	0.23***	0.36***	0.17**	0.18**	0.21***
	(0.02)	(0.03)	(0.03)	(0.06)	(0.05)	(0.09)	(0.09)	(0.08)	(0.07)	(0.07)	(0.07)
CMu	0.06**	0.07**	0.03	0.14**	0.14***	0.24***	0.21***	0.31***	0.12*	0.14**	0.17***
	(0.03)	(0.03)	(0.03)	(0.06)	(0.05)	(0.08)	(0.08)	(0.08)	(0.07)	(0.07)	(0.06)
CMx	0.05**	0.08***	0.05	0.19***	0.18***	0.34***	0.31***	0.39***	0.17**	0.18**	0.22***
	(0.03)	(0.03)	(0.03)	(0.06)	(0.06)	(0.09)	(0.09)	(0.09)	(0.07)	(0.07)	(0.06)
FLc	0.07***	0.10***	0.06	0.18***	0.24***	0.31***	0.29***	0.39***	0.17**	0.18**	0.21***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.07)
FLe	0.09**	0.10**	0.06	0.29***	0.20***	0.33***	0.31***	0.41***	0.23***	0.26***	0.27***
	(0.04)	(0.04)	(0.05)	(0.09)	(0.07)	(0.09)	(0.10)	(0.10)	(0.08)	(0.08)	(0.08)
FLs	0.07**	0.10**	0.05	0.20***	0.25***	0.32***	0.27***	0.40***	0.14*	0.16**	0.20***
	(0.03)	(0.04)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.09)	(0.07)	(0.07)	(0.07)
FRh	0.05*	0.06*	0.03	0.13**	0.17***	0.22***	0.18**	0.28***	0.14*	0.12*	0.15**
	(0.02)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)	(0.06)
FRx	0.06**	0.07**	0.04	0.15***	0.15***	0.23***	0.22***	0.32***	0.14*	0.15**	0.18***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
GG	0.05**	0.08**	0.04	0.17***	0.21***	0.32***	0.28***	0.40***	0.15**	0.17**	0.22***
	(0.03)	(0.03)	(0.03)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
GLe	0.04*	0.06*	0.01	0.12**	0.15***	0.30***	0.26**	0.35***	0.17*	0.18*	0.20**
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.11)	(0.11)	(0.11)	(0.10)	(0.10)	(0.09)
GLi	0.06**	0.08**	0.04	0.15***	0.16**	0.26***	0.22**	0.34***	0.15*	0.16**	0.21***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.08)	(0.07)	(0.06)
GLk	0.09*	0.10*	0.10	0.30***	0.24**	0.33***	0.33***	0.39***	0.13*	0.13*	0.24**
	(0.05)	(0.05)	(0.07)	(0.11)	(0.10)	(0.11)	(0.12)	(0.12)	(0.07)	(0.07)	(0.10)
GLm	0.05*	0.07**	0.02	0.15**	0.20***	0.29***	0.25***	0.34***	0.14**	0.15**	0.18***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.07)
GLt	0.05*	0.07**	0.03	0.14**	0.16***	0.19**	0.18*	1.28***	0.10	0.11	0.13*
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.09)	(0.09)	(0.09)	(0.08)	(0.07)	(0.07)
GRh	0.06**	0.08**	0.04	0.17***	0.20***	0.31***	0.29***	0.39***	0.17**	0.18***	0.25***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.09)	(0.07)	(0.07)	(0.07)
GYh	0.05*	0.06*	0.01	0.12**	0.14**	0.25***	0.21**	0.29***	0.32*	0.32*	0.16**
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.19)	(0.19)	(0.06)
GYk	0.05*	0.06	0.01	0.13**	0.16***	0.27***	0.23**	0.31***	0.15*	0.13*	0.17**
	(0.03)	(0.04)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.08)	(0.07)	(0.07)
GYl	0.04*	0.05	0.00	0.14**	0.15**	0.27***	0.21**	0.30***	0.14*	0.14*	0.15**
	(0.03)	(0.03)	(0.04)	(0.07)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
GYp	0.04*	0.05	−0.00	0.11*	0.13**	0.26***	0.21**	0.29***	0.12*	0.12*	0.15**
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.07)
HSs	0.05*	0.06*	0.01	0.13**	0.18***	0.27***	0.22***	0.32***	0.11	0.12	0.15**
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)

Table C-2 (Continued)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
KSh	0.05**	0.06**	0.03	0.18***	0.18***	0.30***	0.27***	0.37***	0.17**	0.17**	0.21***
	(0.03)	(0.03)	(0.04)	(0.07)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
KSk	0.05**	0.06*	0.02	0.14**	0.17***	0.28***	0.24***	0.33***	0.14**	0.15**	0.18***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
KSl	0.06**	0.07**	0.03	0.15***	0.19***	0.29***	0.25***	0.35***	0.16**	0.16**	0.21***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
LP	0.06**	0.07**	0.04	0.14**	0.16***	0.50**	0.48**	0.57**	0.12*	0.14**	0.19***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.24)	(0.24)	(0.24)	(0.07)	(0.07)	(0.06)
LPd	0.06**	0.07**	0.05	0.16**	0.18***	0.27***	0.24***	0.34***	0.27**	0.28**	0.19***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.12)	(0.12)	(0.06)
LPe	0.06**	0.07**	0.04	0.16***	0.18***	0.30***	0.26***	0.36***	0.15**	0.17**	0.21***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
LPi	0.07**	0.08**	0.05	0.17***	0.22***	0.31***	0.29***	0.39***	0.15**	0.16**	0.22***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
LPk	0.06**	0.09**	0.04	0.19***	0.20***	0.31***	0.26***	0.37***	0.19**	0.19***	0.22***
	(0.03)	(0.04)	(0.04)	(0.06)	(0.06)	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)	(0.07)
LPm	0.06**	0.08**	0.04	0.16***	0.20***	0.30***	0.27***	0.37***	0.14*	0.15**	0.20***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
LVa	0.06**	0.07**	0.04	0.16***	0.19***	0.29***	0.26***	0.35***	0.16**	0.17**	0.20***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
LVg	0.06*	0.11***	0.07*	0.21***	0.25***	0.36***	0.36***	0.44***	0.20***	0.20***	0.24***
	(0.03)	(0.04)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.09)	(0.07)	(0.07)	(0.07)
LVh	0.06**	0.08**	0.05	0.17***	0.20***	0.31***	0.29***	0.39***	0.18**	0.19***	0.23***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
LVj	0.06**	0.09***	0.05	0.18***	0.21***	0.32***	0.30***	0.40***	0.19***	0.20***	0.24***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
LVk	0.06**	0.09***	0.06	0.19***	0.22***	0.33***	0.31***	0.41***	0.18**	0.19***	0.24***
	(0.03)	(0.04)	(0.04)	(0.07)	(0.06)	(0.09)	(0.09)	(0.09)	(0.08)	(0.07)	(0.07)
LVx	0.06**	0.08**	0.06*	0.19***	0.20***	0.34***	0.31***	0.40***	0.22***	0.22***	0.24***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.09)	(0.09)	(0.08)	(0.08)	(0.08)	(0.07)
LXa	0.06**	0.07**	0.04	0.15***	0.17***	0.26***	0.23***	0.32***	0.13*	0.14**	0.19***
	(0.03)	(0.03)	(0.03)	(0.06)	(0.05)	(0.08)	(0.08)	(0.08)	(0.07)	(0.06)	(0.06)
LXf	0.06**	0.08**	0.05	0.21***	0.17***	0.26***	0.23***	0.33***	0.24**	0.26***	0.25***
	(0.03)	(0.03)	(0.04)	(0.07)	(0.06)	(0.08)	(0.08)	(0.08)	(0.10)	(0.10)	(0.08)
NTu	0.05*	0.07**	0.04	0.14**	0.15***	0.28***	0.26***	0.29***	0.18	0.19	0.14**
	(0.03)	(0.03)	(0.04)	(0.06)	(0.05)	(0.09)	(0.10)	(0.08)	(0.12)	(0.12)	(0.07)
PDd	0.06**	0.08**	0.05	0.18***	0.20***	0.32***	0.29***	0.39***	0.17**	0.18**	0.22***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.09)	(0.09)	(0.09)	(0.07)	(0.07)	(0.06)
PHc	0.05*	0.06*	0.02	0.14**	0.17***	0.28***	0.24***	0.33***	0.15**	0.16**	0.19***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.07)
PHg	0.05*	0.06**	0.02	0.15**	0.18***	0.29***	0.24***	0.33***	0.15**	0.16**	0.18***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
PHh	0.05*	0.07**	0.03	0.15**	0.18***	0.29***	0.24***	0.34***	0.16**	0.17**	0.18***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
PHj	0.05**	0.07**	0.04	0.15**	0.18***	0.28***	0.24***	0.33***	0.16**	0.17**	0.19***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
PLd	0.05*	0.06*	0.02	0.12**	0.14**	0.22***	0.19**	0.28***	0.26**	0.27**	0.29***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.12)	(0.12)	(0.11)
PLe	0.06**	0.08**	0.04	0.17***	0.20***	0.27***	0.25***	0.34***	0.12*	0.14**	0.17***
	(0.03)	(0.04)	(0.04)	(0.07)	(0.06)	(0.08)	(0.09)	(0.09)	(0.07)	(0.07)	(0.07)
RGc	0.06**	0.08**	0.06*	0.17***	0.22***	0.34***	0.32***	0.42***	0.17**	0.19***	0.24***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.09)	(0.09)	(0.09)	(0.07)	(0.07)	(0.07)
RGd	0.05*	0.08**	0.04	0.16***	0.17***	0.30***	0.27***	0.36***	0.16**	0.19**	0.20***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.08)	(0.07)	(0.06)
RGe	0.06**	0.07**	0.07*	0.18***	0.21***	0.31***	0.28***	0.38***	0.18**	0.19***	0.22***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)

Table C-2 (Continued)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
RK	0.10*	0.07**	0.03	0.14***	0.17***	0.27***	0.23***	0.34***	0.19**	0.21**	0.24***
	(0.06)	(0.03)	(0.03)	(0.06)	(0.05)	(0.08)	(0.08)	(0.08)	(0.10)	(0.10)	(0.08)
SC	0.04*	0.05	0.00	0.12**	0.15***	0.25***	0.21**	0.30***	0.12*	0.12*	0.15**
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.08)	(0.08)	(0.07)	(0.07)	(0.06)
SCg	0.05*	0.05*	−0.00	0.12*	0.17***	0.26***	0.22**	0.31***	0.10	0.10	0.14**
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.09)	(0.07)	(0.07)	(0.07)
SCh	0.05*	0.06*	0.02	0.15**	0.18***	0.26***	0.27***	0.37***	0.18**	0.19**	0.20***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.10)	(0.09)	(0.08)	(0.08)	(0.08)
SCk	0.05*	0.06*	−0.01	0.12*	0.17***	0.28***	0.22**	0.31***	0.10	0.10	0.15**
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.09)	(0.09)	(0.09)	(0.08)	(0.07)	(0.07)
SCm	0.04	0.04	0.00	0.13**	0.15***	0.25***	0.21**	0.30***	0.13*	0.13*	0.15**
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.07)
SCn	0.04*	0.05*	0.01	0.14**	0.18***	0.28***	0.23***	0.33***	0.15**	0.15**	0.16**
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
SCy	0.04	0.04	−0.01	0.21*	0.12**	0.35***	0.30**	0.39***	0.12*	0.12*	0.15**
	(0.03)	(0.03)	(0.04)	(0.11)	(0.06)	(0.12)	(0.13)	(0.12)	(0.07)	(0.07)	(0.07)
SNg	0.04*	0.05*	0.01	0.24**	0.27**	0.36***	0.31**	0.41***	0.13*	0.13*	0.15**
	(0.03)	(0.03)	(0.04)	(0.12)	(0.11)	(0.13)	(0.13)	(0.13)	(0.07)	(0.07)	(0.06)
SNh	0.05*	0.07**	0.01	0.14**	0.20***	0.30***	0.25***	0.34***	0.13*	0.14*	0.17**
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.09)	(0.09)	(0.09)	(0.08)	(0.07)	(0.07)
ST	0.04*	0.04	−0.00	0.10*	0.12**	0.25***	0.20**	0.30***	0.12	0.12*	0.16**
	(0.02)	(0.03)	(0.03)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
UR	0.04*	0.06*	0.01	0.38*	0.66***	0.51**	0.72***	0.57**	0.38*	0.40*	0.67***
	(0.02)	(0.03)	(0.03)	(0.22)	(0.25)	(0.23)	(0.25)	(0.23)	(0.23)	(0.23)	(0.26)
VRd	0.04	0.04	0.01	0.16***	0.16***	0.21**	0.18*	0.28***	0.12	0.13*	0.13*
	(0.03)	(0.04)	(0.04)	(0.06)	(0.06)	(0.09)	(0.09)	(0.09)	(0.08)	(0.08)	(0.07)
VRe	0.04	0.08**	0.01	0.17***	0.26***	0.32***	0.28***	0.38***	0.11	0.12*	0.19**
	(0.03)	(0.04)	(0.04)	(0.07)	(0.07)	(0.09)	(0.10)	(0.10)	(0.07)	(0.07)	(0.08)
VRk	0.04	0.04	−0.01	0.09	0.10*	0.23***	0.18**	0.27***	0.12	0.11	0.14**
	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
WR	0.04*	0.05*	0.03	0.16***	0.16***	0.26***	0.25***	0.36***	0.13*	0.16**	0.17***
	(0.02)	(0.03)	(0.03)	(0.06)	(0.05)	(0.08)	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)
Landform											
LP wet	−0.01**	−0.02***	−0.02***	−0.04***	−0.06***	−0.06***	−0.09***	−0.09***	−0.05***	−0.06***	−0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
SH	−0.01**	−0.01***	−0.01***	−0.01**	−0.01*	−0.00	−0.01	−0.01*	−0.01*	−0.01*	−0.01**
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
SM	0.00	−0.01***	−0.02***	−0.01	−0.00	0.00	−0.03	−0.04**	−0.03**	−0.03***	−0.02
	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)
SP	−0.01***	−0.02***	−0.02***	−0.02	0.00	−0.00	−0.02	−0.01	−0.02	−0.02	−0.03*
	(0.00)	(0.00)	(0.00)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
TH	−0.00	−0.01***	−0.02***	−0.02***	−0.03***	−0.03**	−0.03***	−0.03***	−0.01	−0.02**	−0.03***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
TM	−0.01**	−0.01***	−0.01***	−0.01	−0.01	−0.01	−0.01	−0.02**	−0.02***	−0.02**	−0.02**
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Lithology											
IA	−0.02*	−0.02*	−0.03**	−0.03**	−0.04**	−0.03	−0.04*	−0.04	−0.02	−0.03	−0.04*
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
IA1	−0.01	−0.01	−0.01	−0.02	−0.01	−0.02	−0.03	−0.02	−0.02	−0.02	−0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
IA2	−0.02**	−0.02	−0.02	−0.05***	−0.07***	−0.09***	−0.10***	−0.08***	−0.07***	−0.07***	−0.09***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
IA4	−0.01	−0.01	−0.01	−0.02	−0.03**	−0.06***	−0.06**	−0.05**	−0.04**	−0.05**	−0.05**
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
IB2	−0.01	−0.02*	−0.02**	−0.03**	−0.03**	−0.03	−0.04	−0.03	−0.01	−0.01	−0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)

Table C-2 (Continued)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
IB3	−0.02*	−0.04**	−0.04**	−0.06**	−0.07***	−0.08***	−0.10***	−0.09***	−0.05**	−0.05**	−0.08***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)
II1	−0.01	−0.01	−0.02	−0.02	−0.01	−0.01	−0.03	−0.02	−0.00	0.00	−0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
II2	−0.02**	−0.03***	−0.04***	−0.04**	−0.03**	−0.05***	−0.07***	−0.06**	−0.04**	−0.04**	−0.06***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
IP2	−0.00	−0.00	−0.00	−0.00	0.01	0.01	0.02	0.03	−0.00	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
IP4	−0.01	−0.01	−0.01	−0.01	−0.01	−0.03	−0.04*	−0.03	−0.02	−0.02	−0.02
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
MA	−0.01	−0.05**	−0.05**	−0.06**	−0.06**	−0.06**	−0.11**	−0.09**	−0.05**	−0.04	−0.06*
	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)	(0.02)	(0.02)	(0.03)
MA1	0.02	−0.01	−0.02	−0.03**	−0.05***	−0.05**	−0.08***	−0.07***	−0.02	−0.06***	−0.05***
	(0.03)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.02)	(0.02)	(0.04)	(0.02)	(0.02)
MA1, SC2	−0.01	−0.03	−0.03*	−0.03	−0.03	0.06	0.03	0.05	0.07	0.07	0.05
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)
MA2	−0.01	−0.01	−0.02	−0.01	0.00	−0.01	−0.02	−0.01	−0.01	−0.01	−0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
MA3, MB1	−0.01	−0.01	−0.01	−0.01	0.00	−0.01	−0.01	−0.00	−0.00	0.00	−0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
MA4, MB2	−0.01	−0.01	−0.01	−0.02	−0.01	−0.03	−0.03	−0.02	−0.03*	−0.03*	−0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
MB1	−0.01	−0.01	−0.01	−0.02	−0.01	−0.00	−0.02	−0.01	−0.02	−0.02	−0.02
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
MB1, MB2	−0.01	−0.01	−0.02*	−0.03	−0.03**	−0.05**	−0.05**	−0.05**	−0.02	−0.02	−0.03
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
SC	−0.01	−0.01	−0.01	−0.00	−0.01	−0.01	−0.01	−0.00	−0.00	0.00	−0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
SC1	−0.01	0.00	0.00	−0.02	−0.02	−0.01	−0.02	−0.01	−0.02	−0.01	−0.02
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
SC16	−0.02**	−0.01	−0.02**	−0.03**	−0.03***	0.02	−0.05*	−0.03	−0.02	0.00	0.05
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
SC2	−0.01	−0.01	−0.01	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
SC2, SC4	−0.01	−0.01	−0.01	0.02	−0.02	−0.04*	0.00	0.01	−0.00	−0.00	−0.03*
	(0.01)	(0.01)	(0.01)	(0.04)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)	(0.04)	(0.02)
SC3	−0.01	−0.01	−0.01	−0.01	0.03	0.03	0.03	0.02	−0.02	−0.01	−0.02
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
SC4	−0.01	−0.01	−0.01	−0.01	−0.01	−0.02	−0.02	−0.02	−0.01	−0.01	−0.02
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
SC5	−0.02**	−0.01	−0.02*	0.00	0.01	−0.05***	−0.06***	−0.03	−0.02	−0.03	0.01
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)
SC7	−0.02*	−0.02**	0.03	0.07	0.06	0.05	0.04	0.04	0.01	0.01	0.00
	(0.01)	(0.01)	(0.05)	(0.10)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
SO1	−0.01*	−0.01	−0.02*	−0.02	−0.03**	−0.02	−0.03	−0.02	−0.01	−0.01	−0.03
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
SO2	−0.01	−0.00	−0.02	−0.03	−0.02	−0.01	−0.00	0.00	−0.03	−0.00	−0.04*
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.02)
SO3	−0.02*	−0.02**	−0.04***	−0.05**	−0.05***	−0.04	−0.04	−0.05*	−0.02	−0.04	−0.04*
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
UA1	−0.02**	−0.03***	−0.04***	−0.07***	−0.08***	−0.08***	−0.10***	−0.10***	−0.08***	−0.08***	−0.09***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
UE1	−0.01	−0.01	−0.01	−0.00	0.00	0.01	0.01	0.02	−0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
UE1, UR1	−0.02*	−0.03***	−0.03***	−0.04**	−0.06***	−0.05**	0.27	0.28	−0.03	−0.02	−0.04**
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.17)	(0.17)	(0.02)	(0.02)	(0.02)

Table C-2 (Continued)

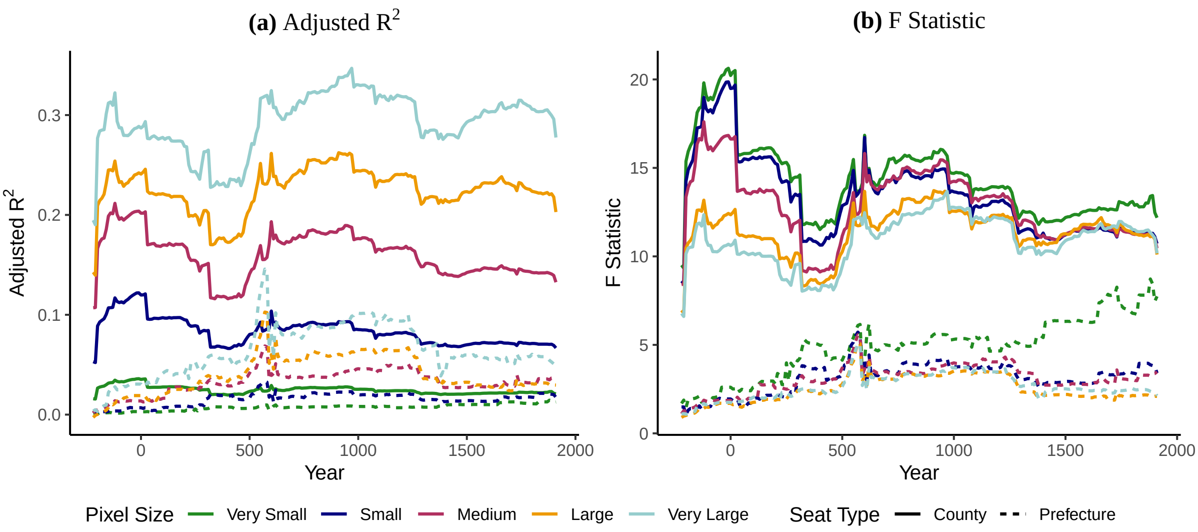
	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
UE2	−0.01	−0.02	−0.02*	−0.03	−0.02	0.01	−0.00	0.01	−0.01	−0.00	−0.03
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)
UF	−0.00	0.00	0.00	−0.00	0.00	0.00	0.01	0.02	−0.00	0.01	0.02
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
UF, UL	0.00	0.01	−0.00	−0.04*	−0.03	0.03	0.02	0.02	−0.00	0.00	0.02
	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)
UF, UM	0.04	0.04	0.04	0.01	−0.02	0.03	0.02	0.03	0.05	0.05	0.05
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
UF1	−0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.03	−0.01	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
UF2	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01	0.00	−0.02	−0.01	0.04
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)
UG	−0.00	0.00	0.01	−0.00	−0.01	−0.00	0.01	0.02	0.01	0.02	0.02
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
UL	−0.01	−0.01	0.01	−0.00	−0.02	−0.00	0.01	0.01	0.02	0.03	0.02
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
UL, UM	−0.02**	−0.02*	0.98***	0.95***	0.94***	0.92***	0.90***	0.92***	0.92***	0.92***	0.91***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
UL2	−0.01	−0.02*	−0.03***	−0.03**	−0.04***	−0.03	−0.04*	−0.03	−0.02	−0.03	−0.03*
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
UM	−0.01	−0.01	−0.01	−0.03*	−0.03*	−0.01	−0.02	−0.01	−0.01	0.01	−0.00
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
UO	0.00	0.02*	0.03**	0.04*	0.03*	0.04*	0.04*	0.06**	0.06**	0.07***	0.05**
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Adj. R^2	0.01	0.01	0.03	0.03	0.04	0.04	0.05	0.05	0.03	0.03	0.03
F Stat.	1.28	1.86	2.56	2.83	3.65	3.48	3.98	4.04	2.72	2.75	3.04
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the prefecture seats. The dependent variable is an indicator that equals one, if the pixel hosts a prefecture seat in that year, and zero otherwise. Conley standard errors using a 150 km radius and a Bartlett kernel are in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. The table omits six indicator variables due to collinearity. Table B-2, Table B-3, and Table B-4 list the full names behind the soil variables' abbreviations.

C.2 Supplementary OLS Results on Varying Pixel Size

In Section 4.1, we discuss the effect of a changing pixel size on the random forest results in line with the MAUP. The following Figure C-1 illustrates that the goodness of fit also differs systematically between county and prefecture seats when using OLS, apart from the classification forests.

Figure C-1: Explanatory Power in OLS Regressions



Notes: The figures plot the adjusted R^2 and F statistics of cross-sectional OLS regressions on local geography. Whereas Table 2 and Table 3 report these values for eleven selected cross-sections, this figure prints a result for one cross-section every ten years. The input resolution and the specification, i.e. eq. (14), are the same as in these baseline tables. Figure 5 and Figure C-4 are closely related figures using random forests rather than econometric techniques.

Shedding more light on coefficient estimates' magnitude and statistical significance at these alternative resolutions, Table C-5 to Table C-12 repeat Table 2 and Table 3 with those differently sized grid cells. They also add heteroskedasticity-robust standard errors as further evidence on top of the baseline Conley standard errors. The results confirm our baseline findings. While county seats show a strong connection to local geography, the link is much weaker, at best, for prefecture seats. The absolute magnitude of the effect depends on the pixel size, providing evidence for the MAUP. In many cases, heteroskedasticity-robust standard errors support Conley results, attributing even higher significance levels to the estimated coefficients.

Table C-3: Local Geography County Seat Regressions (Very Small Pixels)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	−0.07	−0.10	−0.10	−0.14	−0.20	−0.26	−0.19	−0.17	−0.18	−0.23	−0.21
	(0.02)***	(0.02)***	(0.02)***	(0.02)***	(0.02)***	(0.03)***	(0.02)***	(0.02)***	(0.02)***	(0.03)***	(0.03)***
	[0.02]***	[0.03]***	[0.03]***	[0.03]***	[0.04]***	[0.05]***	[0.04]***	[0.03]***	[0.03]***	[0.04]***	[0.03]***
Dist. Coast	−0.00	−0.05	0.01	0.07	0.13	0.13	0.12	0.11	0.04	0.07	0.05
	(0.01)	(0.02)***	(0.01)	(0.01)***	(0.02)***	(0.02)***	(0.02)***	(0.02)***	(0.02)**	(0.02)***	(0.02)***
	[0.02]	[0.03]	[0.02]	[0.02]***	[0.03]***	[0.04]***	[0.04]***	[0.04]***	[0.03]	[0.03]**	[0.03]*
Dist. River	0.01	−0.05	−0.05	−0.14	−0.26	−0.14	−0.06	−0.03	−0.05	−0.06	−0.03
	(0.03)	(0.03)	(0.03)	(0.03)***	(0.03)***	(0.04)***	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
	[0.05]	[0.07]	[0.06]	[0.06]**	[0.07]***	[0.08]*	[0.08]	[0.07]	[0.07]	[0.08]	[0.07]
Ruggedness	−0.07	−0.08	−0.08	−0.12	−0.09	−0.16	−0.15	−0.14	−0.12	−0.14	−0.17
	(0.01)***	(0.02)***	(0.02)***	(0.02)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***
	[0.02]***	[0.03]**	[0.03]***	[0.03]***	[0.04]**	[0.05]***	[0.04]***	[0.04]***	[0.04]***	[0.04]***	[0.04]***
Temperature	0.05	0.09	0.05	−0.01	−0.04	−0.15	−0.08	−0.06	−0.06	−0.08	−0.08
	(0.01)***	(0.02)***	(0.02)***	(0.02)	(0.02)*	(0.03)***	(0.02)***	(0.02)***	(0.02)***	(0.02)***	(0.02)***
	[0.03]**	[0.04]**	[0.03]*	[0.03]	[0.04]	[0.05]***	[0.04]**	[0.03]*	[0.03]**	[0.03]**	[0.03]**
Temperature^2	−0.09	−0.02	−0.09	−0.16	0.32	0.53	0.28	0.20	0.10	−0.06	−0.02
	(0.05)**	(0.06)	(0.06)	(0.07)**	(0.09)***	(0.11)***	(0.10)***	(0.10)**	(0.09)	(0.11)	(0.11)
	[0.10]	[0.15]	[0.13]	[0.12]	[0.16]**	[0.19]***	[0.16]*	[0.16]	[0.15]	[0.17]	[0.17]
Precipitation	−0.19	−0.35	−0.19	−0.07	−0.25	−0.10	0.01	0.03	−0.05	−0.02	−0.03
	(0.03)***	(0.03)***	(0.03)***	(0.03)**	(0.04)***	(0.04)**	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
	[0.05]***	[0.07]***	[0.05]***	[0.05]	[0.08]***	[0.09]	[0.08]	[0.07]	[0.06]	[0.07]	[0.06]
Precipitation^2	0.34	0.64	0.33	0.13	0.44	0.14	−0.06	−0.11	0.04	−0.02	−0.00
	(0.06)***	(0.07)***	(0.07)***	(0.07)*	(0.08)***	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)
	[0.10]***	[0.14]***	[0.10]***	[0.10]	[0.16]***	[0.17]	[0.16]	[0.15]	[0.11]	[0.13]	[0.12]
Elevation	−0.38	−0.32	−0.38	−0.71	−1.19	−1.32	−1.10	−0.94	−0.82	−1.09	−0.93
	(0.07)***	(0.10)***	(0.09)***	(0.09)***	(0.10)***	(0.12)***	(0.11)***	(0.11)***	(0.11)***	(0.13)***	(0.12)***
	[0.11]***	[0.14]**	[0.13]***	[0.13]***	[0.18]***	[0.19]***	[0.17]***	[0.15]***	[0.15]***	[0.18]***	[0.15]***
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.03	0.04	0.03	0.02	0.03	0.03	0.03	0.02	0.02	0.02	0.02
F Stat.	15.37	20.22	15.80	11.52	16.84	15.70	14.69	13.89	12.13	12.60	13.10
Num. obs.	86,257	86,257	86,257	86,257	86,257	86,257	86,257	86,257	86,257	86,257	86,257

Notes: The table reports the regression results of eq. (14) using the county seats. The dependent variable is an indicator that equals one, if the pixel hosts a county seat in that year, and zero otherwise. Heteroskedasticity-robust standard errors are in parentheses and Conley standard errors with a radius of 150 km (great-circle distances computed via the haversine formula (Sinnott, 1984)) and a Bartlett kernel are in brackets ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-4: Local Geography Prefecture Seat Regressions (Very Small Pixels)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	−0.01	−0.01	−0.01	−0.04	−0.06	−0.08	−0.08	−0.10	−0.05	−0.05	−0.05
	(0.00)**	(0.01)**	(0.01)	(0.01)***	(0.01)***	(0.01)***	(0.02)***	(0.02)***	(0.01)***	(0.01)***	(0.01)***
	[0.00]***	[0.01]***	[0.01]	[0.01]***	[0.01]***	[0.02]***	[0.02]***	[0.02]***	[0.01]***	[0.01]***	[0.01]***
Dist. Coast	−0.00	−0.00	0.01	0.03	0.04	0.02	0.03	0.04	0.02	0.02	0.01
	(0.00)	(0.00)	(0.01)	(0.01)***	(0.01)***	(0.01)**	(0.01)***	(0.01)***	(0.01)**	(0.01)**	(0.01)
	[0.00]	[0.00]	[0.01]	[0.01]***	[0.01]***	[0.01]*	[0.01]***	[0.01]***	[0.01]***	[0.01]**	[0.01]
Dist. River	0.00	0.01	0.00	−0.03	−0.04	−0.06	−0.04	−0.05	−0.03	−0.04	−0.04
	(0.01)	(0.01)	(0.01)	(0.02)*	(0.02)**	(0.02)***	(0.02)*	(0.02)**	(0.02)*	(0.02)**	(0.02)*
	[0.01]	[0.01]	[0.01]	[0.02]*	[0.02]**	[0.02]***	[0.03]	[0.02]**	[0.02]*	[0.02]**	[0.02]*
Ruggedness	−0.01	−0.02	−0.02	−0.03	−0.02	−0.06	−0.04	−0.04	−0.07	−0.06	−0.04
	(0.00)**	(0.01)***	(0.01)***	(0.01)***	(0.01)	(0.02)***	(0.02)**	(0.02)**	(0.01)***	(0.02)***	(0.02)***
	[0.00]**	[0.01]***	[0.01]***	[0.01]***	[0.01]	[0.02]***	[0.02]**	[0.02]**	[0.02]***	[0.02]***	[0.02]**
Temperature	−0.00	0.01	0.01	−0.00	−0.01	−0.06	−0.04	−0.06	−0.02	−0.01	−0.02
	(0.00)	(0.01)	(0.01)**	(0.01)	(0.01)	(0.01)***	(0.01)***	(0.01)***	(0.01)	(0.01)	(0.01)*
	[0.00]	[0.01]	[0.01]**	[0.01]	[0.01]	[0.02]***	[0.02]***	[0.01]***	[0.01]	[0.01]	[0.01]*
Temperature^2	0.00	−0.03	−0.03	−0.04	−0.07	0.18	0.02	−0.01	−0.07	−0.07	−0.02
	(0.02)	(0.03)	(0.03)	(0.04)	(0.04)*	(0.06)***	(0.06)	(0.06)	(0.05)	(0.06)	(0.05)
	[0.02]	[0.02]	[0.02]	[0.04]	[0.04]*	[0.07]**	[0.06]	[0.06]	[0.06]	[0.06]	[0.05]
Precipitation	−0.02	−0.04	−0.03	−0.07	−0.04	−0.05	−0.04	−0.04	−0.02	−0.04	−0.04
	(0.01)**	(0.01)***	(0.01)**	(0.02)***	(0.02)*	(0.02)**	(0.02)*	(0.02)*	(0.02)	(0.02)**	(0.02)*
	[0.01]**	[0.01]***	[0.01]**	[0.02]***	[0.02]**	[0.02]**	[0.02]*	[0.02]*	[0.02]	[0.02]**	[0.02]**
Precipitation^2	0.03	0.09	0.05	0.14	0.07	0.09	0.06	0.07	0.02	0.07	0.06
	(0.01)**	(0.02)***	(0.03)**	(0.04)***	(0.04)*	(0.04)**	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)
	[0.01]**	[0.02]***	[0.02]**	[0.04]***	[0.03]**	[0.04]**	[0.05]	[0.05]	[0.04]	[0.04]	[0.04]*
Elevation	−0.04	−0.03	−0.00	−0.12	−0.30	−0.30	−0.34	−0.45	−0.06	−0.07	−0.15
	(0.02)**	(0.03)	(0.03)	(0.05)**	(0.05)***	(0.06)***	(0.07)***	(0.07)***	(0.06)	(0.06)	(0.06)***
	[0.02]**	[0.03]	[0.03]	[0.05]**	[0.05]***	[0.07]***	[0.08]***	[0.07]***	[0.06]	[0.07]	[0.06]***
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
F Stat.	1.45	2.61	2.80	4.98	4.14	5.15	5.31	4.66	5.41	6.31	6.79
Num. obs.	86,257	86,257	86,257	86,257	86,257	86,257	86,257	86,257	86,257	86,257	86,257

Notes: The table reports the regression results of eq. (14) using the prefecture seats. The dependent variable is an indicator that equals one, if the pixel hosts a prefecture seat in that year, and zero otherwise. Heteroskedasticity-robust standard errors are in parentheses and Conley standard errors with a radius of 150 km (great-circle distances computed via the haversine formula (Sinnott, 1984)) and a Bartlett kernel are in brackets ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-5: Local Geography County Seat Regressions (Small Pixels)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	−0.22	−0.30	−0.28	−0.46	−0.61	−0.79	−0.49	−0.48	−0.53	−0.61	−0.56
	(0.06)***	(0.08)***	(0.08)***	(0.08)***	(0.09)***	(0.10)***	(0.10)***	(0.10)***	(0.09)***	(0.10)***	(0.11)***
	[0.10]**	[0.12]**	[0.12]**	[0.12]***	[0.15]***	[0.17]***	[0.14]***	[0.13]***	[0.12]***	[0.13]***	[0.12]***
Dist. Coast	0.01	−0.12	0.07	0.25	0.40	0.42	0.42	0.38	0.14	0.23	0.16
	(0.05)	(0.06)**	(0.06)	(0.06)***	(0.06)***	(0.06)***	(0.06)***	(0.06)***	(0.06)**	(0.06)***	(0.06)**
	[0.08]	[0.11]	[0.09]	[0.09]***	[0.12]***	[0.13]***	[0.12]***	[0.12]***	[0.09]	[0.10]**	[0.09]*
Dist. River	−0.03	−0.21	−0.24	−0.51	−0.96	−0.52	−0.27	−0.17	−0.20	−0.22	−0.10
	(0.10)	(0.12)*	(0.12)**	(0.13)***	(0.12)***	(0.14)***	(0.14)*	(0.14)	(0.14)	(0.15)	(0.16)
	[0.20]	[0.23]	[0.21]	[0.22]**	[0.26]***	[0.27]*	[0.26]	[0.25]	[0.25]	[0.26]	[0.24]
Ruggedness	−0.27	−0.34	−0.33	−0.57	−0.29	−0.73	−0.61	−0.59	−0.54	−0.58	−0.63
	(0.07)***	(0.09)***	(0.09)***	(0.09)***	(0.11)***	(0.12)***	(0.12)***	(0.12)***	(0.12)***	(0.13)***	(0.13)***
	[0.10]***	[0.13]**	[0.12]***	[0.11]***	[0.17]*	[0.19]***	[0.17]***	[0.17]***	[0.16]***	[0.16]***	[0.16]***
Temperature	0.25	0.36	0.28	0.07	−0.09	−0.36	−0.07	−0.05	−0.08	−0.07	−0.03
	(0.05)***	(0.07)***	(0.07)***	(0.07)	(0.08)	(0.09)***	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
	[0.10]***	[0.14]***	[0.12]**	[0.11]	[0.15]	[0.17]**	[0.14]	[0.13]	[0.12]	[0.13]	[0.13]
Temperature^2	−0.38	0.04	−0.23	−0.76	1.36	1.52	0.76	0.47	0.15	−0.36	−0.36
	(0.18)**	(0.25)	(0.25)	(0.29)***	(0.32)***	(0.40)***	(0.37)**	(0.37)	(0.36)	(0.40)	(0.41)
	[0.39]	[0.55]	[0.49]	[0.49]	[0.61]**	[0.71]**	[0.62]	[0.63]	[0.61]	[0.64]	[0.61]
Precipitation	−0.68	−1.10	−0.58	−0.21	−0.79	−0.28	0.17	0.21	−0.03	0.12	0.15
	(0.11)***	(0.13)***	(0.12)***	(0.13)	(0.14)***	(0.16)*	(0.15)	(0.15)	(0.15)	(0.16)	(0.16)
	[0.20]***	[0.23]***	[0.19]***	[0.20]	[0.28]***	[0.30]	[0.26]	[0.25]	[0.21]	[0.24]	[0.22]
Precipitation^2	1.14	1.86	0.89	0.31	1.25	0.20	−0.57	−0.67	−0.18	−0.49	−0.52
	(0.22)***	(0.26)***	(0.25)***	(0.26)	(0.29)***	(0.31)	(0.30)*	(0.30)**	(0.29)	(0.32)	(0.32)*
	[0.38]***	[0.45]***	[0.38]**	[0.41]	[0.56]**	[0.60]	[0.56]	[0.54]	[0.43]	[0.48]	[0.44]
Elevation	−1.26	−1.05	−1.19	−2.36	−3.82	−4.21	−3.27	−3.12	−2.66	−3.16	−2.82
	(0.29)***	(0.38)***	(0.37)***	(0.37)***	(0.39)***	(0.45)***	(0.44)***	(0.45)***	(0.44)***	(0.49)***	(0.48)***
	[0.44]***	[0.54]*	[0.52]**	[0.52]***	[0.66]***	[0.69]***	[0.61]***	[0.57]***	[0.56]***	[0.61]***	[0.54]***
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.09	0.12	0.09	0.07	0.10	0.09	0.09	0.08	0.07	0.07	0.07
F Stat.	14.24	19.49	15.26	10.64	16.72	14.77	13.65	13.07	11.32	11.49	11.42
Num. obs.	21,597	21,597	21,597	21,597	21,597	21,597	21,597	21,597	21,597	21,597	21,597

Notes: The table reports the regression results of eq. (14) using the county seats. The dependent variable is an indicator that equals one, if the pixel hosts a county seat in that year, and zero otherwise. Heteroskedasticity-robust standard errors are in parentheses and Conley standard errors with a radius of 150 km and a Bartlett kernel are in brackets ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-6: Local Geography Prefecture Seat Regressions (Small Pixels)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	−0.03	−0.06	−0.03	−0.14	−0.20	−0.25	−0.25	−0.33	−0.14	−0.14	−0.15
	(0.02)*	(0.02)**	(0.03)	(0.04)***	(0.05)***	(0.06)***	(0.06)***	(0.06)***	(0.05)***	(0.05)***	(0.05)***
	[0.02]*	[0.02]**	[0.03]	[0.04]***	[0.04]***	[0.06]***	[0.07]***	[0.06]***	[0.05]***	[0.05]***	[0.05]***
Dist. Coast	−0.00	0.01	0.03	0.12	0.17	0.08	0.07	0.09	0.04	0.03	0.01
	(0.01)	(0.02)	(0.02)	(0.03)***	(0.03)***	(0.03)**	(0.04)*	(0.04)**	(0.03)	(0.03)	(0.03)
	[0.01]	[0.02]	[0.02]	[0.03]***	[0.03]***	[0.04]**	[0.04]*	[0.04]**	[0.02]	[0.03]	[0.03]
Dist. River	0.00	−0.03	−0.03	−0.11	−0.20	−0.22	−0.14	−0.19	−0.08	−0.08	−0.09
	(0.03)	(0.04)	(0.05)	(0.07)*	(0.06)***	(0.07)***	(0.08)*	(0.08)**	(0.07)	(0.07)	(0.08)
	[0.03]	[0.04]	[0.05]	[0.07]*	[0.06]***	[0.08]***	[0.09]	[0.09]**	[0.07]	[0.07]	[0.07]
Ruggedness	−0.04	−0.07	−0.06	−0.13	−0.12	−0.29	−0.20	−0.17	−0.31	−0.28	−0.20
	(0.01)***	(0.02)***	(0.02)**	(0.04)***	(0.06)**	(0.07)***	(0.07)***	(0.07)**	(0.06)***	(0.06)***	(0.07)***
	[0.02]***	[0.02]***	[0.02]**	[0.05]***	[0.06]**	[0.08]***	[0.08]**	[0.08]**	[0.07]***	[0.06]***	[0.07]***
Temperature	−0.01	0.01	0.04	−0.00	−0.03	−0.16	−0.11	−0.17	−0.03	−0.02	−0.05
	(0.02)	(0.02)	(0.02)*	(0.04)	(0.04)	(0.05)***	(0.06)**	(0.06)***	(0.04)	(0.04)	(0.05)
	[0.02]	[0.02]	[0.02]*	[0.04]	[0.04]	[0.06]***	[0.06]*	[0.06]***	[0.04]	[0.05]	[0.04]
Temperature^2	0.02	−0.03	−0.05	−0.16	−0.18	0.61	0.20	−0.00	−0.13	−0.06	−0.08
	(0.06)	(0.07)	(0.08)	(0.14)	(0.14)	(0.22)***	(0.21)	(0.22)	(0.19)	(0.19)	(0.18)
	[0.05]	[0.07]	[0.08]	[0.14]	[0.15]	[0.24]**	[0.23]	[0.23]	[0.22]	[0.22]	[0.18]
Precipitation	−0.06	−0.12	−0.06	−0.13	−0.07	−0.15	−0.14	−0.16	−0.02	−0.06	−0.06
	(0.03)**	(0.04)***	(0.04)	(0.07)**	(0.07)	(0.08)*	(0.09)	(0.09)*	(0.07)	(0.07)	(0.08)
	[0.03]**	[0.04]***	[0.04]	[0.06]**	[0.06]	[0.08]*	[0.09]	[0.09]*	[0.07]	[0.07]	[0.06]
Precipitation^2	0.10	0.23	0.10	0.24	0.08	0.20	0.18	0.22	−0.04	0.05	0.05
	(0.05)*	(0.08)***	(0.09)	(0.14)*	(0.15)	(0.16)	(0.18)	(0.17)	(0.14)	(0.15)	(0.15)
	[0.05]**	[0.08]***	[0.09]	[0.14]*	[0.13]	[0.17]	[0.18]	[0.18]	[0.15]	[0.15]	[0.13]
Elevation	−0.12	−0.14	−0.07	−0.56	−1.17	−0.96	−1.06	−1.44	−0.09	−0.11	−0.43
	(0.08)	(0.11)	(0.13)	(0.20)***	(0.21)***	(0.26)***	(0.27)***	(0.28)***	(0.23)	(0.24)	(0.24)*
	[0.08]	[0.10]	[0.11]	[0.20]***	[0.21]***	[0.27]***	[0.29]***	[0.29]***	[0.24]	[0.26]	[0.24]*
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.00	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.02
F Stat.	1.04	1.94	2.18	3.55	4.21	4.06	3.78	3.68	2.79	2.88	3.41
Num. obs.	21,597	21,597	21,597	21,597	21,597	21,597	21,597	21,597	21,597	21,597	21,597

Notes: The table reports the regression results of eq. (14) using the prefecture seats. The dependent variable is an indicator that equals one, if the pixel hosts a prefecture seat in that year, and zero otherwise. Heteroskedasticity-robust standard errors are in parentheses and Conley standard errors with a radius of 150 km and a Bartlett kernel are in brackets ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index $\times$ 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-7: Local Geography County Seat Regressions (Medium Pixels)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	−0.23	−0.37	−0.37	−0.70	−0.93	−1.22	−0.80	−0.76	−0.85	−0.99	−0.87
	(0.13)*	(0.17)**	(0.17)**	(0.17)***	(0.18)***	(0.21)***	(0.20)***	(0.20)***	(0.20)***	(0.22)***	(0.22)***
	[0.18]	[0.25]	[0.24]	[0.25]***	[0.29]***	[0.32]***	[0.28]***	[0.26]***	[0.25]***	[0.27]***	[0.26]***
Dist. Coast	−0.14	−0.32	0.07	0.45	0.80	0.86	0.89	0.75	0.30	0.49	0.31
	(0.10)	(0.11)***	(0.11)	(0.11)***	(0.12)***	(0.12)***	(0.13)***	(0.13)***	(0.12)**	(0.13)***	(0.13)**
	[0.16]	[0.21]	[0.18]	[0.17]***	[0.22]***	[0.24]***	[0.23]***	[0.22]***	[0.18]*	[0.19]***	[0.18]*
Dist. River	0.07	−0.13	−0.17	−0.66	−1.49	−0.79	−0.25	−0.10	−0.36	−0.58	−0.26
	(0.20)	(0.24)	(0.24)	(0.25)***	(0.25)***	(0.29)***	(0.29)	(0.30)	(0.30)	(0.31)*	(0.33)
	[0.37]	[0.44]	[0.41]	[0.42]	[0.47]***	[0.55]	[0.55]	[0.54]	[0.52]	[0.53]	[0.50]
Ruggedness	−0.50	−0.57	−0.77	−1.28	−0.66	−1.21	−1.35	−1.33	−1.32	−1.42	−1.42
	(0.15)***	(0.21)***	(0.21)***	(0.21)***	(0.24)***	(0.26)***	(0.26)***	(0.26)***	(0.27)***	(0.28)***	(0.28)***
	[0.20]**	[0.31]*	[0.29]***	[0.27]***	[0.35]*	[0.39]***	[0.39]***	[0.37]***	[0.36]***	[0.37]***	[0.35]***
Temperature	0.58	0.81	0.61	0.27	0.12	−0.48	0.00	0.10	0.07	0.07	0.13
	(0.12)***	(0.16)***	(0.15)***	(0.15)*	(0.17)	(0.19)**	(0.19)	(0.19)	(0.18)	(0.20)	(0.20)
	[0.19]***	[0.27]***	[0.25]**	[0.23]	[0.29]	[0.31]	[0.27]	[0.25]	[0.24]	[0.26]	[0.26]
Temperature^2	−0.25	0.35	−0.14	−1.40	2.42	3.96	2.18	1.36	0.94	0.42	0.20
	(0.42)	(0.57)	(0.58)	(0.61)**	(0.67)***	(0.79)***	(0.75)***	(0.76)*	(0.76)	(0.82)	(0.84)
	[0.79]	[1.10]	[1.02]	[1.00]	[1.16]**	[1.38]***	[1.26]*	[1.22]	[1.21]	[1.26]	[1.21]
Precipitation	−1.15	−1.86	−0.84	−0.16	−1.31	−0.15	0.55	0.59	0.14	0.40	0.45
	(0.22)***	(0.26)***	(0.25)***	(0.26)	(0.29)***	(0.31)	(0.30)*	(0.30)**	(0.31)	(0.32)	(0.33)
	[0.36]***	[0.42]***	[0.36]**	[0.38]	[0.52]**	[0.56]	[0.52]	[0.48]	[0.43]	[0.46]	[0.44]
Precipitation^2	1.86	3.07	1.15	0.10	2.04	−0.32	−1.58	−1.69	−0.92	−1.55	−1.55
	(0.45)***	(0.53)***	(0.52)**	(0.53)	(0.60)***	(0.63)	(0.61)***	(0.62)***	(0.62)	(0.66)**	(0.67)**
	[0.70]***	[0.82]***	[0.73]	[0.78]	[1.06]*	[1.16]	[1.13]	[1.06]	[0.90]	[0.97]	[0.91]*
Elevation	−1.95	−1.33	−2.01	−4.12	−6.42	−7.31	−6.05	−5.39	−4.23	−5.17	−4.56
	(0.62)***	(0.81)*	(0.80)**	(0.79)***	(0.83)***	(0.96)***	(0.93)***	(0.94)***	(0.94)***	(1.03)***	(1.02)***
	[0.87]**	[1.14]	[1.07]*	[1.10]***	[1.28]***	[1.38]***	[1.23]***	[1.17]***	[1.15]***	[1.25]***	[1.19]***
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.17	0.20	0.17	0.12	0.19	0.18	0.18	0.17	0.14	0.15	0.14
F Stat.	13.35	16.63	13.56	9.12	15.82	14.98	14.24	13.35	11.20	11.60	11.27
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the county seats. The dependent variable is an indicator that equals one, if the pixel hosts a county seat in that year, and zero otherwise. Heteroskedasticity-robust standard errors are in parentheses and Conley standard errors with a radius of 150 km and a Bartlett kernel are in brackets ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-8: Local Geography Prefecture Seat Regressions (Medium Pixels)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	−0.08	−0.10	−0.04	−0.27	−0.36	−0.53	−0.46	−0.66	−0.31	−0.32	−0.35
	(0.04)*	(0.06)*	(0.06)	(0.10)***	(0.10)***	(0.13)***	(0.14)***	(0.14)***	(0.12)***	(0.12)***	(0.12)***
	[0.04]*	[0.05]*	[0.06]	[0.10]***	[0.09]***	[0.14]***	[0.15]***	[0.14]***	[0.12]**	[0.12]***	[0.11]***
Dist. Coast	0.02	0.03	0.05	0.23	0.40	0.22	0.22	0.33	0.16	0.16	0.10
	(0.03)	(0.04)	(0.05)	(0.07)***	(0.07)***	(0.08)***	(0.08)***	(0.09)***	(0.07)**	(0.07)**	(0.07)
	[0.03]	[0.04]	[0.05]	[0.07]***	[0.07]***	[0.08]***	[0.10]**	[0.09]***	[0.07]**	[0.07]**	[0.07]
Dist. River	−0.03	−0.04	−0.07	−0.21	−0.36	−0.50	−0.30	−0.39	−0.15	−0.18	−0.09
	(0.07)	(0.09)	(0.10)	(0.15)	(0.14)***	(0.16)***	(0.17)*	(0.18)**	(0.16)	(0.17)	(0.17)
	[0.07]	[0.09]	[0.10]	[0.15]	[0.13]***	[0.17]***	[0.20]	[0.20]**	[0.16]	[0.16]	[0.15]
Ruggedness	0.00	−0.05	−0.12	−0.32	−0.18	−0.49	−0.46	−0.39	−0.51	−0.47	−0.41
	(0.05)	(0.06)	(0.07)*	(0.10)***	(0.13)	(0.15)***	(0.17)***	(0.16)**	(0.16)***	(0.16)***	(0.16)**
	[0.05]	[0.06]	[0.07]*	[0.11]***	[0.12]	[0.17]***	[0.19]**	[0.18]**	[0.17]***	[0.18]***	[0.18]**
Temperature	−0.03	0.04	0.16	−0.03	−0.04	−0.31	−0.17	−0.29	−0.06	−0.02	−0.08
	(0.04)	(0.05)	(0.06)***	(0.09)	(0.09)	(0.12)***	(0.13)	(0.13)**	(0.11)	(0.11)	(0.11)
	[0.04]	[0.05]	[0.06]***	[0.10]	[0.09]	[0.12]**	[0.14]	[0.13]**	[0.12]	[0.12]	[0.10]
Temperature^2	−0.01	−0.22	−0.35	−0.10	0.00	1.29	0.63	0.11	−0.06	−0.01	0.06
	(0.16)	(0.19)	(0.21)*	(0.32)	(0.33)	(0.49)***	(0.48)	(0.50)	(0.44)	(0.45)	(0.44)
	[0.15]	[0.19]	[0.20]*	[0.31]	[0.34]	[0.50]***	[0.52]	[0.52]	[0.48]	[0.48]	[0.43]
Precipitation	−0.08	−0.22	−0.22	−0.27	−0.15	−0.29	−0.25	−0.31	0.05	−0.15	−0.22
	(0.05)	(0.07)***	(0.09)**	(0.13)**	(0.15)	(0.18)	(0.19)	(0.19)*	(0.16)	(0.17)	(0.17)
	[0.05]	[0.07]***	[0.09]**	[0.13]**	[0.14]	[0.20]	[0.20]	[0.20]	[0.17]	[0.17]	[0.15]
Precipitation^2	0.13	0.41	0.39	0.49	0.23	0.46	0.35	0.51	−0.24	0.16	0.31
	(0.11)	(0.15)***	(0.20)*	(0.26)*	(0.31)	(0.37)	(0.38)	(0.38)	(0.33)	(0.35)	(0.34)
	[0.11]	[0.15]***	[0.19]**	[0.27]*	[0.28]	[0.40]	[0.41]	[0.41]	[0.36]	[0.35]	[0.31]
Elevation	−0.37	−0.26	0.05	−0.97	−1.99	−1.83	−1.69	−2.67	−0.11	−0.10	−0.86
	(0.20)*	(0.25)	(0.27)	(0.46)**	(0.47)***	(0.59)***	(0.61)***	(0.66)***	(0.58)	(0.59)	(0.57)
	[0.19]*	[0.24]	[0.25]	[0.42]**	[0.45]***	[0.58]***	[0.68]**	[0.68]***	[0.60]	[0.61]	[0.52]*
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.00	0.01	0.02	0.03	0.04	0.04	0.05	0.05	0.03	0.03	0.03
F Stat.	1.28	1.86	2.56	2.83	3.65	3.48	3.98	4.04	2.72	2.75	3.04
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the prefecture seats. The dependent variable is an indicator that equals one, if the pixel hosts a prefecture seat in that year, and zero otherwise. Heteroskedasticity-robust standard errors are in parentheses and Conley standard errors with a radius of 150 km and a Bartlett kernel are in brackets ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-9: Local Geography County Seat Regressions (Large Pixels)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	−0.22	−0.65	−0.52	−1.15	−1.55	−1.94	−1.05	−1.23	−1.38	−1.61	−1.73
	(0.23)	(0.28)**	(0.29)*	(0.29)***	(0.30)***	(0.34)***	(0.33)***	(0.33)***	(0.34)***	(0.36)***	(0.37)***
	[0.29]	[0.37]*	[0.37]	[0.40]***	[0.46]***	[0.45]***	[0.42]**	[0.41]***	[0.42]***	[0.42]***	[0.41]***
Dist. Coast	−0.07	−0.27	0.09	0.51	1.20	1.06	1.02	0.97	0.39	0.60	0.41
	(0.16)	(0.18)	(0.18)	(0.18)***	(0.19)***	(0.19)***	(0.20)***	(0.20)***	(0.20)**	(0.20)***	(0.21)*
	[0.24]	[0.27]	[0.26]	[0.27]*	[0.33]***	[0.34]***	[0.33]***	[0.32]***	[0.28]	[0.29]**	[0.28]
Dist. River	−0.23	−0.70	−0.69	−1.43	−2.22	−1.62	−0.85	−0.80	−1.15	−1.42	−0.81
	(0.34)	(0.39)*	(0.40)*	(0.41)***	(0.42)***	(0.46)***	(0.46)*	(0.46)*	(0.46)**	(0.47)***	(0.50)
	[0.53]	[0.59]	[0.59]	[0.62]**	[0.68]***	[0.74]**	[0.72]	[0.72]	[0.70]*	[0.72]**	[0.67]
Ruggedness	−0.61	−0.81	−1.00	−1.88	−0.59	−1.70	−1.69	−1.84	−1.93	−1.73	−2.06
	(0.26)**	(0.37)**	(0.37)***	(0.37)***	(0.41)	(0.44)***	(0.43)***	(0.43)***	(0.45)***	(0.46)***	(0.46)***
	[0.33]*	[0.48]*	[0.48]**	[0.45]***	[0.58]	[0.57]***	[0.55]***	[0.57]***	[0.53]***	[0.56]***	[0.56]***
Temperature	1.11	1.14	1.07	0.38	−0.13	−0.80	0.13	0.03	−0.05	−0.11	−0.12
	(0.21)***	(0.27)***	(0.27)***	(0.26)	(0.28)	(0.31)***	(0.30)	(0.30)	(0.30)	(0.31)	(0.32)
	[0.29]***	[0.38]***	[0.37]***	[0.37]	[0.44]	[0.43]*	[0.40]	[0.39]	[0.38]	[0.38]	[0.37]
Temperature^2	−1.51	−0.75	−1.10	−2.05	3.65	4.49	2.63	1.96	0.66	0.92	0.59
	(0.74)**	(0.98)	(0.99)	(1.06)*	(1.12)***	(1.26)***	(1.25)**	(1.25)	(1.26)	(1.33)	(1.35)
	[1.22]	[1.55]	[1.51]	[1.60]	[1.77]**	[1.95]**	[1.80]	[1.76]	[1.76]	[1.82]	[1.76]
Precipitation	−1.34	−2.44	−1.40	−0.52	−1.72	−0.51	0.58	0.67	0.40	0.61	0.31
	(0.35)***	(0.41)***	(0.41)***	(0.42)	(0.44)***	(0.48)	(0.47)	(0.47)	(0.47)	(0.50)	(0.51)
	[0.53]**	[0.58]***	[0.55]**	[0.60]	[0.73]**	[0.76]	[0.70]	[0.69]	[0.63]	[0.66]	[0.64]
Precipitation^2	2.11	4.06	2.11	0.66	2.58	−0.05	−2.01	−2.20	−1.87	−2.44	−1.75
	(0.72)***	(0.84)***	(0.84)**	(0.85)	(0.91)***	(0.99)	(0.97)**	(0.97)**	(0.98)*	(1.02)**	(1.03)*
	[1.06]**	[1.15]***	[1.10]*	[1.24]	[1.48]*	[1.55]	[1.46]	[1.43]	[1.33]	[1.36]*	[1.30]
Elevation	−2.62	−2.84	−3.21	−6.91	−1163	−1249	−9.09	−9.34	−8.88	−9.90	−9.90
	(1.04)**	(1.33)**	(1.34)**	(1.33)***	(1.42)***	(1.54)***	(1.52)***	(1.52)***	(1.56)***	(1.69)***	(1.70)***
	[1.39]*	[1.70]*	[1.66]*	[1.76]***	[2.17]***	[2.07]***	[1.96]***	[1.94]***	[1.99]***	[2.04]***	[2.00]***
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.21	0.24	0.21	0.17	0.26	0.26	0.24	0.24	0.22	0.23	0.22
F Stat.	10.47	12.35	10.77	8.45	13.67	13.27	12.55	12.33	10.82	11.80	11.30
Num. obs.	5,402	5,402	5,402	5,402	5,402	5,402	5,402	5,402	5,402	5,402	5,402

Notes: The table reports the regression results of eq. (14) using the county seats. The dependent variable is an indicator that equals one, if the pixel hosts a county seat in that year, and zero otherwise. Heteroskedasticity-robust standard errors are in parentheses and Conley standard errors with a radius of 150 km and a Bartlett kernel are in brackets ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-10: Local Geography Prefecture Seat Regressions (Large Pixels)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	−0.08	−0.14	−0.08	−0.53	−0.67	−0.93	−0.78	−1.16	−0.54	−0.62	−0.56
	(0.08)	(0.11)	(0.13)	(0.18)***	(0.20)***	(0.23)***	(0.25)***	(0.26)***	(0.22)**	(0.21)***	(0.22)**
	[0.08]	[0.11]	[0.13]	[0.19]***	[0.18]***	[0.24]***	[0.26]***	[0.26]***	[0.23]**	[0.22]***	[0.20]***
Dist. Coast	0.02	0.03	0.09	0.42	0.58	0.33	0.36	0.51	0.18	0.14	−0.02
	(0.05)	(0.07)	(0.09)	(0.12)***	(0.13)***	(0.13)**	(0.15)**	(0.15)***	(0.12)	(0.12)	(0.13)
	[0.05]	[0.07]	[0.09]	[0.13]***	[0.12]***	[0.15]**	[0.17]**	[0.17]***	[0.11]	[0.12]	[0.13]
Dist. River	0.00	0.05	−0.07	−0.49	−0.35	−0.58	−0.34	−0.57	−0.13	−0.38	−0.17
	(0.12)	(0.17)	(0.19)	(0.27)*	(0.25)	(0.31)*	(0.34)	(0.33)*	(0.29)	(0.29)	(0.31)
	[0.11]	[0.17]	[0.20]	[0.27]*	[0.25]	[0.31]*	[0.36]	[0.35]*	[0.28]	[0.30]	[0.29]
Ruggedness	−0.21	−0.13	−0.38	−0.87	−0.50	−0.93	−0.53	−0.65	−1.06	−1.10	−0.87
	(0.08)**	(0.10)	(0.15)***	(0.20)***	(0.26)*	(0.28)***	(0.30)*	(0.31)**	(0.27)***	(0.27)***	(0.30)***
	[0.09]**	[0.10]	[0.15]**	[0.23]***	[0.26]*	[0.32]***	[0.34]	[0.35]*	[0.27]***	[0.28]***	[0.33]***
Temperature	0.03	0.08	0.27	0.10	−0.10	−0.48	−0.20	−0.43	−0.05	−0.05	−0.02
	(0.07)	(0.10)	(0.12)**	(0.16)	(0.17)	(0.20)**	(0.22)	(0.22)*	(0.18)	(0.18)	(0.19)
	[0.07]	[0.10]	[0.12]**	[0.17]	[0.17]	[0.21]**	[0.23]	[0.22]**	[0.18]	[0.19]	[0.18]
Temperature^2	−0.14	−0.23	−0.62	−0.50	−0.04	1.75	0.24	−0.76	−0.78	−0.60	−0.59
	(0.26)	(0.35)	(0.39)	(0.59)	(0.62)	(0.84)**	(0.88)	(0.88)	(0.75)	(0.75)	(0.73)
	[0.25]	[0.32]	[0.37]*	[0.60]	[0.61]	[0.90]*	[0.93]	[0.91]	[0.76]	[0.77]	[0.70]
Precipitation	−0.21	−0.40	−0.47	−0.91	−0.41	−0.49	−0.43	−0.66	−0.12	−0.35	−0.36
	(0.11)*	(0.16)**	(0.20)**	(0.25)***	(0.29)	(0.32)	(0.34)	(0.34)*	(0.28)	(0.28)	(0.29)
	[0.11]*	[0.16]**	[0.19]**	[0.27]***	[0.27]	[0.34]	[0.33]	[0.34]*	[0.28]	[0.28]	[0.27]
Precipitation^2	0.37	0.72	0.87	1.76	0.69	0.55	0.48	1.02	0.01	0.46	0.43
	(0.23)	(0.33)**	(0.41)**	(0.52)***	(0.59)	(0.66)	(0.70)	(0.69)	(0.58)	(0.58)	(0.60)
	[0.22]*	[0.33]**	[0.40]**	[0.59]***	[0.55]	[0.68]	[0.70]	[0.71]	[0.59]	[0.58]	[0.55]
Elevation	−0.30	−0.36	−0.07	−1.74	−3.56	−3.40	−3.22	−5.12	−0.58	−0.85	−1.59
	(0.32)	(0.45)	(0.56)	(0.79)**	(0.85)***	(1.04)***	(1.15)***	(1.19)***	(1.02)	(1.03)	(1.05)
	[0.31]	[0.45]	[0.54]	[0.76]**	[0.83]***	[1.06]***	[1.22]***	[1.22]***	[1.05]	[1.08]	[0.98]
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.00	0.02	0.03	0.03	0.06	0.06	0.07	0.07	0.03	0.03	0.03
F Stat.	1.01	1.70	1.87	2.23	3.29	3.40	3.51	3.45	2.18	2.05	2.14
Num. obs.	5,402	5,402	5,402	5,402	5,402	5,402	5,402	5,402	5,402	5,402	5,402

Notes: The table reports the regression results of eq. (14) using the prefecture seats. The dependent variable is an indicator that equals one, if the pixel hosts a prefecture seat in that year, and zero otherwise. Heteroskedasticity-robust standard errors are in parentheses and Conley standard errors with a radius of 150 km and a Bartlett kernel are in brackets ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-11: Local Geography County Seat Regressions (Very Large Pixels)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	0.18	−0.74	−0.53	−1.00	−1.81	−2.43	−0.85	−1.09	−1.45	−1.51	−1.39
	(0.35)	(0.43)*	(0.43)	(0.43)**	(0.43)***	(0.50)***	(0.48)*	(0.49)**	(0.48)***	(0.52)***	(0.52)***
	[0.42]	[0.53]	[0.52]	[0.56]*	[0.58]***	[0.65]***	[0.60]	[0.59]*	[0.60]**	[0.60]**	[0.60]**
Dist. Coast	−0.30	−0.36	−0.01	0.68	1.40	1.29	1.10	1.13	0.59	0.95	0.54
	(0.23)	(0.25)	(0.25)	(0.25)***	(0.25)***	(0.26)***	(0.26)***	(0.26)***	(0.26)**	(0.27)***	(0.28)*
	[0.33]	[0.37]	[0.36]	[0.35]*	[0.40]***	[0.40]***	[0.39]***	[0.38]***	[0.34]*	[0.36]***	[0.33]*
Dist. River	0.16	−0.73	−1.10	−1.98	−2.80	−2.02	−0.94	−0.91	−0.92	−1.23	−0.73
	(0.47)	(0.55)	(0.57)*	(0.60)***	(0.57)***	(0.65)***	(0.65)	(0.64)	(0.63)	(0.65)*	(0.66)
	[0.66]	[0.75]	[0.74]	[0.85]**	[0.81]***	[0.96]**	[0.96]	[0.95]	[0.94]	[0.99]	[0.87]
Ruggedness	−1.01	−1.03	−1.12	−2.12	−0.44	−1.84	−1.22	−1.15	−1.94	−1.90	−2.57
	(0.43)**	(0.63)	(0.64)*	(0.57)***	(0.64)	(0.68)***	(0.66)*	(0.66)*	(0.68)***	(0.69)***	(0.69)***
	[0.48]**	[0.76]	[0.76]	[0.63]***	[0.80]	[0.84]**	[0.81]	[0.79]	[0.81]**	[0.82]**	[0.77]***
Temperature	1.93	1.70	1.84	1.43	0.28	−0.71	0.50	0.56	0.54	0.69	0.95
	(0.35)***	(0.42)***	(0.42)***	(0.41)***	(0.42)	(0.46)	(0.46)	(0.47)	(0.46)	(0.48)	(0.48)*
	[0.45]***	[0.58]***	[0.58]***	[0.59]**	[0.59]	[0.62]	[0.59]	[0.58]	[0.58]	[0.59]	[0.57]*
Temperature^2	−2.18	−1.77	−2.83	−4.36	2.77	4.60	4.45	2.31	−0.12	−0.04	−1.21
	(1.12)*	(1.43)	(1.43)**	(1.50)***	(1.54)*	(1.74)***	(1.77)**	(1.82)	(1.80)	(1.88)	(1.91)
	[1.61]	[2.07]	[2.05]	[2.15]**	[2.16]	[2.46]*	[2.39]*	[2.51]	[2.43]	[2.48]	[2.43]
Precipitation	−1.56	−2.86	−1.82	−0.64	−1.79	−0.12	1.56	1.43	0.81	1.35	1.11
	(0.49)***	(0.58)***	(0.59)***	(0.60)	(0.61)***	(0.67)	(0.67)**	(0.67)**	(0.68)	(0.69)*	(0.70)
	[0.65]**	[0.73]***	[0.74]**	[0.79]	[0.87]**	[0.97]	[0.95]*	[0.94]	[0.88]	[0.93]	[0.88]
Precipitation^2	2.00	4.22	2.35	0.15	2.07	−1.90	−5.18	−5.01	−3.92	−5.27	−4.58
	(1.03)*	(1.21)***	(1.23)*	(1.23)	(1.27)	(1.41)	(1.47)***	(1.45)***	(1.47)***	(1.51)***	(1.51)***
	[1.34]	[1.48]***	[1.51]	[1.63]	[1.77]	[2.05]	[2.12]**	[2.08]**	[1.97]**	[2.04]**	[1.90]**
Elevation	−1.62	−3.45	−3.42	−7.91	−1350	−1598	−1020	−1102	−9.97	−1099	−9.61
	(1.61)	(2.01)*	(2.02)*	(2.01)***	(2.02)***	(2.30)***	(2.22)***	(2.26)***	(2.25)***	(2.43)***	(2.42)***
	[1.96]	[2.51]	[2.44]	[2.51]***	[2.70]***	[2.96]***	[2.80]***	[2.68]***	[2.72]***	[2.87]***	[2.83]***
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.28	0.29	0.27	0.23	0.32	0.32	0.33	0.32	0.28	0.30	0.31
F Stat.	10.14	10.61	9.77	8.07	12.47	12.42	12.78	12.16	10.19	11.37	11.52
Num. obs.	3,463	3,463	3,463	3,463	3,463	3,463	3,463	3,463	3,463	3,463	3,463

Notes: The table reports the regression results of eq. (14) using the county seats. The dependent variable is an indicator that equals one, if the pixel hosts a county seat in that year, and zero otherwise. Heteroskedasticity-robust standard errors are in parentheses and Conley standard errors with a radius of 150 km and a Bartlett kernel are in brackets ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-12: Local Geography Prefecture Seat Regressions (Very Large Pixels)

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	−0.13	−0.21	0.02	−0.69	−0.55	−1.19	−1.01	−1.62	−0.33	−0.40	−0.33
	(0.12)	(0.16)	(0.18)	(0.28)**	(0.29)*	(0.38)***	(0.40)**	(0.41)***	(0.34)	(0.33)	(0.33)
	[0.11]	[0.15]	[0.17]	[0.28]**	[0.27]**	[0.39]***	[0.45]**	[0.44]***	[0.34]	[0.31]	[0.28]
Dist. Coast	0.00	−0.01	0.02	0.31	0.80	0.47	0.48	0.75	0.23	0.19	0.04
	(0.09)	(0.10)	(0.13)	(0.17)*	(0.19)***	(0.20)**	(0.21)**	(0.22)***	(0.18)	(0.18)	(0.20)
	[0.08]	[0.09]	[0.13]	[0.18]*	[0.18]***	[0.21]**	[0.24]**	[0.24]***	[0.17]	[0.18]	[0.19]
Dist. River	−0.00	0.00	−0.11	−0.83	−0.80	−1.17	−0.70	−0.96	−0.32	−0.53	−0.12
	(0.18)	(0.23)	(0.26)	(0.39)**	(0.37)**	(0.45)***	(0.48)	(0.46)**	(0.41)	(0.44)	(0.46)
	[0.16]	[0.22]	[0.25]	[0.39]**	[0.35]**	[0.45]***	[0.53]	[0.49]**	[0.40]	[0.42]	[0.41]
Ruggedness	−0.13	−0.27	−0.35	−1.11	−0.48	−1.24	−0.90	−1.25	−1.69	−1.48	−1.60
	(0.12)	(0.15)*	(0.19)*	(0.32)***	(0.40)	(0.46)***	(0.50)*	(0.50)**	(0.44)***	(0.42)***	(0.47)***
	[0.11]	[0.15]*	[0.19]*	[0.34]***	[0.38]	[0.48]**	[0.50]*	[0.53]**	[0.45]***	[0.42]***	[0.48]***
Temperature	0.05	0.21	0.58	0.13	0.22	−0.65	−0.24	−0.65	0.31	0.32	0.41
	(0.11)	(0.15)	(0.17)***	(0.25)	(0.27)	(0.35)*	(0.36)	(0.36)*	(0.31)	(0.30)	(0.31)
	[0.10]	[0.15]	[0.17]***	[0.26]	[0.27]	[0.35]*	[0.40]	[0.38]*	[0.30]	[0.28]	[0.28]
Temperature^2	0.03	−0.22	−0.66	−0.50	0.03	3.66	1.34	−0.51	−0.45	0.32	0.08
	(0.37)	(0.47)	(0.55)	(0.78)	(0.89)	(1.27)***	(1.22)	(1.19)	(1.15)	(1.16)	(1.12)
	[0.34]	[0.46]	[0.55]	[0.84]	[0.94]	[1.35]***	[1.33]	[1.23]	[1.19]	[1.24]	[1.10]
Precipitation	−0.43	−0.78	−0.69	−1.37	−0.32	−0.88	−0.50	−0.69	0.10	−0.42	−0.50
	(0.15)***	(0.21)***	(0.28)**	(0.38)***	(0.40)	(0.48)*	(0.51)	(0.50)	(0.43)	(0.43)	(0.45)
	[0.15]***	[0.21]***	[0.27]**	[0.41]***	[0.38]	[0.54]	[0.54]	[0.54]	[0.46]	[0.44]	[0.42]
Precipitation^2	0.78	1.35	1.14	2.42	0.25	0.96	0.24	0.71	−0.91	0.04	0.38
	(0.31)**	(0.41)***	(0.58)**	(0.78)***	(0.83)	(0.99)	(1.06)	(1.05)	(0.90)	(0.88)	(0.92)
	[0.30]***	[0.42]***	[0.56]**	[0.84]***	[0.78]	[1.09]	[1.17]	[1.17]	[1.02]	[0.93]	[0.90]
Elevation	−0.40	−0.23	0.85	−2.45	−3.91	−4.35	−4.31	−7.21	1.38	1.06	0.50
	(0.53)	(0.72)	(0.80)	(1.36)*	(1.36)***	(1.75)**	(1.78)**	(1.86)***	(1.63)	(1.59)	(1.59)
	[0.49]	[0.67]	[0.74]	[1.32]*	[1.32]***	[1.69]**	[1.96]**	[1.91]***	[1.56]	[1.48]	[1.34]
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.00	0.03	0.04	0.05	0.08	0.09	0.10	0.09	0.04	0.06	0.06
F Stat.	1.07	1.75	2.03	2.36	3.11	3.37	3.70	3.47	2.02	2.40	2.54
Num. obs.	3,463	3,463	3,463	3,463	3,463	3,463	3,463	3,463	3,463	3,463	3,463

Notes: The table reports the regression results of eq. (14) using the prefecture seats. The dependent variable is an indicator that equals one, if the pixel hosts a prefecture seat in that year, and zero otherwise. Heteroskedasticity-robust standard errors are in parentheses and Conley standard errors with a radius of 150 km and a Bartlett kernel are in brackets ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

C.3 Robustness Check: Alternative Radii in Conley Standard Errors

Here we provide a robustness check for the main results in Section 4.1 by varying the radii of the Conley standard errors. While the baseline estimations use radii of 150 km in Conley standard error computations, we now test alternative radii of 50, 100, and 500 km. Pixels are weighted using a Bartlett kernel where weights decrease with distance and are zero beyond the specified radius. The "right" radius depends on transportation technology and infrastructure and the frequency of trips. Today, a 150 km radius would be inappropriately small. Motorised vehicles cover that distance in probably two hours and many people regularly visit places beyond that threshold. In historic times, when much of transport between many places was done on foot and taxes had to be paid in kind through much of the imperial period, 150 km was very far (von Glahn, 2016). Nonetheless, we can see that the choice of the radii changes the standard errors of most coefficients by very little, leaving the overall insights intact.

Table C-13: Local Geography County Seat Regressions with Alternative Radii

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	−0.23	−0.37	−0.37	−0.70	−0.93	−1.22	−0.80	−0.76	−0.85	−0.99	−0.87
	(0.14)*	(0.18)**	(0.18)**	(0.18)***	(0.20)***	(0.23)***	(0.21)***	(0.21)***	(0.21)***	(0.23)***	(0.22)***
	[0.16]	[0.22]*	[0.21]*	[0.22]***	[0.25]***	[0.28]***	[0.25]***	[0.23]***	[0.24]***	[0.25]***	[0.24]***
	{0.24}	{0.34}	{0.31}	{0.31}**	{0.39}**	{0.44}***	{0.38}**	{0.31}**	{0.29}***	{0.30}***	{0.29}***
Dist. Coast	−0.14	−0.32	0.07	0.45	0.80	0.86	0.89	0.75	0.30	0.49	0.31
	(0.11)	(0.13)**	(0.12)	(0.12)***	(0.14)***	(0.14)***	(0.14)***	(0.14)***	(0.13)**	(0.13)***	(0.14)**
	[0.14]	[0.17]*	[0.15]	[0.15]***	[0.18]***	[0.19]***	[0.19]***	[0.18]***	[0.15]*	[0.16]***	[0.16]*
	{0.25}	{0.35}	{0.27}	{0.23}*	{0.39}**	{0.43}**	{0.42}**	{0.38}**	{0.25}	{0.28}*	{0.26}
Dist. River	0.07	−0.13	−0.17	−0.66	−1.49	−0.79	−0.25	−0.10	−0.36	−0.58	−0.26
	(0.23)	(0.28)	(0.27)	(0.28)**	(0.29)***	(0.34)**	(0.34)	(0.33)	(0.33)	(0.34)*	(0.34)
	[0.31]	[0.37]	[0.35]	[0.35]*	[0.39]***	[0.45]*	[0.45]	[0.44]	[0.43]	[0.44]	[0.42]
	{0.53}	{0.56}	{0.53}	{0.61}	{0.69}**	{0.82}	{0.82}	{0.80}	{0.79}	{0.80}	{0.78}
Ruggedness	−0.50	−0.57	−0.77	−1.28	−0.66	−1.21	−1.35	−1.33	−1.32	−1.42	−1.42
	(0.15)***	(0.23)**	(0.23)***	(0.22)***	(0.26)**	(0.29)***	(0.29)***	(0.29)***	(0.29)***	(0.30)***	(0.29)***
	[0.18]***	[0.28]**	[0.27]***	[0.24]***	[0.32]**	[0.35]***	[0.35]***	[0.34]***	[0.34]***	[0.34]***	[0.32]***
	{0.24}**	{0.35}	{0.36}**	{0.34}***	{0.35}*	{0.45}***	{0.50}***	{0.46}***	{0.44}***	{0.43}***	{0.42}***
Temperature	0.58	0.81	0.61	0.27	0.12	−0.48	0.00	0.10	0.07	0.07	0.13
	(0.13)***	(0.18)***	(0.17)***	(0.16)*	(0.19)	(0.22)**	(0.20)	(0.20)	(0.19)	(0.21)	(0.21)
	[0.16]***	[0.23]***	[0.21]***	[0.19]	[0.24]	[0.27]*	[0.24]	[0.23]	[0.22]	[0.24]	[0.24]
	{0.30}**	{0.46}*	{0.40}	{0.32}	{0.43}	{0.45}	{0.36}	{0.34}	{0.33}	{0.36}	{0.36}
Temperature^2	−0.25	0.35	−0.14	−1.40	2.42	3.96	2.18	1.36	0.94	0.42	0.20
	(0.48)	(0.66)	(0.64)	(0.67)**	(0.75)***	(0.89)***	(0.83)***	(0.82)*	(0.83)	(0.88)	(0.88)
	[0.64]	[0.88]	[0.83]	[0.83]*	[0.96]**	[1.14]***	[1.04]**	[1.03]	[1.02]	[1.08]	[1.05]
	{1.56}	{2.16}	{1.94}	{1.70}	{1.81}	{2.16}*	{1.99}	{1.95}	{1.96}	{2.04}	{1.86}
Precipitation	−1.15	−1.86	−0.84	−0.16	−1.31	−0.15	0.55	0.59	0.14	0.40	0.45
	(0.25)***	(0.29)***	(0.27)***	(0.28)	(0.33)***	(0.36)	(0.33)*	(0.33)*	(0.32)	(0.34)	(0.33)
	[0.31]***	[0.37]***	[0.32]***	[0.34]	[0.43]***	[0.47]	[0.43]	[0.41]	[0.38]	[0.41]	[0.39]
	{0.58}**	{0.65}***	{0.51}	{0.50}	{0.91}	{0.90}	{0.93}	{0.80}	{0.68}	{0.70}	{0.68}
Precipitation^2	1.86	3.07	1.15	0.10	2.04	−0.32	−1.58	−1.69	−0.92	−1.55	−1.55
	(0.50)***	(0.60)***	(0.56)**	(0.58)	(0.68)***	(0.74)	(0.70)**	(0.69)**	(0.67)	(0.71)**	(0.69)**
	[0.62]***	[0.73]***	[0.66]*	[0.68]	[0.89]**	[0.96]	[0.92]*	[0.88]*	[0.79]	[0.86]*	[0.81]*
	{1.11}*	{1.22}**	{1.00}	{0.98}	{1.79}	{1.84}	{1.95}	{1.67}	{1.38}	{1.42}	{1.36}
Elevation	−1.95	−1.33	−2.01	−4.12	−6.42	−7.31	−6.05	−5.39	−4.23	−5.17	−4.56
	(0.65)***	(0.86)	(0.85)**	(0.84)***	(0.91)***	(1.07)***	(0.99)***	(0.96)***	(0.97)***	(1.05)***	(1.02)***
	[0.76]**	[1.01]	[0.97]**	[0.98]***	[1.11]***	[1.23]***	[1.12]***	[1.08]***	[1.07]***	[1.16]***	[1.09]***
	{1.25}	{1.47}	{1.28}	{1.44}***	{1.95}***	{2.00}***	{1.71}***	{1.44}***	{1.29}***	{1.48}***	{1.47}***
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.17	0.20	0.17	0.12	0.19	0.18	0.18	0.17	0.14	0.15	0.14
F Stat.	13.35	16.63	13.56	9.12	15.82	14.98	14.24	13.35	11.20	11.60	11.27
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the county seats. The dependent variable is an indicator that equals one, if the pixel hosts a county seat in that year, and zero otherwise. The estimations utilise the baseline sample. Conley standard errors using a 50 km radius are in parentheses, using a 100 km radius in brackets, and using a 500 km radius in curly braces ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-14: Local Geography Prefecture Seat Regressions with Alternative Radii

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	−0.08	−0.10	−0.04	−0.27	−0.36	−0.53	−0.46	−0.66	−0.31	−0.32	−0.35
	(0.04)*	(0.05)*	(0.06)	(0.10)***	(0.10)***	(0.12)***	(0.13)***	(0.14)***	(0.12)***	(0.12)***	(0.12)***
	[0.04]*	[0.05]*	[0.06]	[0.10]***	[0.09]***	[0.12]***	[0.14]***	[0.14]***	[0.12]***	[0.12]***	[0.11]***
	{0.03}**	{0.05}**	{0.06}	{0.12}**	{0.10}***	{0.17}***	{0.17}***	{0.15}***	{0.12}**	{0.11}***	{0.10}***
Dist. Coast	0.02	0.03	0.05	0.23	0.40	0.22	0.22	0.33	0.16	0.16	0.10
	(0.03)	(0.04)	(0.05)	(0.06)***	(0.07)***	(0.07)***	(0.08)***	(0.08)***	(0.06)**	(0.06)**	(0.07)
	[0.03]	[0.04]	[0.05]	[0.06]***	[0.06]***	[0.07]***	[0.09]**	[0.08]***	[0.06]**	[0.07]**	[0.07]
	{0.03}	{0.04}	{0.06}	{0.09}***	{0.10}***	{0.12}*	{0.14}	{0.14}**	{0.07}**	{0.07}**	{0.07}
Dist. River	−0.03	−0.04	−0.07	−0.21	−0.36	−0.50	−0.30	−0.39	−0.15	−0.18	−0.09
	(0.07)	(0.09)	(0.10)	(0.14)	(0.13)***	(0.15)***	(0.17)*	(0.17)**	(0.16)	(0.16)	(0.16)
	[0.07]	[0.09]	[0.10]	[0.14]	[0.13]***	[0.15]***	[0.18]	[0.18]**	[0.16]	[0.16]	[0.15]
	{0.06}	{0.08}	{0.08}	{0.16}	{0.16}**	{0.21}**	{0.23}	{0.22}*	{0.18}	{0.20}	{0.18}
Ruggedness	0.00	−0.05	−0.12	−0.32	−0.18	−0.49	−0.46	−0.39	−0.51	−0.47	−0.41
	(0.05)	(0.06)	(0.07)*	(0.10)***	(0.13)	(0.15)***	(0.17)***	(0.17)**	(0.16)***	(0.16)***	(0.16)***
	[0.05]	[0.06]	[0.07]*	[0.11]***	[0.12]	[0.16]***	[0.18]***	[0.17]**	[0.17]***	[0.17]***	[0.17]**
	{0.05}	{0.05}	{0.06}**	{0.11}***	{0.13}	{0.20}**	{0.23}**	{0.20}*	{0.22}**	{0.22}**	{0.21}**
Temperature	−0.03	0.04	0.16	−0.03	−0.04	−0.31	−0.17	−0.29	−0.06	−0.02	−0.08
	(0.04)	(0.05)	(0.05)***	(0.09)	(0.09)	(0.11)***	(0.12)	(0.12)**	(0.11)	(0.11)	(0.11)
	[0.04]	[0.05]	[0.06]***	[0.09]	[0.09]	[0.11]***	[0.13]	[0.12]**	[0.11]	[0.11]	[0.10]
	{0.03}	{0.05}	{0.08}**	{0.12}	{0.10}	{0.14}**	{0.15}	{0.13}**	{0.13}	{0.13}	{0.09}
Temperature^2	−0.01	−0.22	−0.35	−0.10	0.00	1.29	0.63	0.11	−0.06	−0.01	0.06
	(0.16)	(0.19)	(0.20)*	(0.31)	(0.32)	(0.47)***	(0.47)	(0.48)	(0.44)	(0.45)	(0.42)
	[0.15]	[0.19]	[0.20]*	[0.30]	[0.32]	[0.47]***	[0.49]	[0.49]	[0.46]	[0.45]	[0.41]
	{0.13}	{0.19}	{0.22}	{0.32}	{0.40}	{0.51}**	{0.62}	{0.61}	{0.52}	{0.54}	{0.49}
Precipitation	−0.08	−0.22	−0.22	−0.27	−0.15	−0.29	−0.25	−0.31	0.05	−0.15	−0.22
	(0.05)	(0.07)***	(0.09)**	(0.12)**	(0.15)	(0.18)*	(0.18)	(0.18)*	(0.16)	(0.17)	(0.16)
	[0.05]	[0.07]***	[0.09]**	[0.12]**	[0.14]	[0.18]	[0.18]	[0.18]*	[0.16]	[0.16]	[0.15]
	{0.05}	{0.07}***	{0.09}**	{0.16}*	{0.17}	{0.25}	{0.27}	{0.26}	{0.19}	{0.19}	{0.16}
Precipitation^2	0.13	0.41	0.39	0.49	0.23	0.46	0.35	0.51	−0.24	0.16	0.31
	(0.11)	(0.15)***	(0.20)*	(0.26)*	(0.30)	(0.35)	(0.36)	(0.36)	(0.33)	(0.34)	(0.32)
	[0.11]	[0.15]***	[0.20]**	[0.26]*	[0.28]	[0.36]	[0.37]	[0.37]	[0.34]	[0.34]	[0.31]
	{0.10}	{0.15}***	{0.19}**	{0.33}	{0.34}	{0.53}	{0.56}	{0.55}	{0.41}	{0.38}	{0.34}
Elevation	−0.37	−0.26	0.05	−0.97	−1.99	−1.83	−1.69	−2.67	−0.11	−0.10	−0.86
	(0.20)*	(0.25)	(0.26)	(0.44)**	(0.45)***	(0.57)***	(0.60)***	(0.63)***	(0.58)	(0.58)	(0.56)
	[0.20]*	[0.24]	[0.26]	[0.42]**	[0.44]***	[0.56]***	[0.63]***	[0.65]***	[0.58]	[0.58]	[0.53]
	{0.16}**	{0.20}	{0.29}	{0.46}**	{0.52}***	{0.62}***	{0.71}**	{0.73}***	{0.77}	{0.76}	{0.51}*
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.01	0.01	0.03	0.03	0.04	0.04	0.05	0.05	0.03	0.03	0.03
F Stat.	1.28	1.86	2.56	2.83	3.65	3.48	3.98	4.04	2.72	2.75	3.04
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the prefecture seats. The dependent variable is an indicator that equals one, if the pixel hosts a prefecture seat in that year, and zero otherwise. The estimations utilise the baseline sample. Conley standard errors using a 50 km radius are in parentheses, using a 100 km radius in brackets, and using a 500 km radius in curly braces ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

C.4 Robustness Check: Probit and Logit Regressions

In this section, we repeat the OLS linear probability regressions from Section 4.1 measuring direct effects of geography on city locations with logit and probit. The OLS, probit, and logit coefficient estimates do not allow for comparisons in terms of magnitude but in terms of their sign and statistical significance. And in that regard, the following tables are aligned with our baseline results. We see more significant estimates in county seat than in prefecture seat regressions, more pronounced than in the OLS estimations.[1] Distance from the equator, ruggedness, and elevation have negative effects. Distance from the coast reflects the westward expansion of the Tang dynasty discussed in Section 4.1.

[1] An exception are the probit regressions with Conley standard errors where the gap between county and prefecture seats in the number of significant coefficient estimates is smaller than in OLS regressions.

Table C-15: Local Geography County Seat Logit Regressions

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	-14.33	-11.61	-9.15	-10.82	-14.22	-15.60	-9.37	-8.66	-9.74	-8.91	-7.23
	(5.48)***	(4.38)***	(4.23)**	(3.61)***	(3.81)***	(3.28)***	(3.45)***	(3.40)**	(3.17)***	(2.99)***	(2.93)**
	[6.74]**	[5.78]**	[5.35]*	[5.07]**	[5.48]***	[5.07]***	[4.77]**	[4.45]*	[3.96]**	[3.63]**	[3.42]**
Dist. Coast	2.79	-3.34	1.07	5.12	14.63	11.52	11.28	8.71	0.69	2.55	0.42
	(2.51)	(2.05)	(2.03)	(2.13)**	(1.90)***	(1.75)***	(1.85)***	(1.80)***	(1.80)	(1.69)	(1.67)
	[4.07]	[3.24]	[2.93]	[3.30]	[3.16]***	[2.94]***	[3.13]***	[2.93]***	[2.74]	[2.52]	[2.30]
Dist. River	13.70	5.37	4.15	-4.41	-13.16	1.36	6.91	7.22	-0.87	-3.18	-0.60
	(5.62)**	(5.12)	(4.92)	(5.16)	(4.81)***	(4.17)	(4.15)*	(4.16)*	(4.27)	(4.05)	(4.10)
	[8.34]	[7.56]	[6.72]	[7.40]	[7.05]*	[7.07]	[7.03]	[6.84]	[6.63]	[6.24]	[5.43]
Ruggedness	-16.36	-14.84	-20.67	-29.41	-0.53	-12.92	-16.53	-18.16	-20.92	-18.77	-20.27
	(9.19)*	(8.20)*	(8.00)***	(6.63)***	(6.88)	(5.64)**	(6.06)***	(5.94)***	(5.74)***	(5.15)***	(4.91)***
	[11.07]	[10.20]	[10.13]**	[8.28]***	[9.25]	[7.76]*	[8.33]**	[7.86]**	[7.73]***	[6.89]***	[6.21]***
Temperature	65.97	70.35	70.92	74.96	42.86	43.01	55.16	55.02	60.84	57.62	46.92
	(11.00)***	(10.44)***	(10.68)***	(10.41)***	(7.12)***	(6.61)***	(7.28)***	(6.97)***	(7.09)***	(6.69)***	(5.50)***
	[24.42]***	[21.10]***	[20.18]***	[26.60]***	[15.18]***	[16.47]***	[17.79]***	[17.49]***	[19.04]***	[17.45]***	[11.56]***
Temperature^2	-266.92	-237.78	-248.93	-278.14	-129.36	-137.77	-175.67	-181.42	-209.56	-200.36	-165.66
	(39.15)***	(33.30)***	(34.19)***	(33.61)***	(22.49)***	(20.53)***	(22.39)***	(21.56)***	(22.34)***	(20.95)***	(17.67)***
	[89.94]***	[70.86]***	[67.14]***	[88.57]***	[49.92]***	[53.56]**	[57.85]***	[58.14]***	[63.46]***	[58.41]***	[38.81]***
Precipitation	-34.33	-32.53	-16.20	-12.92	-37.56	-19.98	-9.55	-5.08	-4.60	-0.85	1.31
	(7.35)***	(7.43)***	(6.96)**	(4.83)***	(5.06)***	(4.47)***	(4.58)**	(4.85)	(4.84)	(4.55)	(4.54)
	[10.38]***	[11.88]***	[9.72]*	[7.66]*	[7.47]***	[7.73]***	[7.70]	[7.82]	[6.95]	[6.71]	[6.12]
Precipitation^2	58.32	41.93	16.52	27.88	67.48	34.88	16.19	4.16	0.59	-7.84	-11.49
	(17.34)***	(26.00)	(22.37)	(10.06)***	(10.04)***	(9.37)***	(10.25)	(12.12)	(11.77)	(11.12)	(11.41)
	[27.29]**	[45.18]	[34.08]	[17.21]	[16.32]***	[17.20]**	[18.87]	[21.62]	[19.41]	[18.84]	[17.15]
Elevation	-168.83	-56.41	-50.90	-79.20	-174.03	-133.73	-116.93	-89.51	-49.98	-53.36	-41.07
	(33.50)***	(24.70)**	(24.39)**	(21.48)***	(22.14)***	(18.92)***	(20.43)***	(19.18)***	(16.90)***	(15.89)***	(15.20)***
	[50.27]***	[30.91]*	[29.64]*	[29.29]***	[32.42]***	[27.35]***	[29.20]***	[27.18]***	[23.14]**	[21.90]**	[19.72]**
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the county seats. The input pixels are the same as in Table 2. The dependent variable is an indicator that equals one, if the pixel hosts a county seat, and zero otherwise. Heteroskedasticity-robust standard errors are in parentheses and Conley standard errors with a radius of 150 km (great-circle distances computed via the haversine formula (Sinnott, 1984)) and a Bartlett kernel are in brackets ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index $\times$ 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-16: Local Geography Prefecture Seat Logit Regressions

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	-30.12	-22.91	6.15	-22.31	-20.82	-24.31	-19.53	-26.88	-19.84	-22.15	-16.80
	(19.15)	(13.16)*	(12.49)	(8.13)***	(6.77)***	(5.98)***	(5.86)***	(5.79)***	(6.49)***	(6.02)***	(6.14)***
	[20.81]	[13.12]*	[11.65]	[8.69]**	[6.82]***	[7.01]***	[6.52]***	[6.13]***	[7.47]***	[7.03]***	[5.62]***
Dist. Coast	5.85	3.12	6.59	12.98	17.16	8.98	6.13	9.75	7.10	7.57	3.08
	(8.28)	(6.21)	(5.50)	(3.73)***	(3.25)***	(3.18)***	(2.85)**	(2.75)***	(3.51)**	(3.20)**	(2.92)
	[8.10]	[5.85]	[5.48]	[3.81]***	[3.19]***	[3.20]***	[3.23]*	[3.05]***	[3.73]*	[3.47]**	[2.72]
Dist. River	4.39	2.73	-3.01	-6.14	-4.91	-9.92	-0.69	-3.55	6.40	3.77	3.80
	(20.64)	(14.07)	(12.39)	(10.94)	(8.71)	(7.65)	(6.91)	(7.05)	(8.61)	(8.37)	(8.00)
	[19.86]	[14.38]	[12.37]	[10.55]	[7.63]	[7.56]	[7.82]	[7.79]	[9.33]	[8.68]	[7.23]
Ruggedness	3.76	-18.58	-27.82	-29.96	-7.57	-25.35	-20.22	-14.49	-36.95	-32.27	-25.00
	(36.84)	(23.67)	(19.40)	(11.89)**	(10.88)	(9.46)***	(9.40)**	(8.79)*	(9.80)***	(9.88)***	(11.20)**
	[37.88]	[24.87]	[20.35]	[12.80]**	[10.79]	[10.59]**	[10.85]*	[10.04]	[10.97]***	[10.78]***	[12.11]**
Temperature	-15.99	36.57	88.54	11.85	26.19	26.39	25.32	14.74	34.60	40.84	17.40
	(19.23)	(21.08)*	(32.57)***	(12.98)	(12.69)**	(9.82)***	(9.54)***	(8.88)*	(12.17)***	(10.32)***	(8.16)**
	[22.90]	[34.12]	[55.41]	[15.49]	[16.30]	[20.10]	[16.07]	[13.08]	[18.38]*	[19.52]**	[10.38]*
Temperature^2	-8.24	-171.40	-287.38	-71.36	-121.79	-102.89	-108.89	-103.90	-138.30	-154.70	-84.27
	(71.89)	(83.65)**	(135.01)**	(39.02)*	(41.63)***	(30.64)***	(29.34)***	(27.89)***	(37.14)***	(32.64)***	(25.13)***
	[76.57]	[126.01]	[214.47]	[50.00]	[58.71]**	[63.69]	[51.41]**	[44.16]**	[58.37]**	[62.81]**	[35.23]**
Precipitation	-4.58	-24.67	13.84	-20.63	-21.30	-21.47	-15.63	-20.49	11.92	-8.80	-7.33
	(32.59)	(25.15)	(25.80)	(13.58)	(8.96)**	(8.37)**	(8.57)*	(8.03)**	(13.31)	(10.76)	(10.20)
	[35.89]	[27.57]	[25.03]	[16.15]	[9.07]**	[9.60]**	[10.29]	[10.12]**	[16.71]	[12.37]	[10.92]
Precipitation^2	-123.08	-8.99	-132.95	19.51	36.19	35.71	18.91	30.23	-54.98	6.75	-5.17
	(131.10)	(106.88)	(110.19)	(52.22)	(23.03)	(21.12)*	(23.92)	(21.40)	(42.42)	(29.73)	(31.45)
	[151.52]	[127.90]	[121.15]	[65.94]	[25.11]	[25.96]	[32.11]	[30.72]	[58.12]	[37.40]	[38.58]
Elevation	-133.20	-19.05	90.21	-61.37	-119.37	-68.47	-56.81	-95.81	42.41	30.21	-19.89
	(110.49)	(66.52)	(56.36)	(39.66)	(37.17)***	(30.31)**	(29.15)*	(29.80)***	(30.09)	(29.24)	(31.35)
	[124.46]	[68.88]	[59.05]	[38.69]	[38.45]***	[32.84]**	[33.50]*	[32.19]***	[31.46]	[30.38]	[28.37]
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the prefecture seats. The input pixels are the same as in Table 3. The dependent variable is an indicator that equals one, if the pixel hosts a prefecture seat, and zero otherwise. Heteroskedasticity-robust standard errors are in parentheses and Conley standard errors with a radius of 150 km (great-circle distances computed via the haversine formula (Sinnott, 1984)) and a Bartlett kernel are in brackets ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-17: Local Geography County Seat Probit Regressions

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	-6.16	-7.22	-5.33	-5.77	-8.36	-9.14	-5.59	-5.44	-6.05	-5.77	-4.72
	(2.73)**	(2.41)***	(2.32)**	(1.91)***	(2.06)***	(1.81)***	(1.93)***	(1.90)***	(1.75)***	(1.68)***	(1.62)***
	[1.75]***	[1.53]***	[1.42]***	[1.42]***	[1.55]***	[1.53]***	[1.42]***	[1.34]***	[1.19]***	[1.12]***	[1.05]***
Dist. Coast	0.69	-1.27	1.02	3.22	8.23	6.65	6.59	5.26	0.75	1.90	0.67
	(1.34)	(1.16)	(1.15)	(1.16)***	(1.06)***	(0.98)***	(1.05)***	(1.01)***	(1.02)	(0.97)**	(0.96)
	[1.07]	[0.94]	[0.84]	[0.98]***	[0.93]***	[0.93]***	[0.97]***	[0.93]***	[0.85]	[0.80]**	[0.70]
Dist. River	7.34	3.86	3.12	-1.16	-5.72	0.85	4.01	4.49	-0.41	-1.58	0.06
	(2.96)**	(2.81)	(2.69)	(2.69)	(2.67)**	(2.33)	(2.32)*	(2.36)*	(2.35)	(2.26)	(2.36)
	[2.39]***	[2.28]*	[2.02]	[2.07]	[2.12]***	[2.22]	[2.22]*	[2.21]**	[2.08]	[1.99]	[1.73]
Ruggedness	-7.85	-4.33	-7.23	-15.72	0.60	-5.54	-7.59	-8.47	-10.54	-9.49	-10.93
	(4.53)*	(4.73)	(4.62)	(3.38)***	(3.86)	(3.24)*	(3.55)**	(3.45)**	(3.32)***	(2.98)***	(2.67)***
	[2.59]***	[2.44]*	[2.42]***	[2.12]***	[2.45]	[2.24]**	[2.37]***	[2.25]***	[2.18]***	[2.01]***	[1.81]***
Temperature	33.10	30.65	31.13	34.09	20.13	20.89	26.34	26.87	29.98	28.12	22.30
	(5.22)***	(6.09)***	(6.10)***	(5.60)***	(3.87)***	(3.91)***	(4.32)***	(3.99)***	(4.07)***	(3.91)***	(3.07)***
	[5.47]***	[5.00]***	[4.56]***	[5.98]***	[3.76]***	[4.41]***	[4.48]***	[4.60]***	[4.97]***	[4.56]***	[2.99]***
Temperature^2	-128.41	-107.23	-112.68	-128.81	-62.76	-70.30	-86.05	-91.59	-106.91	-102.34	-82.38
	(17.96)***	(17.81)***	(18.13)***	(17.56)***	(11.80)***	(11.92)***	(12.91)***	(12.08)***	(12.58)***	(11.96)***	(9.76)***
	[19.80]***	[16.61]***	[15.01]***	[20.06]***	[12.46]***	[14.72]***	[15.00]***	[15.74]***	[17.01]***	[15.74]***	[10.33]***
Precipitation	-16.86	-15.95	-6.44	-6.62	-19.06	-9.30	-3.98	-1.64	-1.97	0.14	1.47
	(4.01)***	(4.39)***	(4.10)	(2.54)***	(2.79)***	(2.50)***	(2.59)	(2.85)	(2.70)	(2.57)	(2.64)
	[3.15]***	[3.79]***	[3.14]**	[2.17]***	[2.23]***	[2.39]***	[2.43]	[2.62]	[2.18]	[2.16]	[1.99]
Precipitation^2	26.40	13.19	-1.53	13.51	33.04	14.57	4.76	-2.20	-2.13	-7.01	-9.75
	(11.04)**	(16.78)	(14.62)	(5.26)**	(5.58)***	(5.33)***	(6.04)	(7.76)	(6.82)	(6.51)	(7.05)
	[9.17]***	[14.72]	[11.60]	[4.65]***	[4.74]***	[5.25]***	[6.05]	[7.68]	[6.15]	[6.11]	[5.76]*
Elevation	-78.10	-36.16	-30.83	-41.80	-91.44	-73.84	-66.82	-53.53	-30.69	-33.05	-25.07
	(17.05)***	(13.53)***	(13.23)**	(11.15)***	(12.12)***	(10.64)***	(11.33)***	(10.58)***	(9.35)***	(8.95)***	(8.18)***
	[12.21]***	[7.85]***	[7.41]***	[7.89]***	[8.27]***	[7.60]***	[7.98]***	[7.50]***	[6.46]***	[6.31]***	[5.67]***
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the county seats. The input pixels are the same as in Table 2. The dependent variable is an indicator that equals one, if the pixel hosts a county seat, and zero otherwise. Heteroskedasticity-robust standard errors are in parentheses and Conley standard errors with a radius of 150 km (great-circle distances computed via the haversine formula (Sinnott, 1984)) and a Bartlett kernel are in brackets ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index $\times$ 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-18: Local Geography Prefecture Seat Probit Regressions

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	-12.31	-10.33	1.55	-9.83	-10.37	-11.08	-9.35	-12.95	-9.19	-10.42	-7.74
	(8.08)	(5.15)**	(4.96)	(3.61)***	(3.05)***	(2.80)***	(2.82)***	(2.78)***	(2.89)***	(2.67)***	(2.70)***
	[3.73]***	[2.26]***	[2.10]	[1.69]***	[1.40]***	[1.52]***	[1.44]***	[1.36]***	[1.53]***	[1.47]***	[1.14]***
Dist. Coast	2.25	1.30	2.53	5.98	8.33	3.91	2.88	4.97	3.30	3.57	1.66
	(3.43)	(2.76)	(2.56)	(1.68)***	(1.50)***	(1.54)**	(1.42)**	(1.32)***	(1.57)**	(1.44)**	(1.31)
	[1.37]	[1.08]	[1.08]**	[0.80]***	[0.69]***	[0.71]***	[0.76]***	[0.71]***	[0.77]***	[0.73]***	[0.57]***
Dist. River	2.04	1.17	-1.63	-2.61	-3.05	-4.28	-0.25	-1.74	3.73	2.58	2.03
	(8.85)	(6.14)	(5.60)	(5.03)	(4.06)	(3.53)	(3.34)	(3.42)	(3.95)	(3.90)	(3.71)
	[3.51]	[2.73]	[2.54]	[2.15]	[1.66]*	[1.67]**	[1.85]	[1.83]	[1.94]*	[1.86]	[1.59]
Ruggedness	5.92	-7.73	-11.43	-13.42	-4.36	-10.77	-9.04	-6.77	-16.60	-14.06	-9.75
	(15.80)	(10.03)	(7.82)	(4.96)***	(4.67)	(4.51)**	(4.58)**	(4.16)	(4.32)***	(4.42)***	(4.98)**
	[6.42]	[3.94]**	[3.37]***	[2.35]***	[2.10]**	[2.22]***	[2.33]***	[2.13]***	[2.16]***	[2.14]***	[2.26]***
Temperature	-6.97	14.28	33.77	5.06	9.80	12.09	11.42	6.89	14.76	17.98	6.37
	(8.40)	(7.88)*	(11.94)***	(5.50)	(5.44)*	(4.47)***	(4.56)**	(4.18)*	(5.36)***	(4.44)***	(3.50)*
	[4.30]	[5.34]***	[8.61]***	[2.77]*	[2.79]***	[4.01]***	[3.45]***	[2.75]**	[3.25]***	[3.65]***	[1.79]***
Temperature^2	1.88	-70.28	-108.50	-29.88	-50.37	-48.55	-51.35	-50.16	-60.63	-69.55	-33.71
	(29.70)	(29.75)**	(47.81)**	(16.20)*	(17.72)***	(14.25)***	(13.90)***	(13.07)***	(16.59)***	(14.55)***	(10.88)***
	[14.38]	[19.78]***	[32.83]***	[8.94]***	[10.32]***	[13.11]***	[11.38]***	[9.55]***	[10.79]***	[12.13]***	[6.37]***
Precipitation	0.63	-11.59	5.48	-9.72	-9.51	-8.88	-6.21	-9.56	5.87	-4.35	-2.39
	(13.73)	(10.68)	(10.58)	(5.97)	(3.95)**	(4.08)**	(4.49)	(3.89)**	(5.82)	(4.93)	(4.65)
	[6.37]	[4.90]**	[4.53]	[3.22]***	[1.84]***	[2.13]***	[2.60]**	[2.34]***	[3.34]*	[2.67]	[2.40]
Precipitation^2	-68.60	0.97	-57.75	11.37	16.16	13.58	4.38	13.62	-28.14	2.71	-6.95
	(55.62)	(45.89)	(44.76)	(22.09)	(9.59)*	(10.94)	(14.01)	(10.55)	(18.68)	(14.09)	(15.10)
	[27.98]**	[21.71]	[20.89]***	[12.66]	[4.71]***	[5.82]**	[8.78]	[7.15]*	[11.68]**	[8.40]	[8.78]
Elevation	-54.07	-14.32	35.10	-27.41	-57.81	-32.27	-28.89	-47.87	16.48	10.61	-13.42
	(45.16)	(25.71)	(22.69)	(17.16)	(16.04)***	(14.01)**	(14.11)**	(14.31)***	(13.69)	(13.16)	(13.97)
	[21.74]**	[11.37]	[10.61]***	[7.39]***	[7.40]***	[7.03]***	[7.20]***	[6.93]***	[6.46]**	[6.33]*	[5.58]**
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the prefecture seats. The input pixels are the same as in Table 3. The dependent variable is an indicator that equals one, if the pixel hosts a prefecture seat, and zero otherwise. Heteroskedasticity-robust standard errors are in parentheses and Conley standard errors with a radius of 150 km (great-circle distances computed via the haversine formula (Sinnott, 1984)) and a Bartlett kernel are in brackets ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

C.5 Robustness Check: Spatial Durbin Model

Our baseline methods in Section 4.1 only account for spatial correlation in their standard errors. Otherwise, they estimate the effect of geography on city locations non-spatially. In that, they do not account for the effect a city or the geographic conditions in surrounding pixels might have on a grid cell.

Our application does not provide a strong motivation for the use of spatial methods. (i) Given that our pixels are not that large, their geography is similar to that of neighbouring pixels. A spatially weighted environment, therefore, barely adds any information. (ii) In terms of the dependent variable, we observe mostly isolated urban pixels in a vast sea of rural pixels. Spatial spillovers in the urban indicator are much less visible than they are in continuous data such as population size. (iii) The three distance variables already account for important determinants along the spatial dimension. (iv) Our analysis of indirect geography effects on city locations in Section 4.2 focuses on the spatial dimension, but does so in a more structural manner shaped by the theoretical framework and the historical context. To illustrate our results' robustness, this section, nonetheless, tests the mechanisms with spatial econometrics.[2]

We estimate a Spatial Durbin Model, i.e. a model of the form

$$y = \rho W y + X\beta + WX\theta + \varepsilon \tag{(C-1)}$$

where y is the urban indicator and X the geography vector. It is estimated with maximum likelihood and accounts for global spatial spillovers. Observations refer to the baseline pixels. Table 2 and Table 3 in Section 4.1 display the respective results when using a linear probability model.

The estimates in Table C-19 and Table C-20 support the baseline results. We see more significant coefficient estimates in county seat than in prefecture seat regressions. And cities tend to locate in lower, less rugged terrain, and near rivers. Interesting are the lagged variables' reversed signs, e.g. in the case of elevation and ruggedness. When a county seat is set up in a region, it is placed in the locally optimal location, making the surrounding non-selected places appear comparatively worse.

[2]There is currently no counterpart to these spatial econometric methods on the machine learning side. Geographic random forests are a recent innovation that run a set of regional estimations instead of drawing random samples from the full geographic space (Georganos et al., 2021). They can identify regional heterogeneity, but are not spatial in the spatial econometrics sense - they do not address spatial spillovers. And so far, they target continuous outcomes and are not suited to binary classification problems. Given the absence of structural estimation equations and other characteristics in random forests, it is questionable whether it would be possible and reasonable to include different types of spillovers in these methods at all.

Table C-19: Local Geography County Seat Spatial Regressions

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Intercept	1.23	1.87	2.49	1.60	2.34	0.76	3.34*	1.63	−0.24	−0.22	2.11
	(1.45)	(1.67)	(1.70)	(1.77)	(1.82)	(1.99)	(1.97)	(2.00)	(2.02)	(2.12)	(2.16)
Dist. Equator	−4.14	−6.93	−8.93	−5.87	−7.09	−3.67	−9.96	−4.35	0.20	0.05	−4.76
	(5.31)	(6.09)	(6.21)	(6.47)	(6.64)	(7.27)	(7.18)	(7.28)	(7.38)	(7.74)	(7.88)
Dist. Coast	3.49	4.80	8.01	8.76*	11.06**	8.37	7.30	8.49	11.04*	12.61**	8.36
	(4.35)	(4.98)	(5.09)	(5.30)	(5.44)	(5.96)	(5.88)	(5.97)	(6.05)	(6.34)	(6.46)
Dist. River	−8.24**	−6.43*	−7.63**	−6.95*	−11.64***	−3.23	−7.08	−8.81*	−4.13	−1.19	−6.27
	(3.32)	(3.81)	(3.89)	(4.05)	(4.16)	(4.55)	(4.49)	(4.56)	(4.62)	(4.84)	(4.93)
Ruggedness	−1.57***	−2.19***	−2.66***	−3.37***	−3.14***	−4.12***	−4.86***	−4.95***	−4.87***	−5.21***	−5.38***
	(0.44)	(0.51)	(0.52)	(0.54)	(0.55)	(0.60)	(0.60)	(0.61)	(0.61)	(0.64)	(0.65)
Temperature	−0.58	−0.45	−0.55	−0.45	−1.12	−1.43*	−1.91**	−2.20***	−1.57*	−2.11**	−1.87**
	(0.58)	(0.67)	(0.68)	(0.71)	(0.73)	(0.80)	(0.79)	(0.80)	(0.81)	(0.85)	(0.86)
Temperature^2	4.90*	5.40*	7.58**	8.57***	6.25*	9.28**	7.23**	10.54***	11.83***	12.28***	13.68***
	(2.73)	(3.13)	(3.19)	(3.32)	(3.41)	(3.74)	(3.69)	(3.74)	(3.79)	(3.98)	(4.05)
Precipitation	0.96	0.18	−1.06	−1.19	−0.15	0.68	−4.25**	−2.40	−0.78	−0.87	−0.04
	(1.53)	(1.76)	(1.80)	(1.87)	(1.92)	(2.10)	(2.08)	(2.11)	(2.13)	(2.24)	(2.28)
Precipitation^2	0.63	2.90	5.21	5.76	3.66	4.96	13.28***	10.75**	7.82*	8.53*	6.58
	(3.21)	(3.68)	(3.75)	(3.91)	(4.01)	(4.39)	(4.34)	(4.40)	(4.46)	(4.68)	(4.76)
Elevation	−7.33***	−9.91***	−10.75***	−11.39***	−13.13***	−15.42***	−16.25***	−16.89***	−14.09***	−17.07***	−18.08***
	(1.85)	(2.12)	(2.16)	(2.25)	(2.31)	(2.53)	(2.50)	(2.54)	(2.57)	(2.69)	(2.74)
lag. Dist. E.	3.85	6.45	8.46	5.19	6.23	2.58	9.52	3.88	−0.92	−0.75	4.18
	(5.34)	(6.12)	(6.24)	(6.50)	(6.68)	(7.31)	(7.22)	(7.32)	(7.42)	(7.78)	(7.92)
lag. Dist. C.	−3.45	−4.79	−7.78	−8.31	−10.35*	−7.72	−6.57	−7.76	−10.79*	−12.16*	−8.01
	(4.36)	(5.00)	(5.11)	(5.32)	(5.46)	(5.98)	(5.90)	(5.99)	(6.07)	(6.36)	(6.48)
lag. Dist. R.	8.69**	6.46*	7.59*	6.52	10.83**	2.84	7.27	9.21**	4.02	0.75	5.95
	(3.42)	(3.92)	(4.00)	(4.17)	(4.28)	(4.68)	(4.62)	(4.69)	(4.75)	(4.98)	(5.08)
lag. Rugg.	1.60***	2.45***	3.05***	3.04***	3.38***	3.82***	4.76***	4.89***	4.70***	5.21***	5.30***
	(0.61)	(0.70)	(0.71)	(0.74)	(0.76)	(0.83)	(0.82)	(0.84)	(0.85)	(0.89)	(0.90)
lag. Temp.	0.68	0.38	0.49	0.43	0.86	0.66	1.62*	2.00**	1.16	1.91*	1.83*
	(0.68)	(0.78)	(0.79)	(0.82)	(0.85)	(0.93)	(0.91)	(0.93)	(0.94)	(0.99)	(1.00)
lag. Temp.^{2}	−3.84	−2.75	−5.86*	−9.67***	−3.43	−4.95	−3.03	−7.64*	−9.25**	−9.96**	−11.56***
	(3.00)	(3.44)	(3.52)	(3.66)	(3.76)	(4.12)	(4.06)	(4.12)	(4.18)	(4.38)	(4.46)
lag. Prec.	−1.68	−1.06	0.84	1.30	−0.50	−0.52	5.40**	3.51	1.15	1.88	1.05
	(1.68)	(1.93)	(1.97)	(2.05)	(2.10)	(2.30)	(2.27)	(2.30)	(2.34)	(2.45)	(2.49)
lag. Prec.^{2}	0.47	−1.67	−5.29	−6.30	−2.66	−5.88	−16.38***	−13.83***	−9.82**	−12.12**	−10.06*
	(3.58)	(4.10)	(4.18)	(4.36)	(4.47)	(4.90)	(4.84)	(4.91)	(4.97)	(5.21)	(5.31)
lag. Elev.	6.06***	8.02***	8.31***	5.85**	6.85**	7.58**	10.24***	11.12***	8.96***	11.67***	13.15***
	(2.25)	(2.58)	(2.64)	(2.75)	(2.82)	(3.09)	(3.05)	(3.09)	(3.13)	(3.28)	(3.34)
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. ((C-1)) using the county seats. The dependent variables is an indicator that equals one, if the pixel hosts a county seat, and zero otherwise. The input pixels are the same as in Table 2. Heteroskedasticity-robust standard errors are in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Coefficient estimates on categorical soil variables - dominant soil type, landform, lithology - are omitted from the table. Spatial weights are based on queen contiguity.

Table C-20: Local Geography Prefecture Seat Spatial Regressions

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Intercept	0.33	0.32	0.03	1.17	0.99	−0.32	0.53	0.33	0.74	0.36	0.33
	(0.46)	(0.62)	(0.68)	(0.95)	(1.05)	(1.21)	(1.27)	(1.28)	(1.05)	(1.08)	(1.13)
Dist. Equator	−0.89	−0.91	−0.29	−3.56	−3.01	1.77	−0.88	−0.03	−2.26	−1.04	−0.06
	(1.68)	(2.26)	(2.49)	(3.45)	(3.83)	(4.41)	(4.64)	(4.67)	(3.82)	(3.93)	(4.11)
Dist. Coast	−0.43	−1.23	0.39	2.62	3.05	4.55	5.47	3.24	0.31	0.10	2.30
	(1.38)	(1.85)	(2.04)	(2.83)	(3.14)	(3.62)	(3.80)	(3.82)	(3.13)	(3.22)	(3.37)
Dist. River	−0.84	−2.41*	−2.72*	−5.78***	−5.61**	−4.05	−3.96	−5.47*	−2.70	−4.01	−5.06**
	(1.05)	(1.41)	(1.56)	(2.16)	(2.40)	(2.76)	(2.90)	(2.92)	(2.39)	(2.46)	(2.57)
Ruggedness	−0.13	−0.39**	−0.25	−1.08***	−1.13***	−1.92***	−2.18***	−2.06***	−1.53***	−1.63***	−1.88***
	(0.14)	(0.19)	(0.21)	(0.29)	(0.32)	(0.37)	(0.39)	(0.39)	(0.32)	(0.33)	(0.34)
Temperature	−0.07	0.15	0.23	0.17	−0.45	−0.57	−0.72	−0.45	−0.10	−0.04	−0.36
	(0.18)	(0.25)	(0.27)	(0.38)	(0.42)	(0.48)	(0.51)	(0.51)	(0.42)	(0.43)	(0.45)
Temperature^2	0.54	−0.35	0.18	2.13	1.92	3.19	2.55	1.98	−0.48	0.38	1.84
	(0.86)	(1.16)	(1.28)	(1.77)	(1.97)	(2.27)	(2.38)	(2.40)	(1.96)	(2.02)	(2.11)
Precipitation	−0.70	0.24	1.17	−1.04	0.20	−0.16	−1.24	−1.00	−0.05	0.65	−1.04
	(0.49)	(0.65)	(0.72)	(1.00)	(1.11)	(1.28)	(1.34)	(1.35)	(1.10)	(1.14)	(1.19)
Precipitation^2	1.62	0.18	−1.39	2.79	1.06	2.56	4.26	3.72	1.71	0.79	2.95
	(1.01)	(1.36)	(1.51)	(2.08)	(2.31)	(2.67)	(2.80)	(2.82)	(2.31)	(2.38)	(2.48)
Elevation	−0.65	−1.26	−1.82**	−3.69***	−5.71***	−6.51***	−7.29***	−8.17***	−3.40**	−3.84***	−5.59***
	(0.58)	(0.79)	(0.87)	(1.20)	(1.33)	(1.54)	(1.61)	(1.62)	(1.33)	(1.37)	(1.43)
lag. Dist. E.	0.76	0.72	0.18	3.15	2.59	−2.26	0.55	−0.59	2.21	0.90	−0.06
	(1.69)	(2.27)	(2.51)	(3.47)	(3.85)	(4.44)	(4.66)	(4.69)	(3.83)	(3.95)	(4.13)
lag. Dist. C.	0.45	1.30	−0.26	−2.36	−2.53	−4.39	−5.31	−2.93	−0.19	−0.01	−2.31
	(1.38)	(1.86)	(2.05)	(2.84)	(3.15)	(3.63)	(3.81)	(3.84)	(3.14)	(3.23)	(3.38)
lag. Dist. R.	0.81	2.37	2.71*	5.74***	5.52**	3.68	3.98	5.38*	2.67	4.05	5.16*
	(1.08)	(1.45)	(1.61)	(2.22)	(2.47)	(2.84)	(2.99)	(3.01)	(2.46)	(2.53)	(2.65)
lag. Rugg.	0.09	0.29	0.02	0.81**	0.98**	1.64***	2.13***	2.09***	0.94**	1.15**	1.99***
	(0.19)	(0.26)	(0.29)	(0.40)	(0.44)	(0.51)	(0.53)	(0.54)	(0.44)	(0.45)	(0.47)
lag. Temp.	−0.04	−0.26	−0.23	−0.39	0.41	0.21	0.54	0.05	0.18	0.09	0.31
	(0.21)	(0.29)	(0.32)	(0.44)	(0.49)	(0.56)	(0.59)	(0.59)	(0.49)	(0.50)	(0.52)
lag. Temp.^{2}	−0.16	0.83	0.15	−2.33	−1.70	−1.61	−1.76	−1.63	0.92	0.31	−0.98
	(0.95)	(1.28)	(1.41)	(1.95)	(2.17)	(2.50)	(2.62)	(2.64)	(2.16)	(2.23)	(2.33)
lag. Prec.	0.59	−0.54	−1.36*	0.70	−0.34	−0.27	0.98	0.69	0.04	−0.96	0.85
	(0.53)	(0.71)	(0.79)	(1.09)	(1.21)	(1.40)	(1.47)	(1.48)	(1.21)	(1.24)	(1.30)
lag. Prec.^{2}	−1.49	0.29	1.67	−2.25	−0.93	−1.93	−4.04	−3.33	−2.10	−0.62	−2.91
	(1.13)	(1.52)	(1.68)	(2.32)	(2.58)	(2.97)	(3.12)	(3.14)	(2.57)	(2.65)	(2.77)
lag. Elev.	0.09	0.60	1.86*	2.29	3.06*	5.03***	6.27***	5.58***	4.37***	4.32***	5.10***
	(0.71)	(0.96)	(1.06)	(1.46)	(1.62)	(1.87)	(1.97)	(1.98)	(1.62)	(1.67)	(1.74)
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. ((C-1)) using the prefecture seats. The dependent variables is an indicator that equals one, if the pixel hosts a prefecture seat, and zero otherwise. The input pixels are the same as in Table 3. Heteroskedasticity-robust standard errors are in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Coefficient estimates on categorical soil variables - dominant soil type, landform, lithology - are omitted from the table. Spatial weights are based on queen contiguity.

C.6 Supplementary Results on Goodness of Fit Measures

A key insight of Section 4.1 is that county seat locations are more strongly linked to local geography than prefecture seat locations are. We primarily draw that conclusion from OLS regression's adjusted R^2 and F statistics and from random forests' share of correctly classified pixels and R^2. Prefecture seat estimations exhibit a markedly lower goodness of fit than county seat estimations do. Table 2 and Table 3 show this for a few selected cross-sections. Introducing the MAUP's effect, Figure 5 further elaborates on that pattern with classification random forests, Figure C-4 with regression random forests, and Figure C-1 with OLS.

A drawback of these county-prefecture seat comparisons is that these measures are point estimates without a distribution. We can intuitively judge the absolute difference between them but lack a test of statistical significance. To evaluate this discrepancy between the types of administrative cities nonetheless, we borrow a trick that is widely used throughout econometrics: creating distributions via bootstrapping. I.e. we draw samples from the original cross-sectional data sets in Section 4.1. Allowing for replacement, the bootstrapped samples are as large as their source data. Per cross-section, we bootstrap 1,000 samples from the county seat data and 1,000 samples from the prefecture seat data. Repeating the baseline regressions, as specified in eq. (14) and employed in Table 2 and Table 3, for each of these assembled data sets produces 2,000 F statistics and adjusted R^2 for every cross-section.

In exploring the statistical significance of county-prefecture seat disparities, the resulting distributions serve as the dependent variable in following simple OLS regressions

$$F_{cst} = \alpha_t + \beta_t Z_c + \varepsilon_{cst} \tag{(C-2)}$$

where F is the F statistic or the adjusted R^2 from a regression on city type c in year t, using bootstrapped sample s. Z is a binary variable that equals one, if the value is derived from a prefecture seat estimation and zero in case of a county seat estimation.

Table C-21 and Table C-22 document a highly significant difference between county and prefecture seats. The explanatory power of local geography is robustly higher for the lower ranking administrative settlements than it is for the higher ranking ones.

The picture also holds with alternative pixel sizes - as Table C-23 and Table C-24 document - and is robust to the Modifiable Areal Unit Problem.

Showing that the estimated difference between county and prefecture seats is neither simply induced by a relatively higher frequency of ones in county seat regressions' urban indicator variable nor limited to the eleven baseline cross-sections, we test

$$F_{ct} = \beta Z_c + \gamma N^u_{ct} + \delta_t + \varepsilon_{ct} \tag{(C-3)}$$

regressing the adjusted R^2 and F statistic from the 214 cross-sections, used e.g. in Figure 4 on the prefecture seat regression indicator Z_c, the number of urban pixels in that regression N^u, and year fixed effects δ_t. As illustrated in Table C-25, the prefecture seat indicator's coefficient estimate remains highly statistically significant and of a magnitude similar to our other results.

Related to the R^2, we can show how markedly the predicted probabilities, i.e. the linear probability model's fitted values, reflect the difference between county and prefecture seat estimations. For that matter, we extract the mean estimated probability of hosting a city conditional on the pixel actually being urban $Pr(Y = 1|Y = 1)$ and regress it on the prefecture seat regression indicator Z_c, the number of urban pixels in that regression N^u, the mean estimated probability of hosting a city conditional on the pixel being rural $Pr(Y = 1|Y = 0)$, and year fixed effects δ_t. The rationale behind the inclusion of $Pr(Y = 1|Y = 0)$ is to control for the possibility that the model could obtain a higher $Pr(Y = 1|Y = 1)$ by raising $Pr(Y = 1)$ in general.

$$Pr(Y = 1|Y = 1)_{ct} = \beta Z_c + \gamma N^u_{ct} + \phi Pr(Y = 1|Y = 0)_{ct} + \delta_t + \varepsilon_{ct} \tag{(C-4)}$$

The results in Table C-26 confirm what we find throughout the paper: local geography predicts county seat locations significantly better than prefecture seat locations.

Finally, we further explore the role of prefecture fixed effects in the local geography prefecture seat regressions. We extract the adjusted R^2 and F statistic from 57 cross-sectional regressions, with one cross-section every ten years between 1350 and 1910 CE, as in our baseline model. We, then, repeat the estimations with prefecture fixed effects and extract the overall adjusted R^2, within adjusted R^2 (referring to prefecture fixed effects, not soil categorical soil variables), and F statistic. Estimations start in 1350 CE because for some missing prefecture borders before that point. As illustrated in Figure C-2a, the R^2 is quasi zero in both specifications. With the red dashed line narrowly fluctuating around the blue solid line, geographic variation within prefectures explains prefecture seat locations about as well as geographic variation throughout the empire does. The fixed effects regressions' overall adjusted R^2, i.e. the measure that includes the explanatory power of the prefecture fixed effects themselves, is not really informative in this application. It is a little higher, but with a mean of around 0.046 still extremely low. The F statistic is even lower in the fixed effects approach than in the baseline model. When estimating both specifications with variations in empire size, the lines are closer together. A large share of coefficient estimates is also individually statistically insignificant.

Overall, this final test test shows that our baseline estimates are robust to using prefecture fixed effects. Prefecture seat locations are much less driven by local geography than county seat and market town locations are.

Table C-21: F Statistics in Baseline Pixel Estimations

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Intercept	14.84***	18.31***	15.05***	10.39***	17.31***	16.43***	15.62***	14.69***	12.50***	12.93***	12.62***
	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Pref.	−12.79***	−15.68***	−11.70***	−6.52***	−12.72***	−11.97***	−10.60***	−9.63***	−8.64***	−9.06***	−8.55***
	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
R^2	0.99	0.99	0.98	0.96	0.99	0.99	0.98	0.98	0.98	0.98	0.98
Num. obs.	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000

Notes: The dependent variables are F statistics from 1,000 bootstrapped county seat and 1,000 bootstrapped prefecture seat regressions. Observations are drawn with replacement and the bootstrapped sample is as large as the original baseline data set. The estimations generating the F statistics use the specification outlined in eq. (14) and employed e.g. in Table 2 and Table 3. We regress the resulting F statistics on a prefecture seat indicator that equals one for results derived from prefecture seat regressions and zero for county seat regressions. Standard errors are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$).

Table C-22: Adjusted R^2 in Baseline Pixel Estimations

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Intercept	0.18***	0.21***	0.18***	0.13***	0.20***	0.19***	0.19***	0.18***	0.15***	0.16***	0.15***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Pref.	−0.16***	−0.19***	−0.14***	−0.09***	−0.15***	−0.14***	−0.13***	−0.12***	−0.11***	−0.11***	−0.11***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.99	0.99	0.98	0.96	0.99	0.99	0.99	0.98	0.98	0.98	0.98
Num. obs.	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000

Notes: The dependent variables are adjusted R^2 from 1,000 bootstrapped county seat and 1,000 bootstrapped prefecture seat regressions. Observations are drawn with replacement and the bootstrapped sample is as large as the original baseline data set. The estimations generating the adjusted R^2 use the specification outlined in eq. (14) and employed e.g. in Table 2 and Table 3. We regress the resulting adjusted R^2 on a prefecture seat indicator that equals one for results derived from prefecture seat regressions and zero for county seat regressions. Standard errors are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$).

Table C-23: F Statistics and Alternative Pixel Sizes

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
A. Very Small Pixels											
Intercept	16.42***	21.35***	16.90***	12.60***	17.83***	16.64***	15.63***	14.91***	13.18***	13.62***	14.19***
	(0.03)	(0.05)	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.05)	(0.06)
Pref.	−14.27***	−17.39***	−12.53***	−5.99***	−12.58***	−10.32***	−9.28***	−9.30***	−6.37***	−5.88***	−6.09***
	(0.04)	(0.07)	(0.07)	(0.07)	(0.05)	(0.05)	(0.05)	(0.04)	(0.07)	(0.07)	(0.08)
R^2	0.98	0.96	0.95	0.77	0.97	0.96	0.95	0.97	0.82	0.76	0.74
Num. obs.	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000
B. Small Pixels											
Intercept	15.42***	20.89***	16.50***	11.76***	17.95***	15.98***	14.85***	14.28***	12.49***	12.63***	12.58***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Pref.	−13.73***	−18.09***	−13.38***	−7.22***	−12.88***	−11.00***	−10.12***	−9.70***	−8.59***	−8.68***	−8.07***
	(0.03)	(0.05)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
R^2	0.99	0.99	0.98	0.94	0.98	0.98	0.98	0.99	0.98	0.98	0.96
Num. obs.	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000
C. Large Pixels											
Intercept	11.99***	14.06***	12.25***	9.73***	15.32***	14.87***	14.07***	13.84***	12.23***	13.27***	12.77***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Pref.	−10.34***	−11.74***	−9.57***	−6.65***	−11.12***	−10.60***	−9.63***	−9.45***	−9.04***	−10.27***	−9.73***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)
R^2	0.98	0.98	0.98	0.97	0.98	0.98	0.98	0.98	0.98	0.99	0.99
Num. obs.	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000
D. Very Large Pixels											
Intercept	11.74***	12.25***	11.31***	9.49***	14.17***	14.12***	14.44***	13.74***	11.67***	12.97***	13.20***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Pref.	−9.98***	−9.68***	−8.39***	−6.19***	−10.18***	−9.88***	−9.83***	−9.35***	−8.64***	−9.57***	−9.74***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
R^2	0.98	0.98	0.98	0.97	0.98	0.98	0.98	0.98	0.98	0.98	0.99
Num. obs.	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000

Notes: The dependent variables are F statistics from 1,000 bootstrapped county seat and 1,000 bootstrapped prefecture seat regressions. Observations are drawn with replacement and the bootstrapped sample is as large as the original data set. Grid cell resolutions refer to the same pixel sizes as they do throughout the rest of the paper, e.g. in Section C.2. The estimations generating the F statistics use the specification outlined in eq. (14) and employed e.g. in Table 2 and Table 3. We regress the resulting F statistics on a prefecture seat indicator that equals one for results derived from prefecture seat regressions and zero for county seat regressions. Standard errors are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$).

Table C-24: Adjusted R^2 and Alternative Pixel Sizes

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
A. Very Small Pixels											
Intercept	0.03***	0.04***	0.03***	0.02***	0.03***	0.03***	0.03***	0.03***	0.02***	0.02***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Pref.	−0.03***	−0.03***	−0.02***	−0.01***	−0.02***	−0.02***	−0.02***	−0.02***	−0.01***	−0.01***	−0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.98	0.96	0.95	0.77	0.97	0.96	0.95	0.97	0.82	0.76	0.74
Num. obs.	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000
B. Small Pixels											
Intercept	0.09***	0.13***	0.10***	0.07***	0.11***	0.10***	0.09***	0.09***	0.08***	0.08***	0.08***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Pref.	−0.09***	−0.11***	−0.09***	−0.05***	−0.08***	−0.07***	−0.06***	−0.06***	−0.06***	−0.06***	−0.05***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.99	0.99	0.98	0.94	0.98	0.98	0.99	0.99	0.97	0.98	0.96
Num. obs.	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000
C. Large Pixels											
Intercept	0.23***	0.26***	0.23***	0.19***	0.28***	0.27***	0.26***	0.26***	0.23***	0.25***	0.24***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Pref.	−0.21***	−0.22***	−0.19***	−0.14***	−0.20***	−0.19***	−0.18***	−0.17***	−0.18***	−0.20***	−0.19***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.98	0.99	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Num. obs.	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000
D. Very Large Pixels											
Intercept	0.30***	0.31***	0.29***	0.25***	0.35***	0.35***	0.35***	0.34***	0.30***	0.33***	0.33***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Pref.	−0.27***	−0.25***	−0.22***	−0.17***	−0.24***	−0.23***	−0.22***	−0.22***	−0.22***	−0.24***	−0.24***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.98	0.98	0.98	0.97	0.98	0.99	0.99	0.99	0.99	0.99	0.99
Num. obs.	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000

Notes: The dependent variables are adjusted R^2 from 1,000 bootstrapped county seat and 1,000 bootstrapped prefecture seat regressions. Observations are drawn with replacement and the bootstrapped sample is as large as the original data set. Grid cell resolutions refer to the same pixel sizes as they do throughout the rest of the paper, e.g. in Section C.2. The estimations generating the adjusted R^2 use the specification outlined in eq. (14) and employed e.g. in Table 2 and Table 3. We regress the resulting adjusted R^2 on a prefecture seat indicator that equals one for results derived from prefecture seat regressions and zero for county seat regressions. Standard errors are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$).

Table C-25: Goodness of Fit and the Number of Urban Pixels

	Adjusted R^2					F Statistic				
	Very Small	Small	Medium	Large	Very Large	Very Small	Small	Medium	Large	Very Large
Pref.	-0.03^{***}	-0.10^{***}	-0.15^{***}	-0.14^{***}	-0.17^{***}	-17.55^{***}	-14.83^{***}	-11.46^{***}	-6.62^{***}	-5.89^{***}
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(1.85)	(1.55)	(1.08)	(0.51)	(0.42)
Num. Urb. Pixels	-0.02^{***}	-0.03^{***}	-0.03^{*}	0.06^{***}	0.10^{***}	-8.70^{***}	-5.13^{***}	-2.11^{*}	3.47^{***}	4.33^{***}
	(0.00)	(0.01)	(0.01)	(0.01)	(0.02)	(1.90)	(1.64)	(1.24)	(0.69)	(0.66)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.80	0.89	0.92	0.97	0.98	0.80	0.88	0.91	0.97	0.97
Num. obs.	428	428	428	428	428	428	428	428	428	428

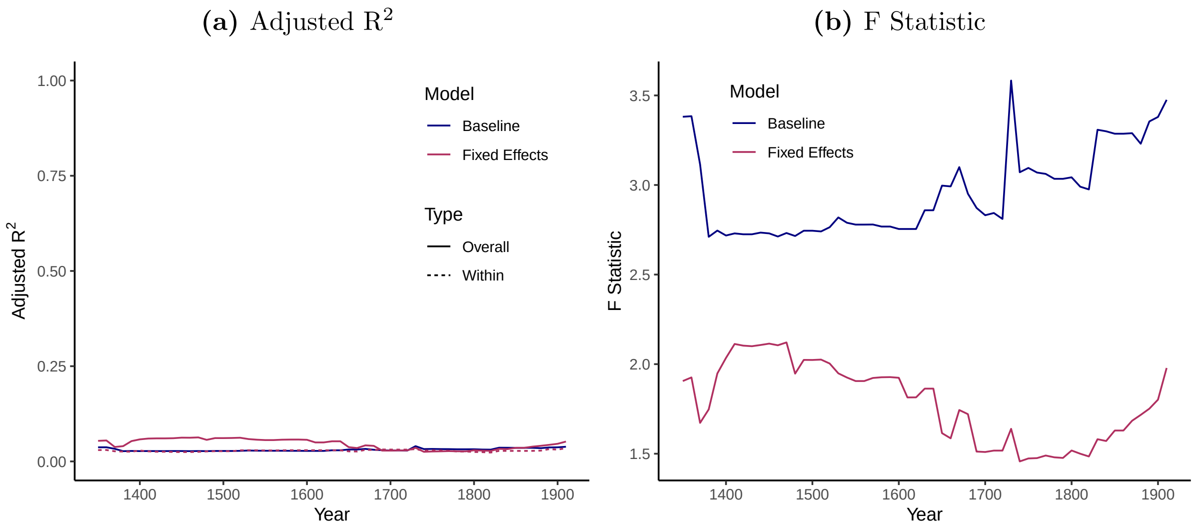
Notes: The dependent variables are adjusted R^2 and F statistics from 214 cross-sectional county seat and 214 cross-sectional prefecture seat regressions, with one cross-section every ten years between 220 BCE and 1910 CE. Grid cell resolutions refer to the same pixel sizes as they do throughout the rest of the paper, e.g. in Section C.2, with *Medium* denoting the baseline pixel size. The estimations generating the adjusted R^2 and the F statistics use the specification outlined in eq. (14) and employed e.g. in Table 2 and Table 3. We regress the resulting adjusted R^2 and F statistics on a prefecture seat indicator, the number of urban pixels in thousands, and year fixed effects. The indicator equals one for results derived from prefecture seat regressions and zero for county seat regressions. The number of urban pixels is the number of cells with a value of one, e.g. the number of cells hosting a county seat in county seat regressions. Standard error clustered at the year level are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$). The prefecture seat indicator coefficient estimate remains statistically significant irrespective of whether we cluster at the year level, the settlement type level, or both.

Table C-26: Predicted Probabilities and the Number of Urban Pixels

	Very Small	Small	Medium	Large	Very Large
Pref.	-0.03^{***}	-0.07^{***}	-0.14^{***}	-0.14^{***}	-0.17^{***}
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Num. Urb. Pixels	1.15^{***}	0.80^{***}	0.50^{***}	0.09^{**}	0.14^{***}
	(0.06)	(0.06)	(0.04)	(0.04)	(0.04)
$Pr(Y = 1 \| Y = 0)$	-101.85^{***}	-18.45^{***}	-5.15^{***}	0.77^{**}	0.75^{***}
	(4.90)	(1.49)	(0.49)	(0.31)	(0.17)
Year FE	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.97	0.97	0.98	0.99	0.99
Num. obs.	428	428	428	428	428

Notes: The dependent variables are the mean predicted probabilities of a pixel to be urban conditional on being urban from 214 cross-sectional county seat and 214 cross-sectional prefecture seat regressions, with one cross-section every ten years between 220 BCE and 1910 CE. Grid cell resolutions refer to the same pixel sizes as they do throughout the rest of the paper, e.g. in Section C.2, with *Medium* denoting the baseline pixel size. The estimations generating the probabilities use the specification outlined in eq. (14) and employed e.g. in Table 2 and Table 3. We regress the resulting mean probabilities within the group of urban pixels on a prefecture seat indicator, the number of urban pixels in thousands, the mean probabilities within the group of rural pixels, and year fixed effects. The indicator equals one for results derived from prefecture seat regressions and zero for county seat regressions. The number of urban pixels is the number of cells with a value of one, e.g. the number of cells hosting a county seat in county seat regressions. Standard error clustered at the year level are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$). The prefecture seat indicator coefficient estimate remains statistically significant irrespective of whether we cluster at the year level, the settlement type level, or both.

Figure C-2: Goodness of Fit and Prefecture Fixed Effects



Notes: The figures plot the adjusted R^2 and F statistics of cross-sectional OLS regressions on local geography. Whereas Table 3 reports these values for eleven selected cross-sections, this figure prints a result for one cross-section every ten years between 1350 and 1910 CE. The computations use the baseline resolution. The fixed effects specification modifies the model from Table 3, and thereby eq. (14), in that it includes prefecture-fixed effects, controls for soil characteristics via indicator variables rather than demeaning-based fixed effects, and employs variations in empire size determined by prefecture borders. The two specifications' F statistics are closer together, if both approaches use the same variable empire size rather than the baseline specification using the default constant empire size.

C.7 Supplementary Results on Random Forest

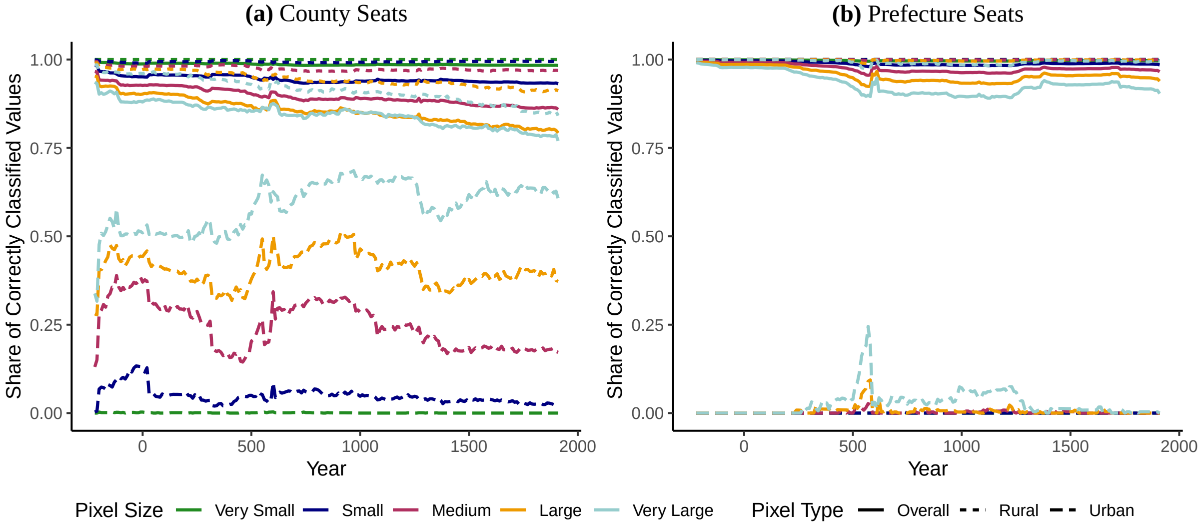
Here we provide supplementary results to the random forest classification presented in Section 4.1.

(i) We first provide further results on how the classification random forest can distinguish urban from rural pixels. Most pixels are rural and the difficult part is to correctly spot the few urban grid cells using geographic characteristics. According to Figure 5, the algorithm manages to correctly classify more than 60% of county seat locations as urban at the coarsest resolution. In contrast, it marks almost no prefecture seat locations as urban.

Technically, the algorithm could reach a high share of correctly classified urban pixels by generally labeling more pixels urban. In a statistical sense, we would speak of a high sensitivity. This might come at the expense of a low specificity, which would be the case if a lower share of rural pixels were to be correctly classified as rural. In the following, we illustrate that the prediction quality on urban pixels does not undermine the prediction quality on rural pixels.

In Figure C-3 we compare the fraction of urban, rural, and overall pixels that were correctly classified. The higher classification quality of county seat locations at low resolutions is accompanied by slightly worse predictions of rural pixels. And rural pixels are, apparently, slightly more difficult to identify in later years than in earlier ones. However, that slope is unrelated to the shape of the urban pixel's prediction quality over the years. More correctly classified urban pixels does not imply fewer correctly classified rural pixels. Moreover, we see that the changes over time are never as large as the differences in the prediction performance between county seats (Figure C-3a) and prefecture seats (Figure C-3b).

Figure C-3: Classification Random Forest Results



Notes: This figure extends Figure 5, adding the share of correctly classified rural pixels and the share of overall correctly classified pixels. Grid cells are predicted to be urban, if the probability is at least 0.5. Otherwise they are classified as rural.

An econometric examination in the form of the following identification strategy confirms those conclusions:

$$\ln Urban_{st} = \alpha\, County_{st} + \beta \ln Rural_{st} + \eta_t + \varepsilon_{st} \qquad \text{((C-5))}$$

We regress the natural logarithm of the share of correctly classified urban pixels U derived from a classification random forest targeting seat type s in year t (the values displayed in Figure 5 and named *Urban* in Figure C-3) on a county seat indicator *County* that equals one for values obtained from county seat estimations and zero for prefecture seat estimations, while controlling for the natural logarithm of the share of correctly classified rural pixels *Rural* (named *Rural* in Figure C-3), and year-fixed effects η. Because the machine learning algorithm does not correctly identify any prefecture seat pixels as urban at the smallest pixel size and only very few at the next larger aggregation, we restrict the elasticity estimations to the top three pixel sizes. Furthermore, we exclude any year in which the share of correctly classified urban or rural pixels is zero. As this drops a lot of observations, the following Table C-27 also reports results computed via Poisson pseudo maximum likelihood (PPML), a method that is common in the trade literature and that allows to include zeros on the left hand side (Santos Silva and Tenreyro, 2006; Head and Mayer, 2014).[3] The coefficient estimates of α vary substantially across data sets and estimation strategies. However, in all cases they are positive and statistically significant. Even after controlling for the algorithm's performance in identifying rural grid

[3]Zeros are not dropped as they do not inherently enter the model as logs. PPML is not the perfect solution to any scenario with zeros on the left hand side, but it provides one additional piece of evidence (Head and Mayer, 2014).

cells, it identifies county seats better than prefecture seats.

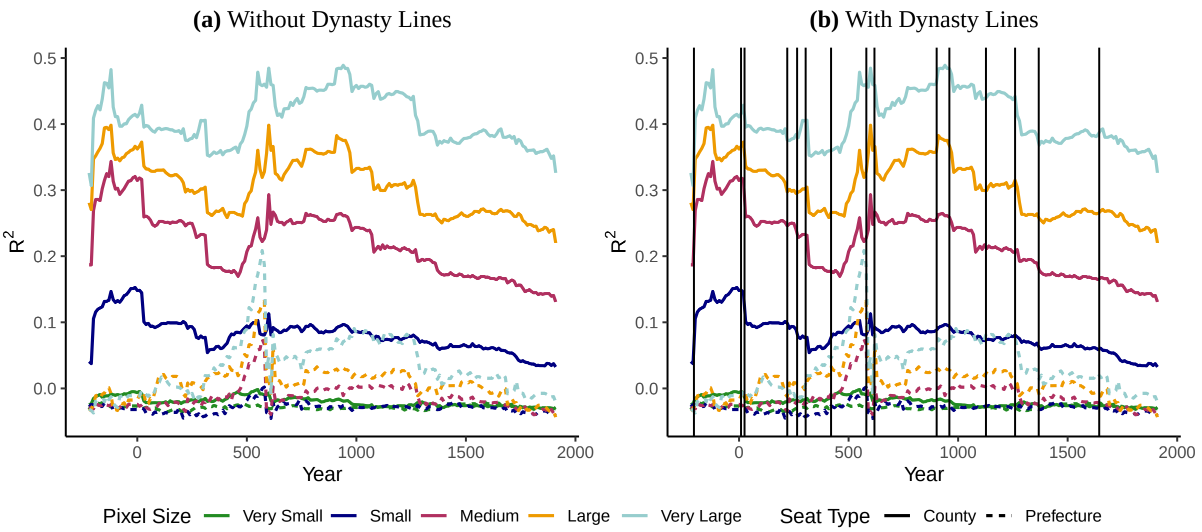
Table C-27: Administrative Status and the Share of Correctly Classified Pixels

	OLS			PPML		
	Medium	Large	Very Large	Medium	Large	Very Large
County	1.43^{*}	2.78^{***}	1.12^{***}	6.79^{***}	4.10^{***}	2.64^{***}
	(0.78)	(0.26)	(0.18)	(0.59)	(0.30)	(0.21)
ln(Rural)	-79.69^{***}	-22.10^{***}	-23.06^{***}	69.26^{***}	-4.29	-6.00^{***}
	(26.78)	(3.93)	(1.59)	(19.23)	(3.72)	(1.94)
Year	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.95	0.94	0.94			
Num. obs.	108	212	304	428	428	428

Notes: The dependent variable is the (natural logarithm of the) share of correctly classified urban pixels. Standard errors clustered at the year level are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$). *Medium*, *Large*, and *Very Large* denote the pixel size, using the same definitions as in other sections of the paper. *County* is an indicator that equals one if the respective share of correctly classified pixels is derived from a county seat estimation, and zero if it refers to a prefecture seat estimation. *Rural* is the share of correctly classified rural pixels.

(ii) The paper visualizes the share of correctly classified county and prefecture seats of the random forest classifier using different resolutions in Figure 5. In Figure C-4, we repeat this analysis using the R^2 from random forest regression rather than the correctly classified share from random forest classification. The pattern is very similar and also coincides with dynastic change.

Figure C-4: R^2 in Regression Random Forests

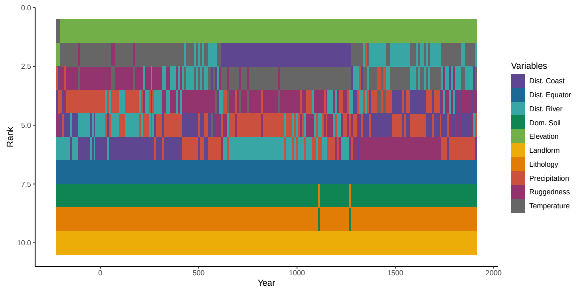


Notes: We derive the results from 214 independently tested cross-sections, with one cross-section every ten years between 220 BCE and 1919 CE. The dependent variable is an indicator that equals one, if the pixel hosts an administrative settlement of the respective type in that year, and zero otherwise. This figure repeats the classification random forest estimations in Section 4.1 with regression random forests. Regression random forests produce - unlike the classification variant - R^2 results. The figure is the regression counterpart of Figure 5. The vertical lines mark the dynastic changes in 206 BCE, 9 CE, 25 CE, 220 CE, 265 CE, 304 CE, 420 CE, 581 CE, 618 CE, 902 CE, 960 CE, 1127 CE, 1260 CE, 1368 CE, and 1644 CE.

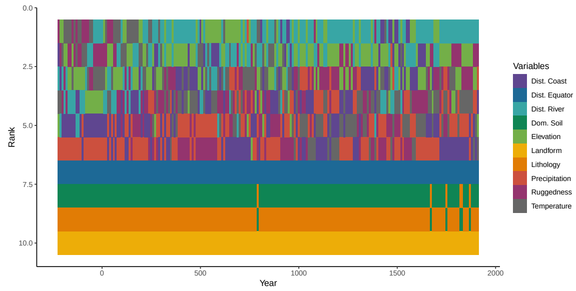
(iii) Figure 4 illustrates the variable importance of geographic factors in predicting city locations at the baseline resolution. The following Figure C-5 plots the corresponding variable rankings when using alternative pixel sizes. It confirms the baseline result of ranks being more stable over time in county seat estimations than in prefecture seat estimations.

Figure C-5: Variable Importance at Different Pixel Sizes

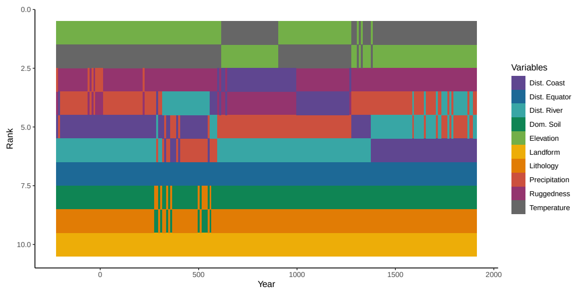
(a) County Seats, Very Small Pixels



(b) Prefecture Seats, Very Small Pixels

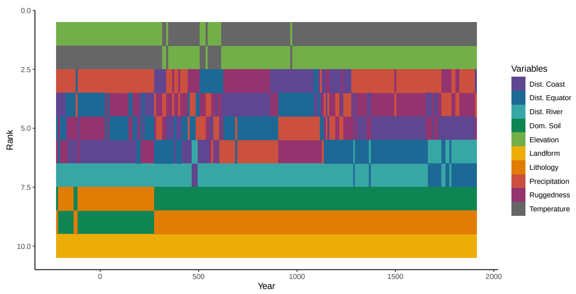


(c) County Seats, Small Pixels

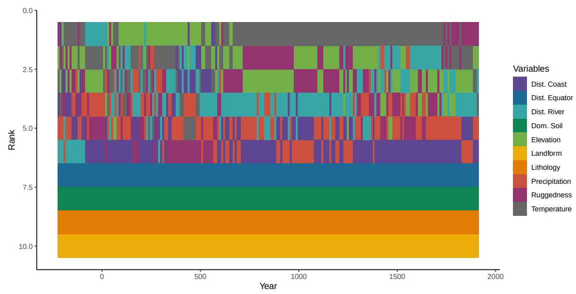


(d) Prefecture Seats, Small Pixels

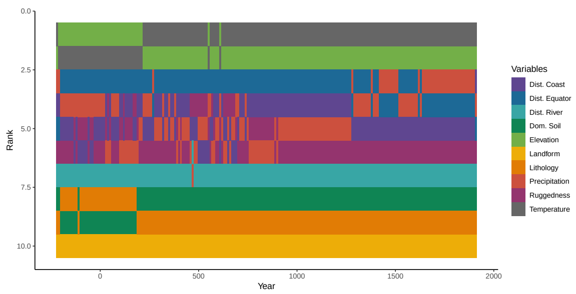
(e) County Seats, Large Pixels



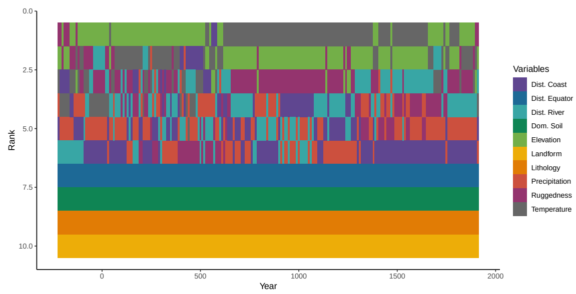
(f) Prefecture Seats, Large Pixels



(g) County Seats, Very Large Pixels



(h) Prefecture Seats, Very Large Pixels



Notes: Rank one denotes the most important and rank ten the least important variable in predicting county seat locations. The dominant soil type, landform, and lithology are categorical variables, the remaining regressors are continuous, and the dependent variable indication whether a pixel hosts a county or prefecture seat is binary. We derive the results from 214 independently tested cross-sections, with one cross-section every ten years between 220 BCE and 1910 CE. The input pixels are the same as those used in the econometric estimations. Figure 4 depicts the corresponding plots using the medium (baseline) pixel size.

(iv) In addition to identifying structural breaks visually in Figure 5b and related graphs, we also run a few econometric tests underlining the results. First, we use Bai and Perron's (2003) method locating multiple breakpoints in the data. We restrict the algorithm to identify at most 15 breaks, which equals the number of dynastic changes depicted in the plots. Comparing the 95% confidence intervals of the estimated breaks with the actual institutional transitions, Bai and Perron's (2003) method confirms what we already see in the plots. In Figure 5 (Figure C-4), the test catches 40% (around 53) of the dynastic transitions in the baseline county seat results and around 7% (33) in the baseline prefecture seat results. It identifies major political events, while being more restricted than a visual analysis of the plot in that it cannot spot close-by breaks (Bai, 1997; Bai and Perron, 1998, 2003).

The second test does not search for the breakpoints itself, but evaluates the statistical significance of a potential discontinuity accompanying dynastic transitions. We use the following spline regression which regresses the goodness of fit measures F_t, i.e. the share of correctly classified urban pixels and the R^2, in year t on a spline indicator $I(T_t \geq \psi_k)$ and its interaction with a time trend T_t. The indicator equals one, if year T_t is equal to or larger than the year ψ_k of dynastic transition k. We therefore allow for jumps and slope changes in the year a new dynasty arises.

$$F_t = \sum_{k=1}^{K} I(T_t \geq \psi_k)(\alpha_k + \beta_k T_t) + \varepsilon_t \tag{(C-6)}$$

The results listed in Table C-28 and Table C-29 are in line with the previous findings. There appear to be significant continuities intersecting with the rise of new dynasties. And these changes stand out in particular for county seats at larger grid cell sizes.

Table C-28: Spline Regressions on Correctly Classified Urban Pixels

	County Seats					Prefecture Seats				
	V. Small	Small	Medium	Large	V. Large	V. Small	Small	Medium	Large	V. Large
Post_{221BCE}	-0.00	0.06	0.57	0.22	-0.15	0.00	-0.00	-0.00	-0.00	-0.00
	(0.02)	(0.16)	(0.42)	(0.46)	(0.41)	(0.00)	(0.01)	(0.07)	(0.19)	(0.35)
Post_{206BCE}	0.00^{***}	0.14^{***}	0.39^{***}	0.45^{***}	0.51^{***}	0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Post_{9CE}	0.00^{*}	0.10^{***}	0.38^{***}	0.44^{***}	0.51^{***}	0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.01)	(0.03)	(0.03)	(0.03)	(0.00)	(0.00)	(0.00)	(0.01)	(0.03)
Post_{25CE}	0.00^{**}	0.05^{***}	0.31^{***}	0.43^{***}	0.51^{***}	0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Post_{220CE}	-0.00	0.04	0.23^{**}	0.37^{***}	0.32^{***}	0.00	0.00	-0.00	-0.05	-0.00
	(0.00)	(0.04)	(0.10)	(0.12)	(0.10)	(0.00)	(0.00)	(0.02)	(0.05)	(0.09)
Post_{265CE}	0.00	0.05	0.30^{*}	0.08	0.06	0.00	-0.00	0.00	0.00	0.00
	(0.01)	(0.07)	(0.18)	(0.19)	(0.17)	(0.00)	(0.00)	(0.03)	(0.08)	(0.15)
Post_{304CE}	0.00	0.04^{**}	0.39^{***}	0.41^{***}	0.53^{***}	0.00	-0.00	0.00	-0.00	0.00
	(0.00)	(0.02)	(0.05)	(0.05)	(0.05)	(0.00)	(0.00)	(0.01)	(0.02)	(0.04)
Post_{420CE}	-0.00^{***}	0.04^{***}	-0.06^{*}	0.01	0.04	0.00	-0.00	-0.06^{***}	-0.23^{***}	-0.54^{***}
	(0.00)	(0.01)	(0.03)	(0.04)	(0.03)	(0.00)	(0.00)	(0.01)	(0.02)	(0.03)
Post_{581CE}	0.06^{***}	-0.12	-0.29	0.26	-0.01	0.00	0.00	0.00	0.14	0.38
	(0.02)	(0.22)	(0.59)	(0.64)	(0.57)	(0.00)	(0.01)	(0.09)	(0.26)	(0.49)
Post_{618CE}	0.00^{**}	0.08^{***}	0.22^{***}	0.24^{***}	0.37^{***}	0.00	0.00	-0.00	0.00	0.06^{***}
	(0.00)	(0.01)	(0.02)	(0.03)	(0.02)	(0.00)	(0.00)	(0.00)	(0.01)	(0.02)
Post_{902CE}	0.00	-0.06	0.52	0.72	0.62	0.00	0.00	-0.00	-0.01	-0.07
	(0.01)	(0.15)	(0.41)	(0.45)	(0.39)	(0.00)	(0.01)	(0.06)	(0.18)	(0.34)
Post_{960CE}	0.00	0.16^{***}	0.83^{***}	0.96^{***}	0.84^{***}	0.00	0.00	-0.04^{***}	0.03	0.01
	(0.00)	(0.03)	(0.07)	(0.08)	(0.07)	(0.00)	(0.00)	(0.01)	(0.03)	(0.06)
Post_{1127CE}	-0.00	-0.04	0.27^{**}	0.34^{**}	0.64^{***}	0.00	0.00	0.02	-0.03	0.03
	(0.00)	(0.05)	(0.12)	(0.13)	(0.12)	(0.00)	(0.00)	(0.02)	(0.05)	(0.10)
Post_{1260CE}	0.00	0.07	0.53^{***}	1.05^{***}	1.30^{***}	0.00	0.00	0.02	0.05	0.51^{***}
	(0.01)	(0.07)	(0.17)	(0.19)	(0.17)	(0.00)	(0.00)	(0.03)	(0.08)	(0.14)
Post_{1368CE}	-0.00	0.00	0.11^{**}	0.07	0.15^{***}	0.00	-0.00	0.02^{**}	-0.00	-0.01
	(0.00)	(0.02)	(0.05)	(0.05)	(0.05)	(0.00)	(0.00)	(0.01)	(0.02)	(0.04)
Post_{1644CE}	-0.00	0.09^{***}	0.20^{***}	0.39^{***}	0.67^{***}	0.00	0.00	0.00	0.04	0.00
	(0.00)	(0.02)	(0.06)	(0.07)	(0.06)	(0.00)	(0.00)	(0.01)	(0.03)	(0.05)
Year $\times$ Post_{221BCE}	-0.00	0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year $\times$ Post_{206BCE}	-0.00	0.00^{***}	0.00^{***}	0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year $\times$ Post_{9CE}	-0.00	0.00^{*}	-0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year $\times$ Post_{25CE}	-0.00	0.00	-0.00	-0.00^{***}	-0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year $\times$ Post_{220CE}	0.00	0.00	0.00	-0.00	0.00^{*}	0.00	-0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year $\times$ Post_{265CE}	-0.00	-0.00	-0.00	0.00	0.00^{***}	0.00	0.00	-0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year $\times$ Post_{304CE}	-0.00	-0.00	-0.00^{***}	-0.00	-0.00	0.00	0.00	-0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year $\times$ Post_{420CE}	0.00^{***}	0.00	0.00^{***}	0.00^{***}	0.00^{***}	0.00	0.00	0.00^{***}	0.00^{***}	0.00^{***}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year $\times$ Post_{581CE}	-0.00^{***}	0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year $\times$ Post_{618CE}	-0.00	-0.00^{**}	0.00^{***}	0.00^{***}	0.00^{***}	0.00	-0.00	0.00	0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year $\times$ Post_{902CE}	-0.00	0.00	-0.00	-0.00	0.00	0.00	-0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year $\times$ Post_{960CE}	-0.00	-0.00^{***}	-0.00^{***}	-0.00^{***}	-0.00^{***}	0.00	-0.00	0.00^{***}	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year $\times$ Post_{1127CE}	0.00	0.00^{*}	-0.00	0.00	0.00	0.00	-0.00	-0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year $\times$ Post_{1260CE}	-0.00	-0.00	-0.00^{*}	-0.00^{***}	-0.00^{***}	0.00	-0.00	-0.00	-0.00	-0.00^{***}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year $\times$ Post_{1368CE}	0.00	0.00^{*}	0.00	0.00^{***}	0.00^{***}	0.00	0.00	-0.00^{**}	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year $\times$ Post_{1644CE}	0.00	-0.00^{***}	-0.00	0.00	-0.00	0.00	-0.00	-0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Adj. R^2	0.74	0.99	1.00	1.00	1.00	–	-0.13	0.73	0.78	0.93
Num. obs.	214	214	214	214	214	214	214	214	214	214

Notes: Standard errors are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$). The dependent variable, the share of correctly classified urban pixels plotted in Figure 5, is zero for all observations in the very small pixel-prefecture seat case. Without variation, no goodness of fit measures etc. are estimated. The 214 observations refer to the 214 cross-sections that Figure 5 is based on, with one cross-section every ten years between 220 BCE and 1910 CE.

Table C-29: Spline Regressions on the R^2

	County Seats					Prefecture Seats				
	V. Small	Small	Medium	Large	V. Large	V. Small	Small	Medium	Large	V. Large
$Post_{221BCE}$	0.01	−0.03	0.19	0.05	−0.09	0.11**	0.06	−0.08	0.03	−0.07
	(0.05)	(0.13)	(0.25)	(0.31)	(0.33)	(0.05)	(0.08)	(0.13)	(0.28)	(0.26)
$Post_{206BCE}$	−0.01***	0.15***	0.32***	0.36***	0.41***	−0.03***	−0.03***	−0.03***	−0.02***	−0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Post_{9CE}$	−0.00	0.15***	0.32***	0.36***	0.40***	−0.03***	−0.03***	−0.03***	−0.01	−0.02
	(0.00)	(0.01)	(0.02)	(0.02)	(0.02)	(0.00)	(0.01)	(0.01)	(0.02)	(0.02)
$Post_{25CE}$	−0.02***	0.10***	0.26***	0.34***	0.40***	−0.03***	−0.04***	−0.03***	−0.02***	−0.01
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
$Post_{220CE}$	−0.04***	0.13***	0.27***	0.31***	0.33***	−0.06***	−0.07***	−0.07**	−0.10	0.05
	(0.01)	(0.03)	(0.06)	(0.08)	(0.08)	(0.01)	(0.02)	(0.03)	(0.07)	(0.07)
$Post_{265CE}$	−0.04*	0.07	0.18*	0.25*	0.14	−0.02	−0.01	0.00	0.09	−0.11
	(0.02)	(0.06)	(0.10)	(0.13)	(0.14)	(0.02)	(0.04)	(0.06)	(0.12)	(0.11)
$Post_{304CE}$	−0.02***	0.06***	0.29***	0.33***	0.42***	−0.02***	−0.06***	−0.03**	−0.01	−0.07**
	(0.01)	(0.02)	(0.03)	(0.03)	(0.04)	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)
$Post_{420CE}$	−0.02***	0.06***	−0.04*	0.02	0.01	−0.03***	−0.12***	−0.27***	−0.30***	−0.41***
	(0.00)	(0.01)	(0.02)	(0.03)	(0.03)	(0.00)	(0.01)	(0.01)	(0.02)	(0.02)
$Post_{581CE}$	0.23***	0.32*	0.01	0.23	0.34	0.05	0.22*	0.04	0.39	1.18***
	(0.07)	(0.19)	(0.35)	(0.43)	(0.46)	(0.07)	(0.12)	(0.18)	(0.39)	(0.36)
$Post_{618CE}$	−0.01***	0.10***	0.25***	0.25***	0.33***	−0.02***	−0.02***	−0.06***	−0.02	0.02
	(0.00)	(0.01)	(0.01)	(0.02)	(0.02)	(0.00)	(0.00)	(0.01)	(0.02)	(0.01)
$Post_{902CE}$	0.06	−0.01	0.32	0.54*	0.43	−0.05	−0.11	0.18	−0.00	0.13
	(0.05)	(0.13)	(0.24)	(0.30)	(0.32)	(0.05)	(0.08)	(0.13)	(0.27)	(0.25)
$Post_{960CE}$	0.02***	0.19***	0.52***	0.67***	0.64***	−0.04***	−0.03**	−0.01	0.06	0.13***
	(0.01)	(0.02)	(0.04)	(0.05)	(0.06)	(0.01)	(0.01)	(0.02)	(0.05)	(0.04)
$Post_{1127CE}$	−0.05***	0.03	0.26***	0.28***	0.49***	−0.01	−0.01	−0.03	0.04	0.12
	(0.01)	(0.04)	(0.07)	(0.09)	(0.09)	(0.01)	(0.02)	(0.04)	(0.08)	(0.08)
$Post_{1260CE}$	−0.04*	0.23***	0.38***	0.85***	0.90***	−0.02	−0.03	0.26***	0.02	0.54***
	(0.02)	(0.06)	(0.10)	(0.13)	(0.13)	(0.02)	(0.03)	(0.05)	(0.12)	(0.11)
$Post_{1368CE}$	−0.02***	0.07***	0.22***	0.17***	0.28***	−0.05***	−0.02*	0.01	0.04	0.00
	(0.01)	(0.02)	(0.03)	(0.04)	(0.04)	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)
$Post_{1644CE}$	0.00	0.23***	0.37***	0.54***	0.67***	0.02***	0.00	0.07***	0.17***	0.18***
	(0.01)	(0.02)	(0.04)	(0.04)	(0.05)	(0.01)	(0.01)	(0.02)	(0.04)	(0.04)
Year × $Post_{221BCE}$	0.00	−0.00	−0.00	−0.00	−0.00	0.00***	0.00	−0.00	0.00	−0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year × $Post_{206BCE}$	0.00***	0.00***	0.00***	−0.00	−0.00***	−0.00***	−0.00***	0.00	−0.00*	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year × $Post_{9CE}$	−0.00	−0.00	−0.00	0.00	0.00	−0.00	−0.00	0.00	−0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year × $Post_{25CE}$	0.00	0.00	−0.00	−0.00**	−0.00	−0.00***	−0.00	0.00*	0.00***	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year × $Post_{220CE}$	0.00**	−0.00	−0.00	−0.00	0.00	0.00**	0.00	0.00	0.00	−0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year × $Post_{265CE}$	0.00	0.00	0.00	0.00	0.00*	−0.00	−0.00	−0.00	−0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year × $Post_{304CE}$	0.00	0.00	−0.00***	−0.00*	−0.00*	−0.00	0.00***	0.00	0.00	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year × $Post_{420CE}$	0.00**	0.00***	0.00***	0.00***	0.00***	0.00	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year × $Post_{581CE}$	−0.00***	−0.00	0.00	0.00	0.00	−0.00	−0.00**	−0.00	−0.00	−0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year × $Post_{618CE}$	−0.00	−0.00	0.00	0.00***	0.00***	−0.00***	−0.00**	0.00***	0.00***	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year × $Post_{902CE}$	−0.00	0.00	−0.00	−0.00	0.00	0.00	0.00	−0.00	0.00	−0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year × $Post_{960CE}$	−0.00***	−0.00***	−0.00***	−0.00***	−0.00***	0.00	0.00	0.00	−0.00	−0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year × $Post_{1127CE}$	0.00	0.00	−0.00	0.00	−0.00	−0.00*	−0.00	0.00	−0.00	−0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year × $Post_{1260CE}$	0.00	−0.00***	−0.00*	−0.00***	−0.00***	−0.00	0.00	−0.00***	−0.00	−0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year × $Post_{1368CE}$	−0.00	−0.00	−0.00*	0.00**	0.00**	0.00***	−0.00	−0.00*	−0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year × $Post_{1644CE}$	−0.00***	−0.00***	−0.00***	−0.00***	−0.00***	−0.00***	−0.00***	−0.00***	−0.00***	−0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Adj. R^2	0.99	1.00	1.00	1.00	1.00	1.00	0.99	0.96	0.89	0.97
Num. obs.	214	214	214	214	214	214	214	214	214	214

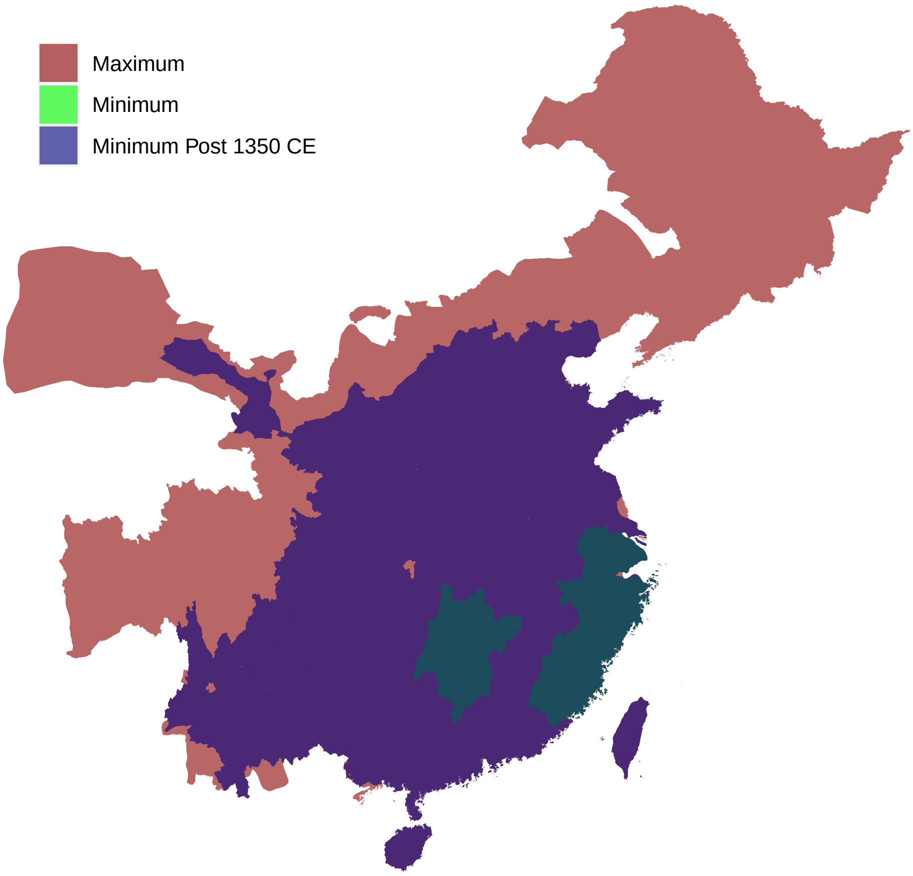
Notes: Standard errors are in parentheses (***$p < 0.01$; **$p < 0.05$; *$p < 0.1$). The dependent variable is the R^2 plotted in Figure C-4. It measures the explanatory power of local geography for county and prefecture seat locations, estimated via regression random forests. The 214 observations refer to the 214 cross-sections that Figure C-4 is based on, with one cross-section every ten years between 220 BCE and 1910 CE.

C.8 Robustness Check: Alternative Empire Size

The baseline estimation in Section 4.1 relies on the empire's maximum extent, as observed in the CHGIS data. Here, we discuss this point and provide robustness results for alternative shapes.

Holding imperial borders constant helps us to overcome missing prefecture borders before 1350 CE. The alternative choice, altering the shape every year and only accounting for pixels that certainly were part of the empire, excludes a lot of cities in earlier years, producing a much smaller, non-representative set of settlements. The maximum extent as a constant shape makes sure that relevant areas are included, at the cost of adding some remote regions that did not belong to the empire in all years. Given the distinctive geographic differences between the periphery and the heartland, adding these additional rural pixels should not pose a major problem to the random forest estimations. Nonetheless, we test alternative shapes in the following paragraphs to evaluate any potential biases.

Figure C-6: Constant Extents of the Empire



Notes: The maximum extent is the space ever covered by a prefecture in any of the years and used in the baseline estimations. The minimum shape is the area covered by prefectures in all years between 221 BCE and 1911 CE. The data on prefecture borders is, unlike the information on county and prefecture seat locations, incomplete before 1350 CE, making the intersection of all years small. The minimum post 1350 CE shape denotes the space covered by prefectures in all years after 1350 CE.

In Figure C-6 the baseline shape, i.e. the areas that ever belonged to a prefecture observed in CHGIS are denoted in red, those that belonged to a prefecture in all years between 221 BCE and 1911 CE in green, and those that belonged to a prefecture in all years between 1350 CE and 1911 CE in purple. The green territories are primarily small because the data on prefecture borders is incomplete prior to 1350 CE.[4] Table C-30 and Table C-31 print summary statistics for these two alternative spatial extents.

[4]County and prefecture seat locations are not plagued by these missing data issues.

Table C-30: Geography Summary Statistics, Minimum Empire Extent

	Mean	St. Dev.	Min	Max
Distance from Coast	257.056	211.459	3.132	786.076
Distance from River	194.002	112.378	5.708	472.203
Distance from Equator	3,079.269	208.066	2,686.042	3,544.824
Elevation	333.737	248.221	0.377	1,189.212
Ruggedness	199,672.159	124,353.471	1,669.436	554,856.250
Temperature	16.573	1.338	12.298	20.974
Precipitation	1,506.589	205.427	991.662	2,022.636

Notes: Distances are in km, temperature in °C, precipitation in mm per year, elevation in meters, ruggedness index in millimeters as defined by Nunn and Puga (2012). Values refer to the Chinese empire's minimum shape with 693 pixels 21.99 x 28.53 km in size. Landform, dominant soil type, and lithology are categorical variables and summarised in Online Appendix B. See Table B-1 for details on variable generation. Variables are differently scaled in subsequent chapters to facilitate readability.

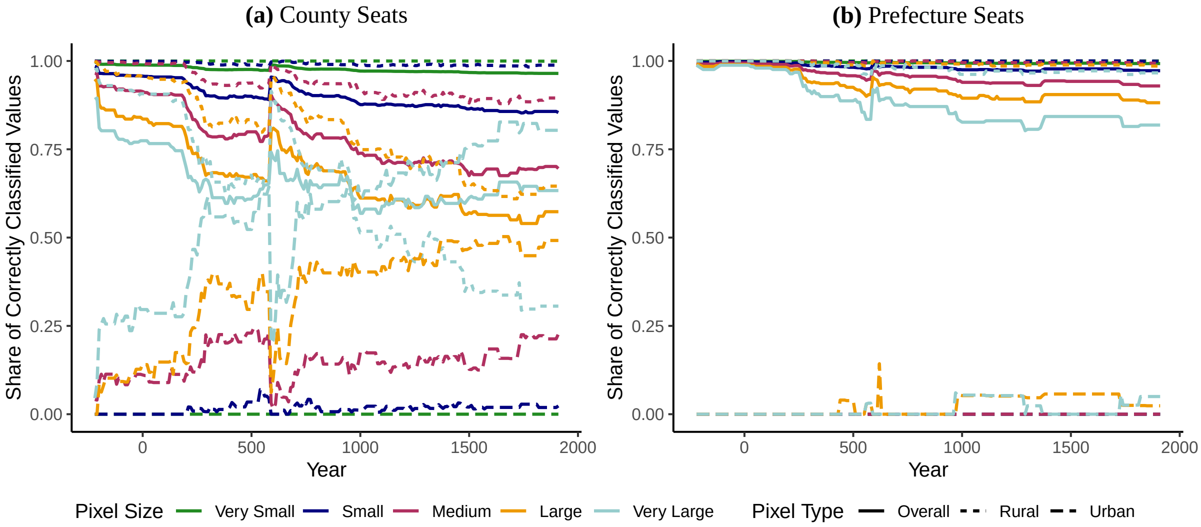
Table C-31: Geography Summary Statistics, Minimum Post 1350 CE Empire Extent

	Mean	St. Dev.	Min	Max
Distance from Coast	518.248	361.798	0.994	1,728.039
Distance from River	117.466	106.216	4.308	686.657
Distance from Equator	3,319.660	559.107	2,029.700	4,490.005
Elevation	756.759	771.196	0.055	4,112.595
Ruggedness	199,764.732	162,492.602	1,669.436	962,372.312
Temperature	14.963	4.392	-8.925	24.982
Precipitation	1,085.605	461.715	96.371	3,762.033

Notes: distances in km, temperature in °C, precipitation in mm per year, elevation in meters, ruggedness index in millimeters as defined by Nunn and Puga (2012). Values refer to the Chinese empire's minimum post 1350 CE shape with 5,345 pixels 21.99 x 28.53 km in size. Landform, dominant soil type, and lithology are categorical variables and summarised in Online Appendix B. See Table B-1 for details on variable generation. Variables are differently scaled in subsequent chapters to facilitate readability.

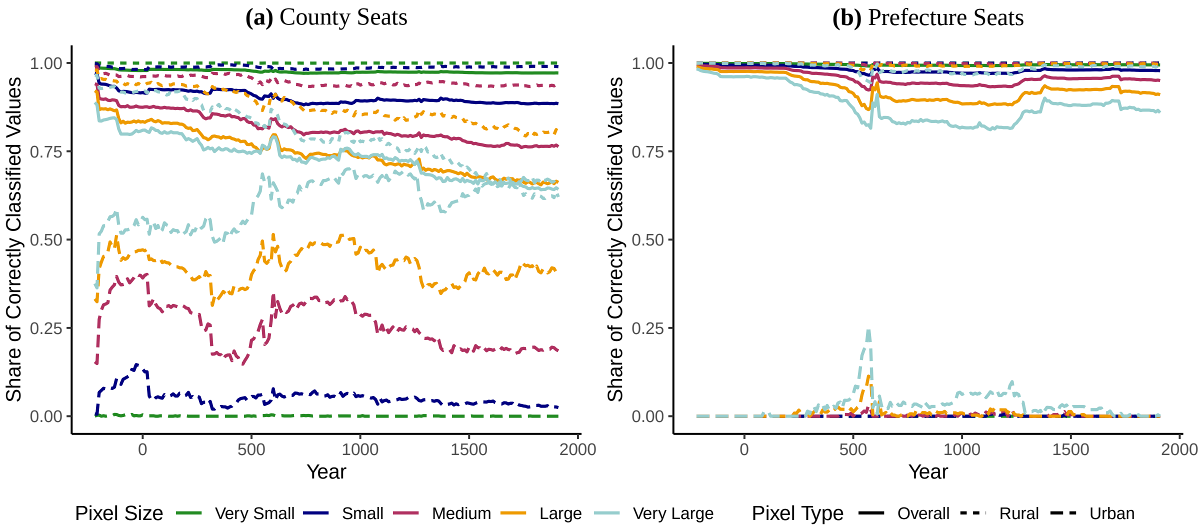
Figure C-7 and Figure C-8 replicate Figure C-3 with the two alternative empire shapes. The post 1350 CE minimum extent mostly excludes the little urbanised periphery and, thus, produces results that are very similar to the baseline outcomes - underlining their robustness. The overall minimum shape, in contrast, omits most cities from the estimations and generates strongly volatile results. These patterns may point to both, the smaller number of observations used in the algorithm and the threat of omitting regions that are crucial to the spatial organisation. Extracting adjusted R^2 and F statistics from OLS regressions supports these findings.

Figure C-7: Classification Random Forest Results (Minimum Empire Extent)



Notes: This figure repeats Figure C-3 with the minimum empire extent depicted in Figure C-6. Grid cells are predicted to be urban, if the probability is at least 0.5. Otherwise they are classified as rural.

Figure C-8: Classification Random Forest Results (Minimum Post 1350 CE Empire Extent)



Notes: This figure repeats Figure C-3 with the minimum post 1350 CE empire extent depicted in Figure C-6. Grid cells are predicted to be urban, if the probability is at least 0.5. Otherwise they are classified as rural.

We do not split the empire into the physiographic macroregions defined by Skinner (1977a) and Skinner (1977b). As von Glahn (2016) criticises, those macroregions only began to resemble actual conditions in imperial China after the crises of the 19th century. We, therefore, stick to our model of an empire-wide process that inter alia reflects the need to transport resources from the interior to the frontier and the disparities between the Yellow River and the Yangzi (Mostern, 2011).

C.9 Robustness Check: Agricultural Suitability Data

In the baseline estimation in Section 4.1, we understand agricultural productivity to be a function of our geographical regressors, such as a temperature, precipitation, distance from the equator. Our empirical analysis evaluates cross-sections separately, allowing the role of geographic characteristics for agricultural suitability to evolve over time, accounting e.g. for technological progress. We do not directly include agricultural productivity in the baseline estimation for two reasons: (i) the potentially strong collinearity with the geographical regressors, (ii) the lack of valid historical data on agricultural productivity, as it is known to change over time.

What we can, nevertheless, do to ensure robustness is to run a regression with modern agricultural data we obtain from the Food and Agriculture Organization of the United Nations and International Institute for Applied Systems Analysis (2015). In particular, we include control variables on agricultural suitability as indices on barley, dryland rice, foxtail millet, soybeans, wetland rice, and wheat between 1961 to 1990, the earliest available period in that data set. The indices, rescaled to range from 0 to 1, refer to the share of a grid cell with at least marginal suitability to growing the crop without CO_2 fertilisation. Table C-32 and Table C-33 contain the regression results, when we repeat our baseline regression and include these agricultural suitability controls. We can see that they barely add to the explanatory power, accompanied by only minor variation in coefficient significance. Even though this is modern and not historical agricultural data, we view this as evidence that the geographical variables already capture a lot of informational value for agricultural suitability, so that their additional inclusion hardly changes the outcome. This makes us confident that the lack of historical agricultural productivity does not constitute an omitted variable bias.

Table C-32: Local Geography Agricultural Suitability County Seat Regressions

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	−0.59***	−0.91***	−0.74***	−0.89***	−1.16***	−1.23***	−0.82***	−0.81***	−0.86***	−0.97***	−0.95***
	(0.21)	(0.25)	(0.25)	(0.27)	(0.29)	(0.33)	(0.30)	(0.27)	(0.26)	(0.27)	(0.27)
Dist. Coast	0.10	0.04	0.36*	0.65***	1.11***	1.13***	1.07***	0.91***	0.39*	0.58***	0.43**
	(0.19)	(0.22)	(0.20)	(0.21)	(0.23)	(0.26)	(0.26)	(0.24)	(0.20)	(0.21)	(0.21)
Dist. River	0.32	0.24	0.19	−0.38	−1.05**	−0.34	0.10	0.24	−0.14	−0.34	−0.03
	(0.36)	(0.43)	(0.40)	(0.40)	(0.44)	(0.52)	(0.53)	(0.52)	(0.50)	(0.51)	(0.48)
Ruggedness	−0.35*	−0.44	−0.41	−0.91***	−0.12	−0.37	−0.58*	−0.64**	−0.71**	−0.74**	−0.99***
	(0.20)	(0.31)	(0.29)	(0.25)	(0.38)	(0.38)	(0.35)	(0.32)	(0.34)	(0.35)	(0.32)
Temperature	0.27	0.32	0.25	0.04	−0.27	−0.73**	−0.16	−0.05	0.03	0.03	0.00
	(0.20)	(0.26)	(0.25)	(0.24)	(0.28)	(0.31)	(0.27)	(0.25)	(0.24)	(0.25)	(0.25)
Temperature^2	1.19	2.35*	1.49	−0.05	5.53***	6.78***	4.02***	2.99**	1.67	1.41	1.58
	(0.97)	(1.31)	(1.22)	(1.19)	(1.31)	(1.57)	(1.41)	(1.36)	(1.35)	(1.37)	(1.33)
Precipitation	−1.26***	−1.83***	−0.98***	−0.35	−1.56***	−0.52	0.05	0.14	−0.36	−0.14	0.11
	(0.34)	(0.40)	(0.35)	(0.38)	(0.51)	(0.54)	(0.47)	(0.44)	(0.41)	(0.44)	(0.43)
Precipitation^2	2.00***	2.90***	1.41**	0.45	2.43**	0.41	−0.55	−0.79	0.17	−0.40	−0.90
	(0.66)	(0.78)	(0.69)	(0.77)	(1.04)	(1.08)	(0.96)	(0.91)	(0.83)	(0.88)	(0.86)
Elevation	−2.07**	−2.05*	−2.05*	−3.70***	−4.89***	−4.87***	−4.06***	−3.90***	−2.84**	−3.48***	−3.37***
	(1.00)	(1.17)	(1.12)	(1.18)	(1.28)	(1.42)	(1.29)	(1.22)	(1.16)	(1.26)	(1.23)
Barley	0.25***	0.22***	0.15**	0.09	0.31***	0.23**	0.21**	0.08	0.17***	0.16**	0.19***
	(0.04)	(0.07)	(0.07)	(0.07)	(0.06)	(0.11)	(0.09)	(0.09)	(0.06)	(0.07)	(0.07)
Dryland Rice	0.10**	0.05	0.00	−0.06	−0.08	−0.11	−0.02	−0.18	0.05	0.00	−0.03
	(0.04)	(0.08)	(0.08)	(0.08)	(0.06)	(0.13)	(0.12)	(0.11)	(0.09)	(0.11)	(0.11)
Foxtail Millet	0.12***	0.16***	0.14***	0.09***	0.12***	0.09***	0.10***	0.08***	0.08***	0.08***	0.07***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Soybean	−0.14***	−0.19***	−0.16***	−0.10***	−0.09***	−0.03	−0.07***	−0.07***	−0.08***	−0.07**	−0.06**
	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Wetland Rice	−0.12***	−0.19***	−0.12***	−0.07	−0.12***	−0.04	−0.01	−0.01	0.02	0.02	−0.03
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Wheat	−0.12***	−0.04	0.01	0.02	−0.17***	−0.11	−0.10	0.03	−0.07	−0.07	−0.10
	(0.03)	(0.06)	(0.06)	(0.06)	(0.06)	(0.12)	(0.10)	(0.09)	(0.06)	(0.07)	(0.07)
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.19	0.23	0.19	0.12	0.21	0.19	0.18	0.17	0.15	0.15	0.15
F Stat.	14.55	18.45	14.62	9.42	16.83	15.36	14.49	13.52	11.23	11.61	11.23
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the county seats. I.e. the dependent variable is an indicator that equals one if a pixel hosts a county seat, and zero otherwise. Observations refer to the baseline pixels. Conley standard errors using a 150 km radius and a Bartlett kernel are in parentheses (***$p < 0.01$, **$p < 0.05$, *$p < 0.1$). Distances in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects. Control variables on agricultural suitability include suitability indices on barley, dryland rice, foxtail millet, soybeans, wetland rice, and wheat in the period 1961 to 1990. The indices, rescaled to range from 0 to 1, refer to the share of a grid cell with at least marginal suitability to growing the crop without CO_2 fertilisation.

Table C-33: Local Geography Agricultural Suitability Prefecture Seat Regressions

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	-0.10^{**}	-0.15^{***}	-0.08	-0.36^{***}	-0.38^{***}	-0.51^{***}	-0.43^{***}	-0.64^{***}	-0.34^{***}	-0.35^{***}	-0.32^{***}
	(0.05)	(0.05)	(0.06)	(0.11)	(0.10)	(0.14)	(0.15)	(0.14)	(0.12)	(0.12)	(0.11)
Dist. Coast	0.05	0.08^{*}	0.12^{**}	0.32^{***}	0.47^{***}	0.32^{***}	0.28^{***}	0.41^{***}	0.21^{***}	0.21^{***}	0.15^{**}
	(0.04)	(0.04)	(0.05)	(0.08)	(0.07)	(0.09)	(0.11)	(0.10)	(0.08)	(0.08)	(0.07)
Dist. River	-0.01	-0.01	-0.01	-0.13	-0.25^{*}	-0.34^{**}	-0.18	-0.28	-0.10	-0.14	0.01
	(0.07)	(0.09)	(0.10)	(0.14)	(0.13)	(0.17)	(0.20)	(0.20)	(0.17)	(0.16)	(0.15)
Ruggedness	0.01	-0.08	-0.10	-0.28^{***}	-0.00	-0.22	-0.18	-0.14	-0.51^{***}	-0.48^{***}	-0.24
	(0.05)	(0.06)	(0.06)	(0.10)	(0.13)	(0.15)	(0.18)	(0.18)	(0.17)	(0.17)	(0.17)
Temperature	-0.06	-0.00	0.09	-0.11	-0.10	-0.37^{***}	-0.19	-0.30^{**}	-0.06	-0.03	-0.10
	(0.04)	(0.05)	(0.05)	(0.10)	(0.09)	(0.12)	(0.14)	(0.13)	(0.12)	(0.11)	(0.10)
Temperature^2	0.17	0.02	0.07	0.31	0.49	1.95^{***}	0.93	0.38	-0.24	-0.13	0.27
	(0.18)	(0.23)	(0.24)	(0.36)	(0.39)	(0.55)	(0.57)	(0.57)	(0.52)	(0.54)	(0.47)
Precipitation	-0.05	-0.15^{**}	-0.14	-0.19	-0.20	-0.31	-0.36^{*}	-0.37^{*}	0.18	-0.01	-0.19
	(0.05)	(0.07)	(0.09)	(0.13)	(0.15)	(0.20)	(0.20)	(0.20)	(0.18)	(0.18)	(0.15)
Precipitation^2	0.06	0.25^{*}	0.21	0.33	0.32	0.53	0.62	0.68^{*}	-0.48	-0.13	0.30
	(0.11)	(0.15)	(0.19)	(0.26)	(0.30)	(0.41)	(0.41)	(0.40)	(0.40)	(0.38)	(0.31)
Elevation	-0.37^{*}	-0.40	0.03	-1.09^{**}	-1.71^{***}	-1.31^{**}	-1.17^{*}	-2.23^{***}	-0.59	-0.57	-0.69
	(0.22)	(0.24)	(0.27)	(0.46)	(0.48)	(0.58)	(0.67)	(0.68)	(0.59)	(0.59)	(0.53)
Barley	0.01^{***}	0.01	0.02	0.03	-0.01	-0.01	-0.00	0.01	-0.13^{**}	-0.14^{***}	-0.08^{*}
	(0.00)	(0.02)	(0.02)	(0.03)	(0.04)	(0.05)	(0.05)	(0.04)	(0.06)	(0.05)	(0.05)
Dryland Rice	0.00	-0.01	0.00	-0.01	-0.06	-0.06	-0.02	0.00	-0.10	-0.13^{**}	-0.08
	(0.01)	(0.02)	(0.02)	(0.04)	(0.04)	(0.07)	(0.07)	(0.05)	(0.07)	(0.06)	(0.06)
Foxtail Millet	0.00	0.00	0.01	0.01	0.02^{**}	0.01	0.02	0.00	-0.02^{**}	-0.03^{**}	-0.00
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Soybean	-0.00	-0.00	-0.00	-0.02	-0.01	0.00	-0.01	-0.00	0.00	0.01	0.01
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Wetland Rice	-0.01^{**}	-0.02^{***}	-0.03^{***}	-0.04^{***}	-0.01	-0.01	0.01	0.00	-0.01	-0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Wheat	-0.00	0.00	-0.01	0.01	0.04	0.05	0.04	0.03	0.15^{***}	0.16^{***}	0.10^{**}
	(0.00)	(0.02)	(0.02)	(0.02)	(0.03)	(0.06)	(0.05)	(0.04)	(0.06)	(0.05)	(0.05)
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.00	0.01	0.03	0.03	0.04	0.04	0.05	0.05	0.03	0.03	0.03
F Stat.	1.29	1.90	2.61	2.92	3.64	3.47	3.93	3.97	2.79	2.83	2.99
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

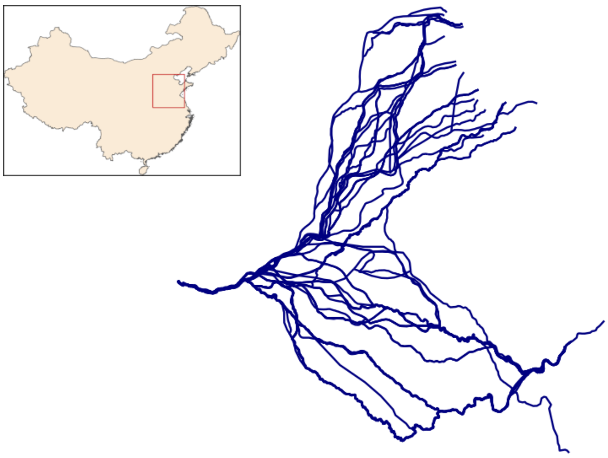
Notes: The table reports the regression results of eq. (14) using the prefecture seats. I.e. the dependent variable is an indicator that equals one if a pixel hosts a prefecture seat, and zero otherwise. Observations refer to the baseline pixels. Conley standard errors using a 150 km radius and a Bartlett kernel are in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects. Control variables on agricultural suitability include suitability indices on barley, dryland rice, foxtail millet, soybeans, wetland rice, and wheat in the period 1961 to 1990. The indices, rescaled to range from 0 to 1, refer to the share of a grid cell with at least marginal suitability to growing the crop without CO_2 fertilisation.

C.10 Robustness Check: Paleo Data

Our baseline model uses modern geography data. In contrast to paleoclimatic data, it has the advantages of being available for many variables and of being accurately identified at high resolutions. Nonetheless, some geographic factors changed over time. In this section, we illustrate that using paleoclimatic data does not change the results. The adjustments that we account for are the changes in Yellow River's lower path and developments in temperature and precipitation.[5]

Figure C-9 shows how the Yellow River changed its course over time. The data comes from Chen et al. (2012) and Chen et al. (2015). The many path changes and floods were largely a result of deforestation causing soil erosion and consequently increased sediment uptake in the river (Elvin, 2004). Bursting dikes on 1593 occasions, many floods had a devastating impact on the population and the institutional setting. To name a few examples: a flood in 2 CE directly killed tens of thousands of people and triggered starvation and disease with an even higher death toll (Major and Cook, 2017); a few years later, in 11 CE, again thousands died from a flood that resulted in mass migration, famine, emerging bandit gangs killing county officials and forming armies, and turned out to be a key reason for the fall of the Xin dynasty (Major and Cook, 2017); a flood with a particularly high death toll was the one in 1117 CE that killed over a million people (Elvin, 2004; Tuotuo, 1346).

Figure C-9: All Observed Lower Yellow River Paths



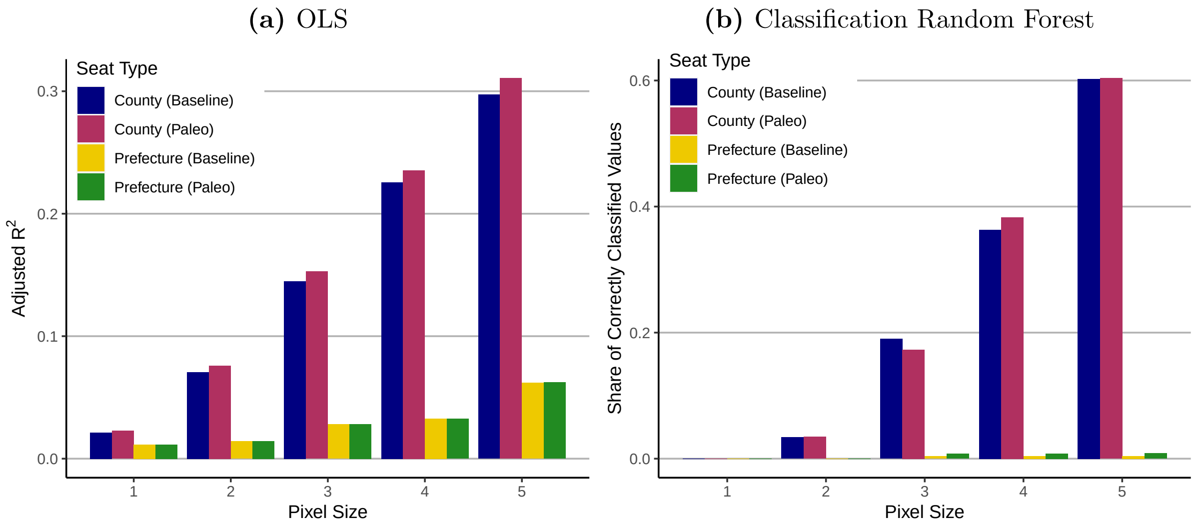
Notes: The paths of the lower Yellow River according to Chen et al. (2012) and Chen et al. (2015). The red shape marks the extent of the plotted river segments on a map of modern China (Center for Spatial Science, University of California, Davis, 2018).

We use multiproxy warm season temperature reconstructions by Zhang et al. (2018),

[5] The baseline data on temperature and precipitation refers to observations between 1900 and 1950 and should be less affected by man-made climate change than 21st century data. Nonetheless, it does not account for the historic fluctuations in climate.

which is published at a 5 x 5 degree resolution in decadal intervals, as information on historic temperatures. The precipitation data is reconstructed summer precipitation by Shi et al. (2018) published as 2 x 2 degree pixels at an annual level. With both, a 5 x 5 degree and a 2 x 2 degree resolution, the grid cells are so large that the whole empire just contains a few pixels. To increase the sample size to a level that can meaningfully be used in estimations, we disaggregate cells to the baseline resolution via bilinear interpolation. The severe measurement error this introduces and the fact that the temperature is measured in anomalies rather than levels are the reasons why we add the paleoclimatic data on temperature and precipitation as additional variables, without replacing the modern counterparts.[6]

Figure C-10: Local Geography Estimations in 1500 CE Using Paleo Data



Notes: Grid cells are predicted to be urban, if the probability is at least 0.5. Otherwise they are classified as rural. The right figure displays the share of urban pixels correctly classified as urban. Prefecture seats and county seats in the baseline and the paleo setting are evaluated in separate estimations. The paleo data differs from the baseline data in that it accounts for the changing path of the lower Yellow River and historic climatic conditions in terms of temperature and precipitation. Aggregations one to five refer to very small, small, medium (baseline), large, and very large pixels respectively.

Figure C-10 compares results with the baseline outcomes. It turns out that accounting for geographic changes barely affects the estimates, underlining the robustness of our baseline results.

[6]Climatic shocks had marked effects on the empire, often setting off a series of events. Climatic disturbances and cold weather led to harvest failures between 1638 and 1642 CE and the Chinese population size shrank by 20% in response to the ensuing natural disasters, wars, and epidemics (von Glahn, 2016; Atwell, 1986, 1990; Marks, 2012). Many dynasties set up grain reserves to counter famines resulting from geographic shocks, such as floods and changes in climate, which were such a severe threat that they could lead to the downfall of a dynasty (Mostern, 2011; von Glahn, 2016). Even the Qing dynasty that intervened comparatively little in other regards was heavily invested in a system of grain reserves (von Glahn, 2016; Will, 1990; Will and Wong, 1991). Despite the effort that was put into this system, the Qing's granary reserves were depleted after the Tambora volcanic eruption induced a series of failed harvests at a global scale (von Glahn, 2016). Climatic changes were without doubt important. However, as long as there is only data on large macro regions, the incentives to relocate administrative cities between more fine-grain locations are not traceable.

C.11 Robustness Check: Man-Made Infrastructure

All of the baseline model's explanatory variables are exogenous geographic factors. We omit any man-made transport infrastructure given its potential endogeneity. Does the imperial administration build the town because of a nearby (planned) road or does it build the road because of the settlement? Potential biases from reverse or simultaneous causality may impede the identification strategy. Nonetheless, there is an important piece of transport infrastructure that played a vital role in various dynasties and needs to be mentioned in the context of this paper: the Grand Canal.

The Grand Canal was not one persistent water way once built and unmodified throughout imperial times. It was a series of segments with early ones dating back to the Spring and Autumn period (770 - 476 BCE), long before the imperial period starting with the unification under the Qin dynasty in 221 BCE. Over the following millennia, dynasties added a number of extensions and modifications (Tan et al., 2019; Porter, 2016; Mostern, 2011; Wilkinson, 2013).

While in the baseline regression, we include *distance to rivers* as a regressor, here, we look at *distance to waterways* instead. This comprises both natural rivers and the Grand Canal. As the construction of the Grand Canal changed over time, we look at three different points in time with three different shapefiles of the Grand Canal: the Sui, Song, and Yuan dynasties (Bol, 2021c,a,b). In line with the baseline specification, Table C-34 and Table C-35 first repeat the cross-sectional estimations for county and prefecture seats with all regressors assumed to be time-constant - incl. the Grand Canal. The outcomes are similar to Table 2 and Table 3.

In Table C-36 to Table C-41, we zoom in on the three dynasties for which we have the geocoded canal paths. These tables again make use of the mentioned distance to waterways, but using the dynasty-specific rather than all recorded canal segments. For each dynasty, we run three cross-sectional regressions: one around a decade into the dynasty, one around a decade before the end of the dynasty, and one around a decade after its fall. Facilitating comparisons, the right three columns report respective baseline results using rivers without canals. The impact on other covariates' coefficient estimates is small. And even the *distance from waterway* effect is usually not significantly different from the pure river effect when accounting for standard errors.

Another important piece of man-made infrastructure that we add to our analysis are courier routes from the Ming dynasty provided by Berman and Zhang (2017). Courier routes are even more endogenous than the Grand Canal, as artificial waterways are somewhat restricted by the topology - given that it is costly and technologically demanding to allow water to pass through upward sloping terrain. Table C-42 and Table C-43 show that distance to courier routes has a significantly negative impact on both prefecture and county seats, suggesting that man-made infrastructure did play a role. However, the coefficients on the other, exogeneous, geographical regressors and the goodness of fit results do not change by a lot, so that our main results remain valid in the presence of man-made infrastructure. Without overcoming the endogeneity concern, any of these findings should be taken with a grain of salt.[7]

Table C-34: Local Geography County Seat Regressions with the Grand Canal

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	-0.31^{*}	-0.45^{*}	-0.46^{*}	-0.74^{***}	-0.99^{***}	-1.30^{***}	-0.87^{***}	-0.82^{***}	-0.92^{***}	-1.07^{***}	-0.94^{***}
	(0.18)	(0.25)	(0.24)	(0.25)	(0.29)	(0.32)	(0.27)	(0.25)	(0.25)	(0.26)	(0.26)
Dist. Coast	-0.12	-0.30	0.08	0.48^{***}	0.84^{***}	0.89^{***}	0.91^{***}	0.77^{***}	0.32^{*}	0.52^{***}	0.32^{*}
	(0.16)	(0.21)	(0.18)	(0.17)	(0.22)	(0.24)	(0.23)	(0.22)	(0.18)	(0.19)	(0.18)
Dist. Waterway	-0.99^{***}	-1.21^{***}	-1.26^{***}	-1.10^{***}	-1.96^{***}	-1.64^{***}	-1.11^{**}	-0.97^{*}	-1.23^{***}	-1.50^{***}	-1.14^{**}
	(0.31)	(0.40)	(0.36)	(0.36)	(0.46)	(0.50)	(0.52)	(0.50)	(0.47)	(0.47)	(0.44)
Ruggedness	-0.50^{***}	-0.57^{*}	-0.77^{***}	-1.28^{***}	-0.67^{*}	-1.21^{***}	-1.36^{***}	-1.33^{***}	-1.32^{***}	-1.42^{***}	-1.43^{***}
	(0.19)	(0.31)	(0.30)	(0.27)	(0.35)	(0.39)	(0.39)	(0.37)	(0.36)	(0.38)	(0.35)
Temperature	0.49^{***}	0.71^{***}	0.51^{**}	0.22	0.04	-0.58^{*}	-0.08	0.02	-0.01	-0.02	0.05
	(0.19)	(0.26)	(0.24)	(0.23)	(0.29)	(0.31)	(0.26)	(0.25)	(0.24)	(0.25)	(0.26)
Temperature^2	-0.08	0.53	0.04	-1.31	2.54^{**}	4.12^{***}	2.32^{*}	1.50	1.09	0.59	0.35
	(0.78)	(1.08)	(0.99)	(0.98)	(1.14)	(1.36)	(1.24)	(1.21)	(1.19)	(1.24)	(1.19)
Precipitation	-1.31^{***}	-2.02^{***}	-1.00^{***}	-0.22	-1.37^{***}	-0.27	0.43	0.47	0.02	0.27	0.32
	(0.37)	(0.43)	(0.36)	(0.37)	(0.52)	(0.56)	(0.52)	(0.48)	(0.42)	(0.46)	(0.43)
Precipitation^2	2.31^{***}	3.53^{***}	1.61^{**}	0.29	2.23^{**}	0.04	-1.22	-1.32	-0.55	-1.17	-1.18
	(0.74)	(0.87)	(0.72)	(0.75)	(1.08)	(1.14)	(1.08)	(1.00)	(0.86)	(0.92)	(0.87)
Elevation	-2.24^{***}	-1.64	-2.32^{**}	-4.28^{***}	-6.63^{***}	-7.59^{***}	-6.30^{***}	-5.63^{***}	-4.49^{***}	-5.46^{***}	-4.81^{***}
	(0.87)	(1.14)	(1.08)	(1.11)	(1.29)	(1.39)	(1.23)	(1.17)	(1.15)	(1.24)	(1.19)
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.17	0.20	0.17	0.12	0.19	0.19	0.18	0.17	0.14	0.15	0.14
F Stat.	13.48	16.77	13.71	9.18	15.93	15.12	14.33	13.42	11.30	11.72	11.34
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the county seats. The dependent variable is an indicator that equals one, if the pixel hosts a county seat in that year, and zero otherwise. The estimations use the baseline pixels. Conley standard errors using a 150 km radius and a Bartlett kernel are in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. The distance to the nearest waterway accounts for the baseline rivers and the Grand Canal. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

[7] We could introduce instruments in the vein of Donaldson and Hornbeck (2016) or Jedwab and Storeygard (2022), trying to derive exogenous effects from endogenous infrastructure. However, any such analysis doing justice to the historical context would add tens of pages to this appendix. In keeping the already long appendix in a reader-friendly format, we defer such study to another paper.

Table C-35: Local Geography Prefecture Seat Regressions with the Grand Canal

	200 BCE	1 CE	200 CE	400 CE	600 CE	800 CE	1000 CE	1200 CE	1400 CE	1600 CE	1800 CE
Dist. Equator	−0.08**	−0.11**	−0.05	−0.28***	−0.37***	−0.55***	−0.49***	−0.69***	−0.32***	−0.33***	−0.36***
	(0.04)	(0.05)	(0.06)	(0.10)	(0.09)	(0.14)	(0.15)	(0.14)	(0.12)	(0.12)	(0.10)
Dist. Coast	0.02	0.04	0.06	0.23***	0.41***	0.24***	0.23**	0.35***	0.16**	0.17**	0.10
	(0.03)	(0.04)	(0.05)	(0.07)	(0.07)	(0.08)	(0.10)	(0.09)	(0.07)	(0.07)	(0.07)
Dist. Waterway	−0.10	−0.15*	−0.23**	−0.38***	−0.44***	−0.69***	−0.65***	−0.68***	−0.26	−0.30*	−0.20
	(0.07)	(0.08)	(0.09)	(0.14)	(0.13)	(0.16)	(0.17)	(0.17)	(0.17)	(0.17)	(0.15)
Ruggedness	0.00	−0.05	−0.12*	−0.32***	−0.18	−0.50***	−0.47**	−0.39**	−0.51***	−0.47***	−0.41**
	(0.05)	(0.06)	(0.07)	(0.11)	(0.12)	(0.17)	(0.19)	(0.18)	(0.17)	(0.18)	(0.18)
Temperature	−0.04	0.03	0.15***	−0.05	−0.06	−0.34***	−0.21	−0.32**	−0.07	−0.04	−0.09
	(0.04)	(0.05)	(0.05)	(0.09)	(0.09)	(0.12)	(0.14)	(0.13)	(0.12)	(0.12)	(0.10)
Temperature^2	0.00	−0.20	−0.32	−0.07	0.02	1.33***	0.70	0.17	−0.04	0.02	0.08
	(0.15)	(0.19)	(0.20)	(0.31)	(0.33)	(0.49)	(0.52)	(0.52)	(0.48)	(0.48)	(0.42)
Precipitation	−0.09*	−0.24***	−0.24**	−0.29**	−0.16	−0.32	−0.30	−0.35*	0.04	−0.17	−0.23
	(0.05)	(0.07)	(0.09)	(0.13)	(0.14)	(0.20)	(0.20)	(0.20)	(0.17)	(0.17)	(0.15)
Precipitation^2	0.16	0.45***	0.45**	0.56**	0.26	0.54	0.50	0.63	−0.20	0.21	0.35
	(0.10)	(0.15)	(0.20)	(0.27)	(0.28)	(0.39)	(0.41)	(0.42)	(0.35)	(0.34)	(0.30)
Elevation	−0.39**	−0.30	0.01	−1.03**	−2.03***	−1.91***	−1.80***	−2.77***	−0.15	−0.14	−0.89*
	(0.19)	(0.23)	(0.25)	(0.42)	(0.45)	(0.58)	(0.67)	(0.67)	(0.60)	(0.60)	(0.52)
Soil	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.00	0.01	0.02	0.03	0.04	0.04	0.05	0.05	0.03	0.03	0.03
F Stat.	1.29	1.87	2.58	2.86	3.67	3.52	4.03	4.09	2.73	2.77	3.05
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the prefecture seats. The dependent variable is an indicator that equals one, if the pixel hosts a prefecture seat in that year, and zero otherwise. The estimations use the baseline pixels. Conley standard errors using a 150 km radius and a Bartlett kernel are in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. The distance to the nearest waterway accounts for the baseline rivers and the Grand Canal. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-36: Local Geography County Seat Regressions Targeting the Sui Dynasty

	Waterways incl. Grand Canal			Baseline		
	590 CE	610 CE	620 CE	590 CE	610 CE	620 CE
Dist. Equator	−1.00***	−0.77***	−0.90***	−0.95***	−0.72***	−0.85***
	(0.26)	(0.27)	(0.30)	(0.26)	(0.27)	(0.30)
Dist. Coast	1.02***	0.89***	1.04***	0.99***	0.85***	1.00***
	(0.20)	(0.21)	(0.23)	(0.20)	(0.21)	(0.23)
Dist. Waterway	−1.77***	−1.58***	−1.95***			
	(0.41)	(0.40)	(0.44)			
Dist. River				−1.42***	−1.16***	−1.60***
				(0.41)	(0.41)	(0.45)
Ruggedness	−0.63*	−0.60*	−0.73**	−0.63*	−0.60*	−0.72**
	(0.32)	(0.33)	(0.36)	(0.32)	(0.33)	(0.36)
Temperature	−0.10	0.16	0.17	−0.03	0.23	0.24
	(0.26)	(0.27)	(0.29)	(0.26)	(0.27)	(0.29)
Temperature^2	2.14**	2.06**	1.67	2.04**	1.96*	1.58
	(1.02)	(1.04)	(1.12)	(1.03)	(1.05)	(1.13)
Precipitation	−0.82*	−0.86*	−0.90*	−0.78*	−0.81*	−0.86*
	(0.45)	(0.48)	(0.52)	(0.45)	(0.48)	(0.52)
Precipitation^2	1.15	1.17	1.16	1.01	1.00	1.02
	(0.93)	(0.98)	(1.06)	(0.92)	(0.98)	(1.07)
Elevation	−6.93***	−5.80***	−6.27***	−6.77***	−5.63***	−6.10***
	(1.19)	(1.18)	(1.37)	(1.19)	(1.17)	(1.37)
Soil	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.17	0.18	0.18	0.17	0.18	0.18
F Stat.	13.81	14.30	14.65	13.73	14.21	14.57
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the county seats. The dependent variable is an indicator that equals one, if the pixel hosts a county seat in that year, and zero otherwise. The estimations use the baseline pixels. Conley standard errors using a 150 km radius and a Bartlett kernel are in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. The distance to the nearest waterway accounts for the baseline rivers and the Grand Canal. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-37: Local Geography County Seat Regressions Targeting the Song Dynasty

	Waterways incl. Grand Canal			Baseline		
	970 CE	1270 CE	1290 CE	970 CE	1270 CE	1290 CE
Dist. Equator	−0.98***	−0.60**	−0.99***	−0.96***	−0.59**	−0.97***
	(0.31)	(0.26)	(0.27)	(0.31)	(0.26)	(0.27)
Dist. Coast	0.81***	0.57***	0.26	0.80***	0.57***	0.25
	(0.23)	(0.22)	(0.18)	(0.23)	(0.22)	(0.18)
Dist. Waterway	−1.04*	−0.19	−0.75			
	(0.53)	(0.55)	(0.51)			
Dist. River				−0.68	0.04	−0.47
				(0.56)	(0.56)	(0.53)
Ruggedness	−1.20***	−1.11***	−1.07***	−1.20***	−1.11***	−1.07***
	(0.40)	(0.36)	(0.37)	(0.40)	(0.36)	(0.37)
Temperature	−0.30	0.18	−0.06	−0.28	0.19	−0.04
	(0.30)	(0.26)	(0.25)	(0.30)	(0.26)	(0.25)
Temperature^2	4.09***	1.56	1.31	4.02***	1.53	1.26
	(1.36)	(1.19)	(1.21)	(1.38)	(1.20)	(1.22)
Precipitation	0.03	0.44	−0.25	0.10	0.47	−0.20
	(0.54)	(0.48)	(0.44)	(0.54)	(0.49)	(0.45)
Precipitation^2	−0.51	−1.39	−0.12	−0.70	−1.49	−0.26
	(1.09)	(1.02)	(0.90)	(1.12)	(1.04)	(0.93)
Elevation	−6.16***	−4.71***	−4.68***	−6.09***	−4.66***	−4.63***
	(1.32)	(1.20)	(1.20)	(1.32)	(1.20)	(1.20)
Soil	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.19	0.16	0.15	0.19	0.16	0.15
F Stat.	15.33	12.49	11.77	15.28	12.48	11.74
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the county seats. The dependent variable is an indicator that equals one, if the pixel hosts a county seat in that year, and zero otherwise. The estimations use the baseline pixels. Conley standard errors using a 150 km radius and a Bartlett kernel are in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. The distance to the nearest waterway accounts for the baseline rivers and the Grand Canal. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-38: Local Geography County Seat Regressions Targeting the Yuan Dynasty

	Waterways incl. Grand Canal			Baseline		
	1270 CE	1360 CE	1380 CE	1270 CE	1360 CE	1380 CE
Dist. Equator	-0.65^{**}	-0.99^{***}	-0.71^{***}	-0.59^{**}	-0.92^{***}	-0.63^{**}
	(0.26)	(0.27)	(0.25)	(0.26)	(0.27)	(0.25)
Dist. Coast	0.58^{***}	0.24	0.43^{**}	0.57^{***}	0.22	0.41^{**}
	(0.22)	(0.18)	(0.18)	(0.22)	(0.18)	(0.18)
Dist. Waterway	-0.75	-1.14^{**}	-1.16^{**}			
	(0.51)	(0.48)	(0.47)			
Dist. River				0.04	-0.36	-0.30
				(0.56)	(0.53)	(0.52)
Ruggedness	-1.12^{***}	-1.12^{***}	-1.28^{***}	-1.11^{***}	-1.12^{***}	-1.28^{***}
	(0.36)	(0.37)	(0.36)	(0.36)	(0.37)	(0.36)
Temperature	0.12	-0.08	0.19	0.19	-0.00	0.27
	(0.25)	(0.25)	(0.24)	(0.26)	(0.26)	(0.25)
Temperature^2	1.65	1.45	1.10	1.53	1.31	0.95
	(1.19)	(1.22)	(1.18)	(1.20)	(1.23)	(1.19)
Precipitation	0.35	-0.22	0.32	0.47	-0.10	0.45
	(0.48)	(0.44)	(0.43)	(0.49)	(0.45)	(0.44)
Precipitation^2	-1.15	-0.21	-1.26	-1.49	-0.56	-1.64^{*}
	(1.01)	(0.90)	(0.90)	(1.04)	(0.94)	(0.94)
Elevation	-4.89^{***}	-4.72^{***}	-3.85^{***}	-4.66^{***}	-4.46^{***}	-3.57^{***}
	(1.19)	(1.20)	(1.17)	(1.20)	(1.20)	(1.17)
Soil	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.16	0.15	0.14	0.16	0.15	0.14
F Stat.	12.52	11.86	11.26	12.48	11.78	11.17
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the county seats. The dependent variable is an indicator that equals one, if the pixel hosts a county seat in that year, and zero otherwise. The estimations use the baseline pixels. Conley standard errors using a 150 km radius and a Bartlett kernel are in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. The distance to the nearest waterway accounts for the baseline rivers and the Grand Canal. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-39: Local Geography Prefecture Seat Regressions Targeting the Sui Dynasty

	Waterways incl. Grand Canal			Baseline		
	590 CE	610 CE	620 CE	590 CE	610 CE	620 CE
Dist. Equator	-0.38^{***}	-0.31^{***}	-0.45^{***}	-0.36^{***}	-0.30^{***}	-0.43^{***}
	(0.09)	(0.08)	(0.12)	(0.09)	(0.08)	(0.12)
Dist. Coast	0.42^{***}	0.32^{***}	0.54^{***}	0.41^{***}	0.31^{***}	0.53^{***}
	(0.06)	(0.05)	(0.08)	(0.06)	(0.05)	(0.08)
Dist. Waterway	-0.53^{***}	-0.29^{***}	-0.53^{***}			
	(0.13)	(0.11)	(0.16)			
Dist. River				-0.45^{***}	-0.25^{**}	-0.45^{***}
				(0.13)	(0.11)	(0.17)
Ruggedness	-0.24^{*}	-0.20^{**}	-0.31^{**}	-0.23^{*}	-0.19^{**}	-0.30^{**}
	(0.13)	(0.09)	(0.14)	(0.13)	(0.09)	(0.14)
Temperature	-0.13	-0.13^{*}	-0.02	-0.11	-0.12	-0.00
	(0.09)	(0.08)	(0.10)	(0.09)	(0.08)	(0.11)
Temperature^2	0.23	0.08	-0.50	0.21	0.07	-0.52
	(0.35)	(0.29)	(0.39)	(0.35)	(0.29)	(0.39)
Precipitation	-0.09	-0.11	-0.24	-0.08	-0.11	-0.23
	(0.13)	(0.13)	(0.18)	(0.13)	(0.13)	(0.18)
Precipitation^2	0.14	0.23	0.38	0.11	0.21	0.35
	(0.27)	(0.26)	(0.37)	(0.27)	(0.26)	(0.37)
Elevation	-2.19^{***}	-1.49^{***}	-2.64^{***}	-2.14^{***}	-1.46^{***}	-2.60^{***}
	(0.44)	(0.38)	(0.54)	(0.44)	(0.38)	(0.54)
Soil	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.04	0.04	0.05	0.04	0.04	0.05
F Stat.	3.28	3.51	4.34	3.26	3.51	4.33
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the prefecture seats. The dependent variable is an indicator that equals one, if the pixel hosts a prefecture seat in that year, and zero otherwise. The estimations use the baseline pixels. Conley standard errors using a 150 km radius and a Bartlett kernel are in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index $\times$ 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. The distance to the nearest waterway accounts for the baseline rivers and the Grand Canal. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-40: Local Geography Prefecture Seat Regressions Targeting the Song Dynasty

	Waterways incl. Grand Canal			Baseline		
	970 CE	1270 CE	1290 CE	970 CE	1270 CE	1290 CE
Dist. Equator	-0.52^{***}	-0.47^{***}	-0.47^{***}	-0.52^{***}	-0.46^{***}	-0.47^{***}
	(0.15)	(0.12)	(0.11)	(0.15)	(0.12)	(0.11)
Dist. Coast	0.25^{***}	0.40^{***}	0.25^{***}	0.25^{***}	0.39^{***}	0.25^{***}
	(0.09)	(0.09)	(0.07)	(0.09)	(0.09)	(0.07)
Dist. Waterway	-0.43^{**}	-0.42^{**}	-0.11			
	(0.20)	(0.17)	(0.16)			
Dist. River				-0.39^{**}	-0.29^{*}	-0.08
				(0.19)	(0.17)	(0.16)
Ruggedness	-0.52^{***}	-0.35^{**}	-0.42^{***}	-0.52^{***}	-0.35^{**}	-0.42^{***}
	(0.19)	(0.15)	(0.15)	(0.19)	(0.15)	(0.15)
Temperature	-0.25^{*}	-0.25^{**}	-0.21^{**}	-0.24^{*}	-0.24^{**}	-0.21^{**}
	(0.14)	(0.10)	(0.11)	(0.14)	(0.10)	(0.11)
Temperature^2	0.85	0.51	0.21	0.84	0.48	0.20
	(0.52)	(0.46)	(0.46)	(0.53)	(0.47)	(0.46)
Precipitation	-0.31	0.18	0.04	-0.30	0.20	0.05
	(0.21)	(0.18)	(0.16)	(0.21)	(0.18)	(0.16)
Precipitation^2	0.52	-0.44	-0.15	0.49	-0.51	-0.17
	(0.42)	(0.38)	(0.33)	(0.42)	(0.39)	(0.33)
Elevation	-1.89^{***}	-2.02^{***}	-1.28^{**}	-1.89^{***}	-2.00^{***}	-1.27^{**}
	(0.68)	(0.56)	(0.55)	(0.68)	(0.56)	(0.55)
Soil	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.05	0.05	0.04	0.05	0.05	0.04
F Stat.	3.96	4.08	3.43	3.96	4.07	3.43
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the prefecture seats. The dependent variable is an indicator that equals one, if the pixel hosts a prefecture seat in that year, and zero otherwise. The estimations use the baseline pixels. Conley standard errors using a 150 km radius and a Bartlett kernel are in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. The distance to the nearest waterway accounts for the baseline rivers and the Grand Canal. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-41: Local Geography Prefecture Seat Regressions Targeting the Yuan Dynasty

	Waterways incl. Grand Canal			Baseline		
	1270 CE	1360 CE	1380 CE	1270 CE	1360 CE	1380 CE
Dist. Equator	-0.47^{***}	-0.54^{***}	-0.24^{**}	-0.46^{***}	-0.53^{***}	-0.22^{**}
	(0.12)	(0.11)	(0.11)	(0.12)	(0.11)	(0.11)
Dist. Coast	0.40^{***}	0.24^{***}	0.18^{***}	0.39^{***}	0.24^{***}	0.17^{***}
	(0.09)	(0.07)	(0.07)	(0.09)	(0.07)	(0.06)
Dist. Waterway	-0.39^{**}	-0.23	-0.26^{*}			
	(0.17)	(0.17)	(0.15)			
Ruggedness	-0.35^{**}	-0.37^{**}	-0.34^{**}	-0.35^{**}	-0.37^{**}	-0.34^{**}
	(0.15)	(0.17)	(0.16)	(0.15)	(0.17)	(0.16)
Temperature	-0.25^{**}	-0.25^{**}	-0.05	-0.24^{**}	-0.24^{**}	-0.03
	(0.10)	(0.11)	(0.11)	(0.10)	(0.11)	(0.11)
Temperature^2	0.51	0.27	0.07	0.48	0.26	0.05
	(0.46)	(0.47)	(0.42)	(0.47)	(0.47)	(0.42)
Precipitation	0.19	-0.09	0.08	0.20	-0.07	0.10
	(0.18)	(0.16)	(0.17)	(0.18)	(0.16)	(0.17)
Precipitation^2	-0.46	0.09	-0.25	-0.51	0.06	-0.30
	(0.39)	(0.33)	(0.35)	(0.39)	(0.34)	(0.36)
Elevation	-2.05^{***}	-1.31^{**}	-0.07	-2.00^{***}	-1.27^{**}	-0.03
	(0.56)	(0.54)	(0.56)	(0.56)	(0.54)	(0.56)
Dist. River				-0.29^{*}	-0.15	-0.15
				(0.17)	(0.17)	(0.15)
Soil	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.05	0.04	0.03	0.05	0.04	0.03
F Stat.	4.08	3.39	2.72	4.07	3.38	2.71
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the prefecture seats. The dependent variable is an indicator that equals one, if the pixel hosts a prefecture seat in that year, and zero otherwise. The estimations use the baseline pixels. Conley standard errors using a 150 km radius and a Bartlett kernel are in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index $\times$ 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. The distance to the nearest waterway accounts for the baseline rivers and the Grand Canal. Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-42: Local Geography County Seat Regressions with Ming Courier Routes

	Extension			Baseline		
	1400 CE	1500 CE	1600 CE	1400 CE	1500 CE	1600 CE
Dist. Equator	−0.61**	−0.73**	−0.72**	−0.85***	−0.97***	−0.99***
	(0.30)	(0.32)	(0.32)	(0.25)	(0.26)	(0.27)
Dist. Coast	0.31*	0.47**	0.51***	0.30*	0.46**	0.49***
	(0.18)	(0.19)	(0.19)	(0.18)	(0.19)	(0.19)
Dist. River	−0.50	−0.60	−0.74	−0.36	−0.46	−0.58
	(0.53)	(0.55)	(0.54)	(0.52)	(0.54)	(0.53)
Dist. Courier Routes	−0.58**	−0.60**	−0.67**			
	(0.25)	(0.26)	(0.26)			
Ruggedness	−1.29***	−1.30***	−1.39***	−1.32***	−1.32***	−1.42***
	(0.36)	(0.36)	(0.37)	(0.36)	(0.36)	(0.37)
Temperature	−0.04	−0.06	−0.05	0.07	0.05	0.07
	(0.23)	(0.23)	(0.24)	(0.24)	(0.25)	(0.26)
Temperature^2	1.90	1.41	1.54	0.94	0.42	0.42
	(1.19)	(1.19)	(1.21)	(1.21)	(1.24)	(1.26)
Precipitation	0.36	0.67	0.65	0.14	0.45	0.40
	(0.45)	(0.50)	(0.49)	(0.43)	(0.47)	(0.46)
Precipitation^2	−1.20	−1.95*	−1.87*	−0.92	−1.67*	−1.55
	(0.92)	(1.02)	(0.99)	(0.90)	(1.00)	(0.97)
Elevation	−3.55***	−4.22***	−4.38***	−4.23***	−4.92***	−5.17***
	(1.23)	(1.28)	(1.33)	(1.15)	(1.20)	(1.25)
Soil	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.14	0.14	0.15	0.14	0.14	0.15
F Stat.	11.18	11.24	11.58	11.20	11.26	11.60
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the county seats. The dependent variable is an indicator that equals one, if the pixel hosts a county seat in that year, and zero otherwise. The estimations use the baseline pixels. Conley standard errors using a 150 km radius and a Bartlett kernel are in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. The distance to the nearest courier routes refers to the Ming courier routes published by Berman and Zhang (2017). Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

Table C-43: Local Geography Prefecture Seat Regressions with Ming Courier Routes

	Extension			Baseline		
	1400 CE	1500 CE	1600 CE	1400 CE	1500 CE	1600 CE
Dist. Equator	−0.22	−0.22	−0.26*	−0.31**	−0.31**	−0.32***
	(0.14)	(0.15)	(0.14)	(0.12)	(0.13)	(0.12)
Dist. Coast	0.16**	0.18***	0.17**	0.16**	0.18**	0.16**
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Dist. River	−0.20	−0.23	−0.22	−0.15	−0.18	−0.18
	(0.16)	(0.17)	(0.17)	(0.16)	(0.17)	(0.16)
Dist. Courier Routes	−0.21**	−0.22**	−0.16*			
	(0.08)	(0.09)	(0.09)			
Ruggedness	−0.50***	−0.47***	−0.46***	−0.51***	−0.48***	−0.47***
	(0.17)	(0.18)	(0.18)	(0.17)	(0.18)	(0.18)
Temperature	−0.10	−0.04	−0.06	−0.06	0.01	−0.02
	(0.11)	(0.12)	(0.11)	(0.12)	(0.13)	(0.12)
Temperature^2	0.29	0.28	0.26	−0.06	−0.08	−0.01
	(0.49)	(0.52)	(0.48)	(0.48)	(0.51)	(0.48)
Precipitation	0.13	−0.03	−0.09	0.05	−0.12	−0.15
	(0.18)	(0.20)	(0.18)	(0.17)	(0.18)	(0.17)
Precipitation^2	−0.34	−0.06	0.08	−0.24	0.05	0.16
	(0.37)	(0.40)	(0.36)	(0.36)	(0.39)	(0.35)
Elevation	0.14	0.42	0.09	−0.11	0.17	−0.10
	(0.63)	(0.68)	(0.64)	(0.60)	(0.65)	(0.61)
Soil	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.03	0.03	0.03	0.03	0.03	0.03
F Stat.	2.72	2.75	2.75	2.72	2.74	2.75
Num. obs.	9,590	9,590	9,590	9,590	9,590	9,590

Notes: The table reports the regression results of eq. (14) using the prefecture seats. The dependent variable is an indicator that equals one, if the pixel hosts a prefecture seat in that year, and zero otherwise. The estimations use the baseline pixels. Conley standard errors using a 150 km radius and a Bartlett kernel are in parentheses ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$). Distances are in 10,000 km, Ruggedness in Ruggedness Index × 10,000,000, Temperature in 100°C, Precipitation in 10 m, Elevation in 100 km. The distance to the nearest courier routes refers to the Ming courier routes published by Berman and Zhang (2017). Categorical soil variables - dominant soil type, landform, lithology - are included as fixed effects.

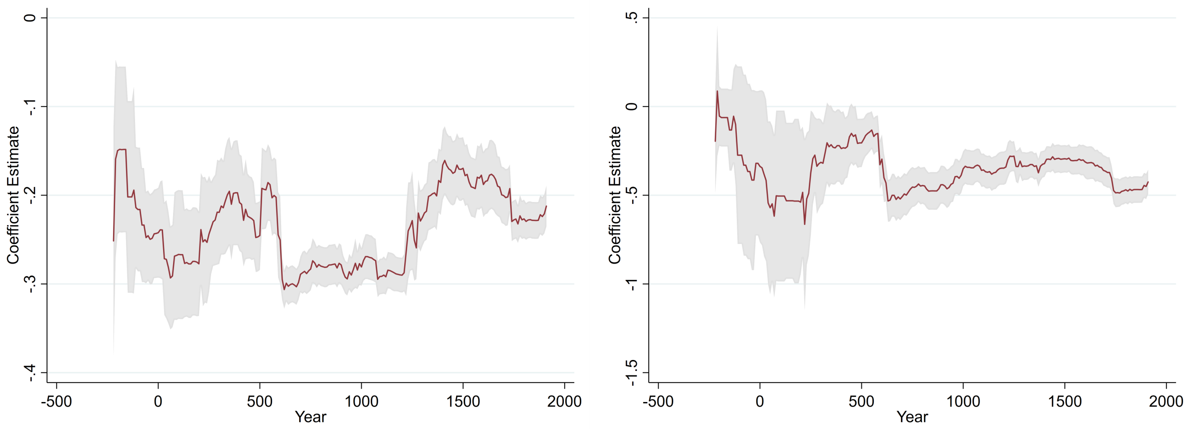
C.12 Robustness Check: Alternative Hiking Functions, Resolutions, and Standard Errors for Distance

In Section 4.2 we compute the travel time between locations using Tobler's (1993) hiking function. To illustrate the robustness of those results, we consider two alternative functions: (i) the one by Márquez-Pérez et al. (2017), (ii) a simple Euclidean distance.

When repeating the estimation with the alternative cost function by Márquez-Pérez et al. (2017) who modified Tobler's specification, the results are so similar that Figure 7a and Figure C-11a look quasi identical. The same holds comparing Figure 7b and Figure C-11b.

Figure C-11: Indirect Effects and a Márquez-Pérez et al. (2017) Hiking Function

(a) Inter-City Distance and Administrative Status **(b)** Pixel Centrality and Administrative Status

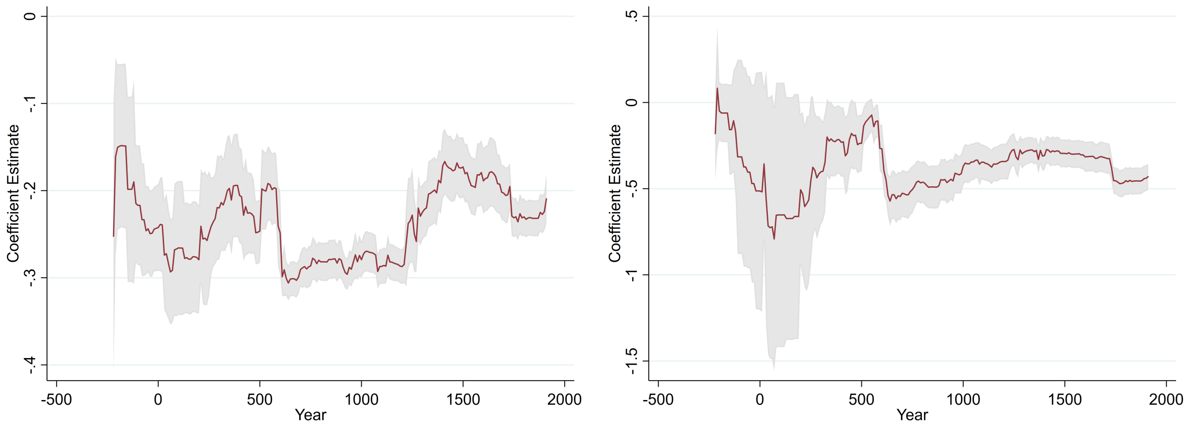


Notes: a) The figure plots the estimate of coefficient β in eq. (16) and its 95% confidence intervals. The dependent variable is a city's total distance to all other cities in the prefecture. b) The figure plots the estimate of coefficient β in eq. (17) and its 95% confidence intervals. The dependent variable is the pixel's distance to prefecture centroid. Unlike the applications depicted in Figure 7a and Figure 7b, the estimations in a) and b) use the hiking function by Márquez-Pérez et al. (2017) rather than the one by Tobler (1993). We derive the estimates from 214 cross-sectional regressions, with one cross-section every ten years between 220 BCE and 1910 CE.

We then turn to Euclidean distance. Unlike travel time computed via hiking functions, the Euclidean distance measure does not take the topography into account. It is simply the length of a straight line between two locations. The outcomes depicted in Figure C-12a and Figure C-12b are similar to Figure 7a and Figure 7b, confirming the robustness of the results to alternative distance measures.

Figure C-12: Indirect Effects and Euclidean Distances

(a) Inter-City Distance and Administrative Status **(b)** Pixel Centrality and Administrative Status

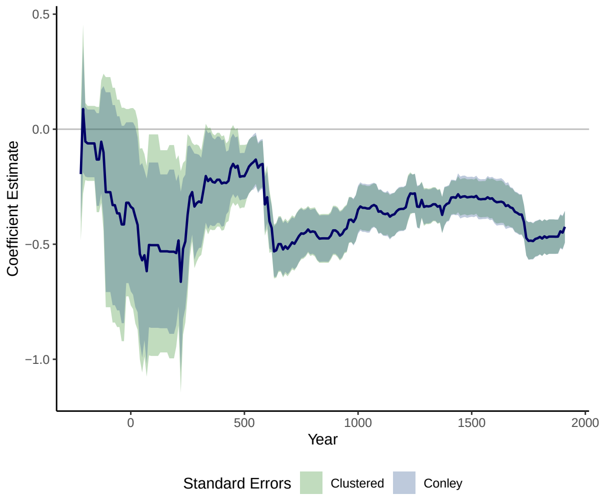


Notes: a) The figure plots the estimate of coefficient β in eq. (16) and its 95% confidence intervals. The dependent variable is a city's total distance to all other cities in the prefecture. b) The figure plots the estimate of coefficient β in eq. (17) and its 95% confidence intervals. The dependent variable is the pixel's distance to prefecture centroid. Unlike the applications depicted in Figure 7a and Figure 7b, the estimations in a) and b) use Euclidean distances. We derive the estimates from 214 cross-sectional regressions, with one cross-section every ten years between 220 BCE and 1910 CE.

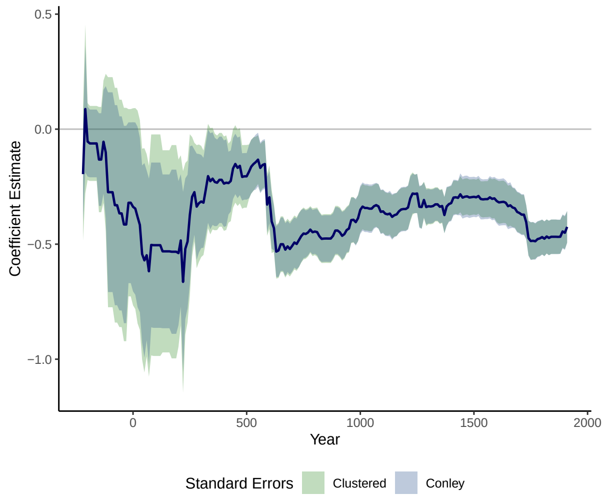
Local geography regressions in Section 4.1 and related estimations employ Conley standard errors by default. By contrast, the 95% confidence intervals in Figure 7b build on standard errors clustered at the prefecture level. We choose this alternative method because the location decision is made within prefectures, making prefecture-level clustering mandatory. Combining Conley standard errors with clustering at a higher spatial level is econometrically underexplored. An in-depth investigation of when such combined standard errors would be valid and what adjustments their estimation would require is still missing from the literature. For consistency reasons, we repeat the analysis with Conley standard errors using a Bartlett kernel and a 150 km radius nonetheless, omitting the prefecture-level clustering. Figure C-13 confirms the baseline results, with Conley standard errors often being even smaller than their clustered counterpart.

Figure C-13: Pixel Centrality and Administrative Status with Very Small Pixels

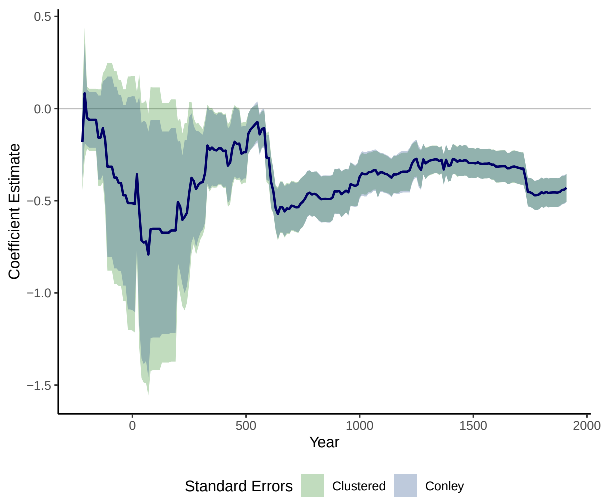
(a) Tobler (1993) Hiking Function



(b) Márquez-Pérez et al. (2017) Hiking Function



(c) Euclidean Distance Function



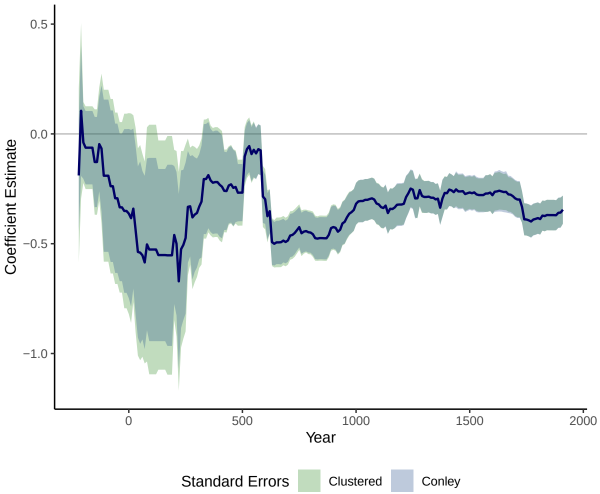
Notes: The figures plot the estimate of coefficient β in eq. (17) and its 95% confidence intervals with different distance functions. The dependent variable is the pixel's distance to the prefecture's centroid. Like Figure 7b, Figure C-11b, and Figure C-12b, the analyses also use distances based on very small pixels. We derive the estimates from 214 cross-sectional regressions, with one cross-section every ten years between 220 BCE and 1910 CE.

Whereas the baseline models on local geography in Section 4.1 use medium sized cells, the remote geography estimations in Section 4.2 and this section are based on the very small pixels. The reason is that with larger grid cells travel paths become excessively unrealistic. (i) Averaging elevation over large rugged terrain does not pose much of a problem in local geography regressions which also include a ruggedness index, apart from elevation. In hiking functions, the elevation layer is the sole input. Smoothing out the mountains in it, eradicates the variation that drives their results. (ii) As travel paths connect pixel centroids, larger cells also mean that the resulting connections between places also adopt increasingly angular shapes. To come up with baseline resolution results nonetheless, we aggregate the finer grid distances into the larger pixels where the distance value of the larger pixel is the average of the distance values of the smaller pixels within its area inside the prefecture borders. With this strategy, we keep the necessary smaller pixel size in hiking functions, but repeat the pixel centrality estimations of Figure 7b, Figure C-11b, and Figure C-12b with medium size pixels.[8] The averaged distance values provide a better picture of the larger cells' position than a centroid-based distance would. However, this also implies that the grid cell hosting the prefecture seat does not have a distance of zero. As Figure C-14 illustrates, the results look nonetheless similar.

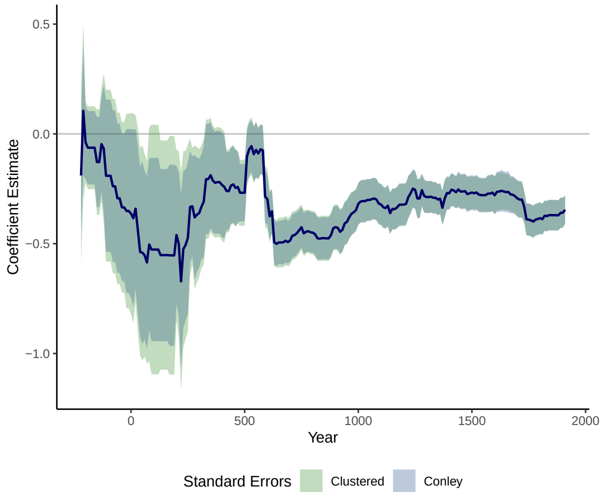
[8]Some of the prefectures that existed at some point of the imperial period were small, consisting of only a few baseline pixels. How exactly border cells are assigned to prefectures can have a notable effect in pixel centrality regressions. When using very small pixels in Section 4.2, we employ the default strategy of assigning a grid cell to the prefecture that the cell's centroid falls into. At 1/9 of medium cells' size, a much smaller share of very small pixels touches any prefecture border, making regressions less sensitive to border pixel assignments. And because border shapes are usually less complex at small than at large scales, the centroid location is a good measure of which region those cells predominantly belong to. In the medium pixel scenario, the centroid rule also identifies the predominant prefecture in most cases. However, there are a few grid cells where most of the cell lies in one prefecture, but because of a complex border shape the centroid lies in the other prefecture. That is why in the subsequent application illustrated in Figure C-14, we assign cells based on area instead of centroid location. A pixel is assigned to a prefecture, if at least 50% of its area falls into the prefecture.

Figure C-14: Pixel Centrality and Administrative Status with Medium Pixels

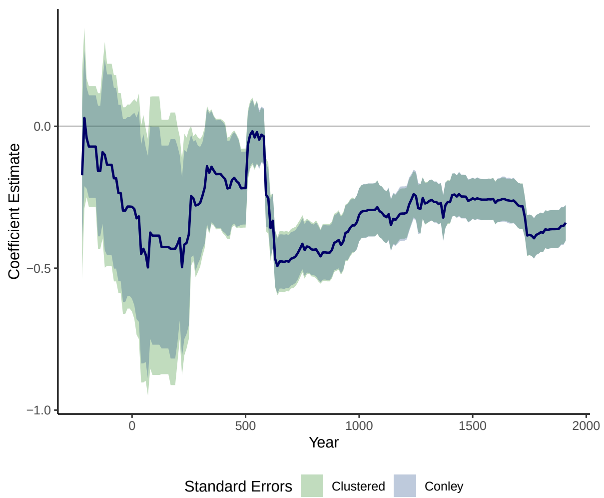
(a) Tobler (1993) Hiking Function



(b) Márquez-Pérez et al. (2017) Hiking Function



(c) Euclidean Distance Function



Notes: The figures plot the estimate of coefficient β in eq. (17) and its 95% confidence intervals with different distance functions. The dependent variable is the pixel's distance to the prefecture's centroid. Like Figure 7b, Figure C-11b, and Figure C-12b, the analyses also use distances based on very small pixels. The difference is that they aggregate these finer distances to the baseline resolution of medium pixels by averaging across the smaller pixels contained in the larger grid cells. We derive the estimates from 214 cross-sectional regressions, with one cross-section every ten years between 220 BCE and 1910 CE.

C.13 Supplementary Results on Modern China

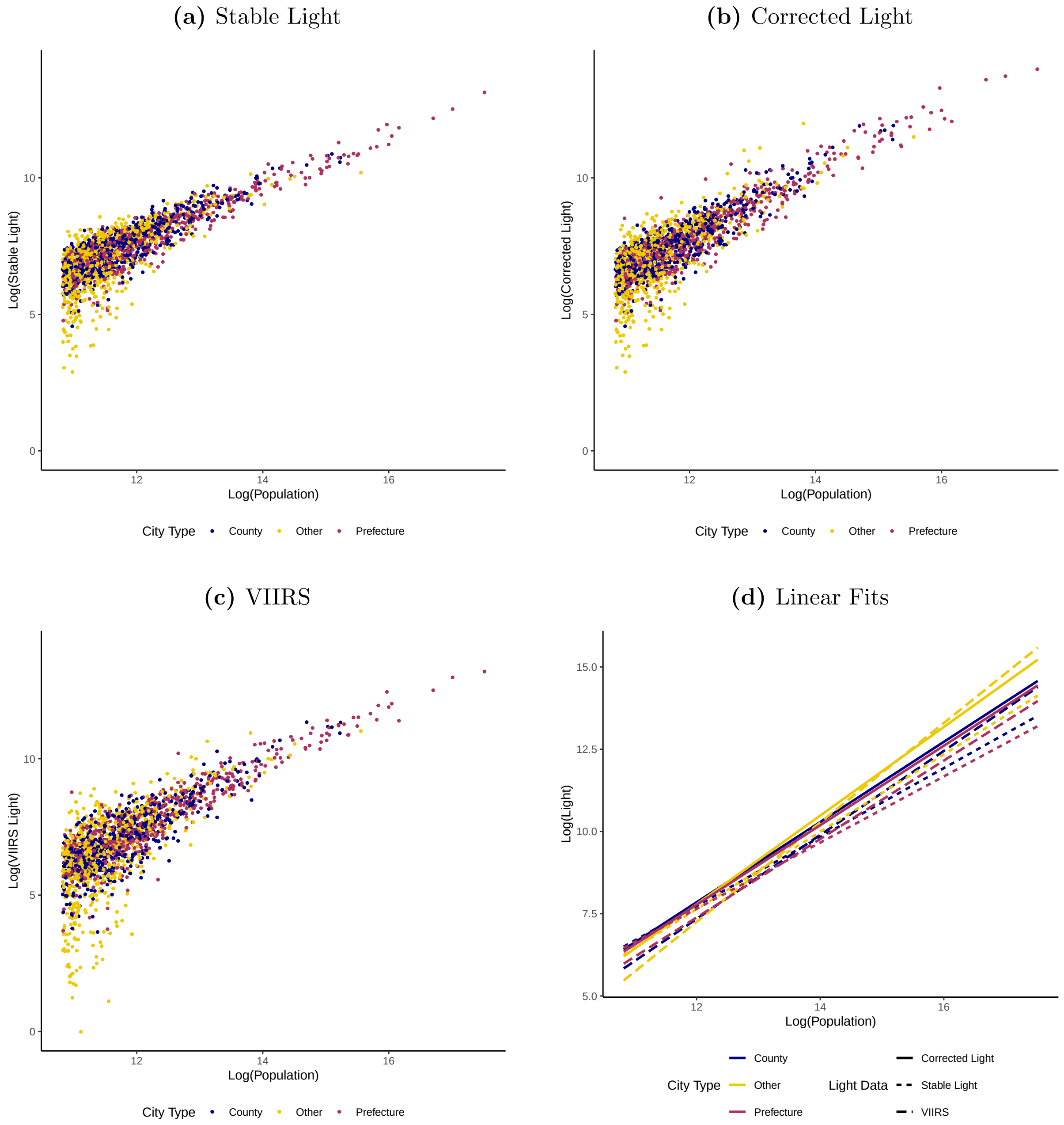
Figure C-15 illustrates the locations of the 29 current province and autonomous region capitals that fall into the territory of former Chinese empire observed in CHGIS. As described in the main text, 28 of these 29 cities were also a prefecture seat some time in imperial China.

Figure C-15: Selected Modern Provincial Capitals

Notes: The depicted capitals are Beijing, Changchun, Changsha, Chengdu, Chongqing, Fuzhou, Guangzhou, Guiyang, Haikou, Hangzhou, Harbin, Hefei, Hohhot, Jinan, Kunming, Lanzhou, Nanchang, Nanjing, Nanning, Shanghai, Shenyang, Shijiazhuang, Taiyuan, Tianjin, Wuhan, Xi'an, Xining, Yinchuan, and Zhengzhou. City locations are derived from OpenStreetMap (OpenStreetMap Contributors, 2019) and Chinese borders from the Global Administrative Areas data base (Center for Spatial Science, University of California, Davis, 2018).

In Figure C-16 we show scatter plots on the relation between population and satellite data of nighttime light as described in the main paper. The figures illustrate the strong association between the two values, underlining our finding that historical prefecture and county seats which are more populous are also more economically active as measured by nighttime lights.

Figure C-16: Modern Chinese Cities' Population Size and Nighttime Light Emissions



Notes: The nighttime lights refer to the 2013 satellite F18 DMSP OLS stable lights image (Elvidge et al., 2014), its top coding corrected version by Bluhm and Krause (2022), and the 2016 VIIRS nighttime lights (Earth Observation Group, NOAA National Centers for Environmental Information, 2016; Elvidge et al., 2017). For a comparison of the different light images, see Bluhm and Krause (2022) and Düben and Krause (2021). We compute a city's luminosity by aggregating its nighttime light emissions within the agglomeration as defined by the GHS Urban Centre Database (Florczyk et al., 2019). Its population is derived from the same source and refers to the identical area.

Geo-spatial population data has been around for some years. The GHS data (Florczyk et al., 2019) that we use is currently the state of the art in that field and improves on many flaws that earlier releases have. The Gridded Population of the World (GPW) data (Center for International Earth Science Information Network - CIESIN - Columbia University, 2018) e.g. uniformly distributes population within census regions.

In places where census regions are large and do not just contain a city, but also its rural hinterland, the city's population is underestimated and the hinterland's population overestimated. The GHS data overcomes this sometimes heavy measurement error by redistributing population within census regions based on built up structures. People live where the houses are. The GHS authors follow a similar approach in determining the spatial extent of agglomerations: they check how far the contiguously built-up space reaches. Despite the improvement relative to earlier alternatives, there is, of course, still some uncertainty involved, in both the spatial population distribution and agglomeration borders. This is why we also test our hypothesis with alternative data sets. In particular, we use population data from the already mentioned GPW (Center for International Earth Science Information Network - CIESIN - Columbia University, 2018) and the older Global Rural-Urban Mapping Project (GRUMP) (Center for International Earth Science Information Network - CIESIN - Columbia University et al., 2011; Balk et al., 2006). Administrative borders provided by the Global Administrative Areas Database (Center for Spatial Science, University of California, Davis, 2018)[9] and functional urban areas by Ma and Long (2020) serve as substitutes for agglomeration shapes. Administrative borders are usually economically less meaningful than agglomerations or functional urban areas. Many settlements have long outgrown their administrative boundaries, rendering its legal border an arbitrary line cutting through a contiguous settlement. Administrative borders especially tend to underestimate the economically relevant size of large and fast-growing urban spaces.

Table C-44, Table C-47, and Table C-50 correspond to Table 4, Table C-45, Table C-48, and Table C-51 to Table 5, and Table C-46, Table C-49, and Table C-52 to Table 6. Comparing them highlights three conclusions. (i) Different population data sets produce similar results. (ii) Varying the spatial definition of cities influences coefficient magnitudes more substantially, as administrative regions, functional urban areas, and built-up- and population density-based agglomerations measure conceptually different structures. And (iii) they robustly support our findings in Section 5. Modern cities (or regions) that hosted an administrative town in imperial times are significantly more populous and brighter than those without such institutional background. Those gains are larger for prefecture than for county seats, which exhibit smaller and sometimes insignificant coefficient estimates in regression with few observations. The higher luminosity appears to be caused by the larger population and not higher productivity.

A valid concern that persists after discussing these robustness checks relates to the role of historic, urban population sizes. If administrative cities hosted larger populations in imperial times than other settlements or regions did, is the population size the channel

[9]The analyses test both ADM 2 and ADM 3 regions.

driving modern outcomes or is it the former political status? Because historic Chinese population counts are unreliable (see Section 3.1), we cannot address that questions directly. Instead, we choose an indirect method that does not provide the deep insights that two millennia of census data would, but adds at least one piece of evidence on that matter. We use the separate 1820 CE cross-sectional CHGIS data set that Section 4.3 builds on and compare administrative to market towns. This rather late cross-section might not be representative of earlier times and market towns could also be smaller than county and prefecture seats were in 1820 CE. Despite these caveats, market towns mark population centers and we can make sure that the other estimations in this section are not driven by comparing former administrative cities to places that did not host any urban settlement in imperial times. The estimation strategy is similar to the one used in Section 5 and in this appendix section. We regress the natural logarithm of a city's total population and nighttime light emissions on historic city type indicators. The regressors are mutually exclusive indicators marking the highest ranked administrative settlement it hosted. That implies that a modern city is only labeled as a former market town, if it neither hosted a prefecture seat nor a county seat. Other than before, we only consider a city's status in 1820 CE, not throughout the entire imperial period.

According to the results in Table C-53, modern cities with a prefecture seat background are significantly larger than other former population centers, whereas there is no robust evidence of a bonus for former county seats over historic market towns.

Table C-44: Imperial History and Population Size of ADM 2 Regions

	OLS		Median Regressions	
	Pop.	Log(Pop.)	Pop.	Log(Pop.)
A. GPW				
Intercept	920,442.89***	13.40***	607,760.23***	13.32***
	(196,533.10)	(0.19)	(210,172.42)	(0.31)
Historic Prefecture Seat	3,598,276.73***	1.69***	2,942,881.67***	1.77***
	(287,092.71)	(0.19)	(291,371.78)	(0.32)
Historic County Seat	1,328,314.42***	0.82***	924,652.28***	0.92***
	(407,857.52)	(0.27)	(313,988.51)	(0.35)
Adj. R^2	0.08	0.26		
Num. obs.	331	331	331	331
B. GRUMP				
Intercept	767,669.02***	13.18***	473,647.62**	13.07***
	(172,288.38)	(0.20)	(201,603.44)	(0.38)
Historic Prefecture Seat	3,316,550.55***	1.83***	2,953,203.76***	1.98***
	(244,270.07)	(0.21)	(272,060.08)	(0.39)
Historic County Seat	1,098,059.68***	0.90***	1,040,562.26***	1.16***
	(313,971.44)	(0.27)	(301,357.29)	(0.41)
Adj. R^2	0.10	0.30		
Num. obs.	331	331	331	331
C. GHS				
Intercept	981,182.53***	13.44***	612,657.54**	13.33***
	(214,967.33)	(0.19)	(283,628.70)	(0.39)
Historic Prefecture Seat	3,375,678.43***	1.61***	3,007,440.92***	1.78***
	(295,958.12)	(0.20)	(349,641.83)	(0.39)
Historic County Seat	1,157,421.29***	0.75***	1,083,836.72***	1.02**
	(378,368.52)	(0.27)	(414,387.20)	(0.43)
Adj. R^2	0.08	0.24		
Num. obs.	331	331	331	331

Notes: Observations refer to the around 96% of modern ADM 2 regions that intersect with a prefecture in imperial times as reported by CHGIS. The dependent variables are ADM 2 regions' total population. GPW population refers to the year 2015, GRUMP population to 2000, and GHS population to 2015. The explanatory indicator variables are mutually exclusive, marking the modern city according to the highest ranked administrative settlement that it hosted. I.e. modern cities are only treated as historic county seat, if they never hosted a prefecture seat. Heteroskedasticity-robust standard errors are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$). Most results remain statistically significant at the thresholds denoted by the asterisks when repeating the OLS estimations using Bartlett kernel Conley standard errors with 150 km and 500 km radii. Exceptions are the county seat estimates in data set C which are significant at the 5% level when using a 500 km radius.

Table C-45: Imperial History and Nighttime Light Emissions of ADM 2 Regions

	OLS			Median Regressions		
	Stable	Corrected	VIIRS	Stable	Corrected	VIIRS
Intercept	10.15***	10.22***	9.00***	10.63***	10.78***	9.28***
	(0.29)	(0.30)	(0.30)	(0.44)	(0.38)	(0.75)
Historic Prefecture Seat	1.23***	1.24***	1.31***	0.74*	0.63	0.95
	(0.29)	(0.31)	(0.31)	(0.44)	(0.38)	(0.76)
Historic County Seat	0.55	0.57	0.67*	0.15	0.06	0.31
	(0.35)	(0.37)	(0.39)	(0.49)	(0.45)	(0.79)
Adj. R^2	0.13	0.11	0.09			
Num. obs.	331	331	331	331	331	331

Notes: Observations refer to the around 96% of modern ADM 2 regions that intersect with a prefecture in imperial times as reported by CHGIS. The dependent variables are the natural logarithm of ADM 2 regions' total nighttime light emissions. The stable and corrected light refer to the last available year (2013) and the VIIRS light data to 2016. The explanatory indicator variables are mutually exclusive, marking the modern city according to the highest ranked administrative settlement that it hosted. I.e. modern cities are only treated as historic county seat, if they never hosted a prefecture seat. Heteroskedasticity-robust standard errors are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$). The results remain statistically significant at the thresholds denoted by the asterisks when repeating the OLS estimations using Bartlett kernel Conley standard errors with 150 km and 500 km radii.

Table C-46: Imperial History, Nighttime Light Emissions, and Population Size of ADM 2 Regions

	OLS			Median Regressions		
	Stable	Corrected	VIIRS	Stable	Corrected	VIIRS
A. GPW						
Intercept	−1.34**	−2.51***	−5.06***	−1.87**	−3.41***	−6.65***
	(0.58)	(0.64)	(0.85)	(0.76)	(0.76)	(0.85)
Historic Prefecture Seat	−0.22	−0.37*	−0.47**	−0.34	−0.72***	−0.31
	(0.19)	(0.19)	(0.21)	(0.26)	(0.25)	(0.39)
Historic County Seat	−0.16	−0.21	−0.19	−0.26	−0.50*	−0.05
	(0.22)	(0.22)	(0.24)	(0.30)	(0.26)	(0.39)
Log(Population)	0.86***	0.95***	1.05***	0.90***	1.04***	1.14***
	(0.04)	(0.04)	(0.06)	(0.05)	(0.05)	(0.06)
Adj. R^2	0.61	0.62	0.61			
Num. obs.	331	331	331	331	331	331
B. GRUMP						
Intercept	−0.44	−1.31*	−3.41***	−1.84**	−2.10**	−3.60***
	(0.69)	(0.76)	(1.00)	(0.79)	(0.84)	(1.19)
Historic Prefecture Seat	−0.24	−0.36*	−0.42*	−0.65***	−0.85***	−0.07
	(0.19)	(0.20)	(0.24)	(0.23)	(0.26)	(0.64)
Historic County Seat	−0.18	−0.21	−0.18	−0.61**	−0.67**	0.11
	(0.22)	(0.23)	(0.27)	(0.24)	(0.27)	(0.65)
Log(Population)	0.80***	0.88***	0.94***	0.93***	0.96***	0.93***
	(0.05)	(0.05)	(0.07)	(0.06)	(0.06)	(0.08)
Adj. R^2	0.54	0.53	0.49			
Num. obs.	331	331	331	331	331	331
C. GHS						
Intercept	−1.94***	−3.13***	−5.44***	−2.09***	−3.40***	−6.27***
	(0.54)	(0.61)	(0.82)	(0.69)	(0.72)	(0.69)
Historic Prefecture Seat	−0.21	−0.36**	−0.42**	−0.55**	−0.66**	−0.16
	(0.17)	(0.17)	(0.21)	(0.25)	(0.29)	(0.38)
Historic County Seat	−0.13	−0.17	−0.14	−0.55**	−0.47	0.04
	(0.20)	(0.20)	(0.24)	(0.27)	(0.29)	(0.38)
Log(Population)	0.90***	0.99***	1.07***	0.93***	1.03***	1.11***
	(0.04)	(0.04)	(0.06)	(0.05)	(0.05)	(0.04)
Adj. R^2	0.66	0.67	0.63			
Num. obs.	331	331	331	331	331	331

Notes: Observations refer to the around 96% of modern ADM 2 regions that intersect with a prefecture in imperial times as reported by CHGIS. The dependent variables are the natural logarithm of ADM 2 regions' total nighttime light emissions. The stable and corrected light refer to the last available year (2013) and the VIIRS light data to 2016. Data sets A, B, and C differ in the population data that they use, where GPW population alludes to 2015, GRUMP to 2000, and GHS to 2015. The explanatory indicator variables are mutually exclusive, marking the modern city according to the highest ranked administrative settlement that it hosted. I.e. modern cities are only treated as historic county seat, if they never hosted a prefecture seat. Heteroskedasticity-robust standard errors are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$). Most results remain statistically significant at the thresholds denoted by the asterisks when repeating the OLS estimations using Bartlett kernel Conley standard errors with 150 km and 500 km radii. Exceptions are the intercept in specification one of data set A which is significant at the 10% level when using a 500 km radius, the intercept in specification 2 of data set B which is insignificant when using a 500 km radius, the prefecture seat estimate in specification 3 of data set B which is insignificant when using a 500 km radius, the prefecture seat estimate in specification two of data set C which is significant at the 5% level when using a 500 km radius, and the prefecture seat estimate in specification three of data set C which is significant at the 5% level when using 150 km and 500 km radii.

Table C-47: Imperial History and Population Size of ADM 3 Regions

	OLS		Median Regressions	
	Pop.	Log(Pop.)	Pop.	Log(Pop.)
A. GPW				
Intercept	420,370.80***	12.35***	256,664.30***	12.46***
	(30,432.73)	(0.05)	(19,798.72)	(0.08)
Historic Prefecture Seat	253,769.85***	0.77***	246,973.14***	0.67***
	(38,540.78)	(0.06)	(24,214.46)	(0.08)
Historic County Seat	169,408.25***	0.62***	189,210.09***	0.55***
	(37,557.81)	(0.06)	(25,738.02)	(0.09)
Adj. R^2	0.02	0.10		
Num. obs.	2,274	2,274	2,274	2,274
B. GRUMP				
Intercept	364,744.28***	12.24***	248,970.53***	12.43***
	(24,165.66)	(0.06)	(16,850.35)	(0.07)
Historic Prefecture Seat	241,864.45***	0.82***	231,071.28***	0.66***
	(29,377.33)	(0.06)	(21,025.34)	(0.07)
Historic County Seat	173,313.32***	0.67***	165,948.45***	0.51***
	(29,973.90)	(0.06)	(22,695.56)	(0.08)
Adj. R^2	0.03	0.11		
Num. obs.	2,274	2,274	2,274	2,274
C. GHS				
Intercept	417,281.37***	12.36***	264,880.18***	12.49***
	(28,265.57)	(0.05)	(19,631.69)	(0.08)
Historic Prefecture Seat	231,137.73***	0.73***	220,705.92***	0.61***
	(36,677.66)	(0.06)	(24,002.77)	(0.08)
Historic County Seat	147,350.04***	0.58***	171,201.53***	0.50***
	(34,989.68)	(0.06)	(25,353.03)	(0.08)
Adj. R^2	0.02	0.09		
Num. obs.	2,274	2,274	2,274	2,274

Notes: Observations refer to the around 95% of modern ADM 3 regions that intersect with a prefecture in imperial times as reported by CHGIS. The dependent variables are ADM 3 regions' total population. GPW population refers to the year 2015, GRUMP population to 2000, and GHS population to 2015. The explanatory indicator variables are mutually exclusive, marking the modern city according to the highest ranked administrative settlement that it hosted. I.e. modern cities are only treated as historic county seat, if they never hosted a prefecture seat. Heteroskedasticity-robust standard errors are in parentheses (***$p < 0.01$; **$p < 0.05$; *$p < 0.1$). The results remain statistically significant at the 1% level when repeating the OLS estimations using Bartlett kernel Conley standard errors with 150 km and 500 km radii.

Table C-48: Imperial History and Nighttime Light Emissions of ADM 3 Regions

	OLS			Median Regressions		
	Stable	Corrected	VIIRS	Stable	Corrected	VIIRS
Intercept	8.82***	8.87***	7.49***	9.06***	9.10***	7.46***
	(0.07)	(0.07)	(0.08)	(0.09)	(0.10)	(0.10)
Historic Prefecture Seat	0.46***	0.48***	0.53***	0.28***	0.28***	0.51***
	(0.08)	(0.08)	(0.09)	(0.10)	(0.11)	(0.11)
Historic County Seat	0.36***	0.35***	0.35***	0.24**	0.23**	0.35***
	(0.08)	(0.08)	(0.09)	(0.11)	(0.11)	(0.12)
Adj. R^2	0.02	0.02	0.02			
Num. obs.	2,274	2,274	2,274	2,274	2,274	2,274

Notes: Observations refer to the around 95% of modern ADM 3 regions that intersect with a prefecture in imperial times as reported by CHGIS. The dependent variables are the natural logarithm of ADM 3 regions' total nighttime light emissions. The stable and corrected light refer to the last available year (2013) and the VIIRS light data to 2016. The explanatory indicator variables are mutually exclusive, marking the modern city according to the highest ranked administrative settlement that it hosted. I.e. modern cities are only treated as historic county seat, if they never hosted a prefecture seat. Heteroskedasticity-robust standard errors are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$). Most results remain statistically significant at the 1% level when repeating the OLS estimations using Bartlett kernel Conley standard errors with 150 km and 500 km radii. Exceptions are the prefecture seat estimates in the stable and corrected lights specifications which remain significant at the 5% level with 500 km radii, and the county seat estimates which remain significant at the 10% level with 500 km radii in all three specifications.

Table C-49: Imperial History, Nighttime Light Emissions, and Population Size of ADM 3 Regions

	OLS			Median Regressions		
	Stable	Corrected	VIIRS	Stable	Corrected	VIIRS
A. GPW						
Intercept	−3.01***	−4.08***	−6.92***	−2.67***	−4.33***	−7.65***
	(0.29)	(0.29)	(0.31)	(0.28)	(0.27)	(0.38)
Historic Prefecture Seat	−0.27***	−0.33***	−0.37***	−0.29***	−0.38***	−0.40***
	(0.05)	(0.05)	(0.06)	(0.06)	(0.05)	(0.07)
Historic County Seat	−0.24***	−0.30***	−0.37***	−0.23***	−0.32***	−0.39***
	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.08)
Log(Population)	0.96***	1.05***	1.17***	0.94***	1.08***	1.23***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Adj. R^2	0.54	0.58	0.55			
Num. obs.	2,274	2,274	2,274	2,274	2,274	2,274
B. GRUMP						
Intercept	−2.19***	−2.95***	−5.27***	−2.64***	−3.39***	−5.80***
	(0.34)	(0.35)	(0.39)	(0.28)	(0.30)	(0.38)
Historic Prefecture Seat	−0.27***	−0.31***	−0.32***	−0.28***	−0.32***	−0.33***
	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)	(0.08)
Historic County Seat	−0.24***	−0.30***	−0.34***	−0.23***	−0.30***	−0.36***
	(0.05)	(0.06)	(0.07)	(0.06)	(0.06)	(0.08)
Log(Population)	0.90***	0.97***	1.04***	0.94***	1.01***	1.08***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)
Adj. R^2	0.47	0.49	0.43			
Num. obs.	2,274	2,274	2,274	2,274	2,274	2,274
C. GHS						
Intercept	−3.42***	−4.50***	−7.15***	−3.20***	−4.64***	−7.83***
	(0.28)	(0.28)	(0.31)	(0.25)	(0.24)	(0.36)
Historic Prefecture Seat	−0.26***	−0.31***	−0.33***	−0.29***	−0.32***	−0.29***
	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.07)
Historic County Seat	−0.21***	−0.28***	−0.33***	−0.22***	−0.25***	−0.27***
	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.07)
Log(Population)	0.99***	1.08***	1.18***	0.98***	1.10***	1.23***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Adj. R^2	0.58	0.62	0.57			
Num. obs.	2,274	2,274	2,274	2,274	2,274	2,274

Notes: Observations refer to the around 95% of modern ADM 3 regions that intersect with a prefecture in imperial times as reported by CHGIS. The dependent variables are the natural logarithm of ADM 3 regions' total nighttime light emissions. The stable and corrected light refer to the last available year (2013) and the VIIRS light data to 2016. Data sets A, B, and C differ in the population data that they use, where GPW population alludes to 2015, GRUMP to 2000, and GHS to 2015. The explanatory indicator variables are mutually exclusive, marking the modern city according to the highest ranked administrative settlement that it hosted. I.e. modern cities are only treated as historic county seat, if they never hosted a prefecture seat. Heteroskedasticity-robust standard errors are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$). Most results remain statistically significant at the 1% level when repeating the OLS estimations using Bartlett kernel Conley standard errors with 150 km and 500 km radii. Exceptions are the prefecture and county seat estimates which remain significant at the 5% level in the first specification of all three data sets when using 500 km radii.

Table C-50: Imperial History and Population Size of Functional Urban Areas

	OLS		Median Regressions	
	Pop.	Log(Pop.)	Pop.	Log(Pop.)
A. GPW				
Intercept	466,694.15***	12.55***	328,827.47***	12.70***
	(79,031.00)	(0.18)	(64,358.40)	(0.21)
Historic Prefecture Seat	1,504,798.30***	1.07***	381,983.51***	0.77***
	(351,820.79)	(0.20)	(91,613.95)	(0.23)
Historic County Seat	209,002.06*	0.38*	82,442.97	0.22
	(119,704.29)	(0.20)	(86,076.72)	(0.26)
Adj. R^2	0.02	0.11		
Num. obs.	320	320	320	320
B. GRUMP				
Intercept	318,517.19***	12.10***	223,870.20***	12.32***
	(54,885.82)	(0.20)	(51,928.10)	(0.23)
Historic Prefecture Seat	1,073,497.14***	1.22***	355,095.30***	0.95***
	(237,371.51)	(0.22)	(68,700.08)	(0.24)
Historic County Seat	163,894.65**	0.44*	107,575.19	0.39
	(78,084.61)	(0.23)	(68,898.88)	(0.27)
Adj. R^2	0.03	0.12		
Num. obs.	320	320	320	320
C. GHS				
Intercept	488,515.97***	12.59***	311,441.38***	12.65***
	(83,422.32)	(0.18)	(43,915.13)	(0.13)
Historic Prefecture Seat	1,430,460.73***	1.04***	372,388.77***	0.79***
	(331,073.86)	(0.20)	(78,309.57)	(0.16)
Historic County Seat	173,086.47	0.34*	84,860.15	0.24
	(118,428.25)	(0.20)	(72,833.32)	(0.20)
Adj. R^2	0.03	0.12		
Num. obs.	320	320	320	320

Notes: Observations refer to the around 99% of modern functional urban areas that intersect with a prefecture in imperial times as reported by CHGIS. The dependent variables are functional urban areas' total population. GPW population refers to the year 2015, GRUMP population to 2000, and GHS population to 2015. The explanatory indicator variables are mutually exclusive, marking the modern city according to the highest ranked administrative settlement that it hosted. I.e. modern cities are only treated as historic county seat, if they never hosted a prefecture seat. Heteroskedasticity-robust standard errors are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$). Most results remain statistically significant at the thresholds denoted by the asterisks when repeating the OLS estimations using Bartlett kernel Conley standard errors with 150 km and 500 km radii. Exceptions are the county seat estimate in specification two of data set A which is significant at the 5% level when using a 500 km radius, and the county seat estimate in specification two of data set C which is significant at the 10% level when using a 500 km radius.

Table C-51: Imperial History and Nighttime Light Emissions of Functional Urban Areas

	OLS			Median Regressions		
	Stable	Corrected	VIIRS	Stable	Corrected	VIIRS
Intercept	9.14***	9.33***	8.57***	9.24***	9.40***	8.49***
	(0.19)	(0.20)	(0.21)	(0.26)	(0.25)	(0.32)
Historic Prefecture Seat	0.88***	0.92***	1.04***	0.59**	0.65**	0.99***
	(0.20)	(0.22)	(0.23)	(0.27)	(0.27)	(0.33)
Historic County Seat	0.24	0.20	0.19	0.13	0.09	0.19
	(0.21)	(0.23)	(0.24)	(0.28)	(0.27)	(0.33)
Adj. R^2	0.09	0.09	0.11			
Num. obs.	320	320	320	320	320	320

Notes: Observations refer to the around 99% of modern functional urban areas that intersect with a prefecture in imperial times as reported by CHGIS. The dependent variables are the natural logarithm of functional urban areas' total nighttime light emissions. The stable and corrected light refer to the last available year (2013) and the VIIRS light data to 2016. The explanatory indicator variables are mutually exclusive, marking the modern city according to the highest ranked administrative settlement that it hosted. I.e. modern cities are only treated as historic county seat, if they never hosted a prefecture seat. Heteroskedasticity-robust standard errors are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$). All results remain statistically significant at the thresholds denoted by the asterisks when repeating the OLS estimations using Bartlett kernel Conley standard errors with 150 km and 500 km radii.

Table C-52: Imperial History, Nighttime Light Emissions, and Population Size of Functional Urban Areas

	OLS			Median Regressions		
	Stable	Corrected	VIIRS	Stable	Corrected	VIIRS
A. GPW						
Intercept	-2.25^{***}	-3.46^{***}	-4.03^{***}	-2.37^{***}	-3.37^{***}	-4.29^{***}
	(0.35)	(0.38)	(0.36)	(0.12)	(0.32)	(0.35)
Historic Prefecture Seat	-0.09	-0.17	-0.04	0.02	-0.03	-0.05
	(0.11)	(0.11)	(0.12)	(0.07)	(0.09)	(0.15)
Historic County Seat	-0.11	-0.19^{*}	-0.19	-0.00	-0.11	-0.14
	(0.11)	(0.12)	(0.12)	(0.07)	(0.10)	(0.15)
Log(Population)	0.91^{***}	1.02^{***}	1.00^{***}	0.91^{***}	1.00^{***}	1.02^{***}
	(0.02)	(0.03)	(0.03)	(0.01)	(0.02)	(0.02)
Adj. R^2	0.89	0.88	0.86			
Num. obs.	320	320	320	320	320	320
B. GRUMP						
Intercept	-0.56	-1.38^{**}	-1.81^{***}	-1.55^{***}	-2.39^{***}	-3.00^{***}
	(0.48)	(0.55)	(0.60)	(0.29)	(0.37)	(0.51)
Historic Prefecture Seat	-0.09	-0.16	-0.00	0.01	-0.08	-0.05
	(0.13)	(0.14)	(0.16)	(0.13)	(0.15)	(0.23)
Historic County Seat	-0.12	-0.20	-0.19	0.02	-0.06	-0.19
	(0.13)	(0.15)	(0.17)	(0.13)	(0.16)	(0.23)
Log(Population)	0.80^{***}	0.89^{***}	0.86^{***}	0.87^{***}	0.95^{***}	0.95^{***}
	(0.04)	(0.04)	(0.05)	(0.02)	(0.03)	(0.04)
Adj. R^2	0.81	0.78	0.74			
Num. obs.	320	320	320	320	320	320
C. GHS						
Intercept	-2.66^{***}	-3.95^{***}	-4.50^{***}	-2.75^{***}	-4.11^{***}	-4.76^{***}
	(0.33)	(0.35)	(0.34)	(0.19)	(0.28)	(0.39)
Historic Prefecture Seat	-0.09	-0.17^{*}	-0.04	0.06	-0.06	0.03
	(0.10)	(0.10)	(0.11)	(0.08)	(0.09)	(0.19)
Historic County Seat	-0.09	-0.17	-0.16	0.08	-0.08	-0.02
	(0.10)	(0.11)	(0.12)	(0.08)	(0.09)	(0.20)
Log(Population)	0.94^{***}	1.06^{***}	1.04^{***}	0.93^{***}	1.06^{***}	1.05^{***}
	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.03)
Adj. R^2	0.89	0.90	0.87			
Num. obs.	320	320	320	320	320	320

Notes: Observations refer to the around 99% of modern functional urban areas that intersect with a prefecture in imperial times as reported by CHGIS. The dependent variables are the natural logarithm of functional urban areas' total nighttime light emissions. The stable and corrected light refer to the last available year (2013) and the VIIRS light data to 2016. Data sets A, B, and C differ in the population data that they use, where GPW population alludes to 2015, GRUMP to 2000, and GHS to 2015. The explanatory indicator variables are mutually exclusive, marking the modern city according to the highest ranked administrative settlement that it hosted. I.e. modern cities are only treated as historic county seat, if they never hosted a prefecture seat. Heteroskedasticity-robust standard errors are in parentheses ($^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$). Most results remain statistically significant at the thresholds denoted by the asterisks when repeating the OLS estimations using Bartlett kernel Conley standard errors with 150 km and 500 km radii. Exceptions are the county seat estimate in specification two of data set A which is insignificant when using a 500 km radius and the intercept in specification three of data set B which is significant at the 5% level when using a 500 km radius.

Table C-53: Modern Cities' History as Administrative and Other Population Centers

	OLS			Median Regressions		
	Population		Light	Population		Light
	GHS	GPW	VIIRS	GHS	GPW	VIIRS
A. GHS Agglomerations						
Intercept	11.50***	10.22***	6.45***	11.33***	10.26***	6.57***
	(0.03)	(0.06)	(0.07)	(0.03)	(0.07)	(0.06)
Prefecture Seat in 1820 CE	1.27***	1.94***	1.89***	1.16***	1.88***	1.71***
	(0.09)	(0.13)	(0.13)	(0.12)	(0.17)	(0.14)
County Seat in 1820 CE	0.30***	0.58***	0.69***	0.32***	0.53***	0.47***
	(0.04)	(0.08)	(0.08)	(0.05)	(0.09)	(0.08)
Market Town in 1820 CE	0.31***	0.64***	0.72***	0.24***	0.56***	0.52***
	(0.05)	(0.10)	(0.10)	(0.06)	(0.11)	(0.11)
Adj. R^2	0.17	0.13	0.14			
Num. obs.	1,899	1,899	1,899	1,899	1,899	1,899
B. Functional Urban Areas						
Intercept	12.40***	12.32***	8.32***	12.58***	12.44***	8.45***
	(0.18)	(0.18)	(0.19)	(0.23)	(0.22)	(0.30)
Prefecture Seat in 1820 CE	1.47***	1.56***	1.49***	1.10***	1.27***	1.08***
	(0.21)	(0.21)	(0.22)	(0.25)	(0.25)	(0.34)
County Seat in 1820 CE	0.69***	0.75***	0.69***	0.45*	0.62**	0.40
	(0.20)	(0.20)	(0.22)	(0.25)	(0.24)	(0.33)
Market Town in 1820 CE	0.64***	0.70***	0.71***	0.46	0.56*	0.40
	(0.22)	(0.22)	(0.24)	(0.32)	(0.29)	(0.35)
Adj. R^2	0.17	0.18	0.13			
Num. obs.	323	323	323	323	323	323
C. ADM 3 Regions						
Intercept	12.27***	12.23***	7.39***	12.48***	12.36***	7.60***
	(0.09)	(0.09)	(0.12)	(0.12)	(0.09)	(0.15)
Prefecture Seat in 1820 CE	1.03***	1.10***	1.16***	0.77***	0.94***	0.96***
	(0.10)	(0.10)	(0.15)	(0.13)	(0.11)	(0.19)
County Seat in 1820 CE	0.58***	0.66***	0.34***	0.45***	0.61***	0.13
	(0.09)	(0.09)	(0.13)	(0.12)	(0.10)	(0.16)
Market Town in 1820 CE	0.40***	0.45***	0.37***	0.31**	0.44***	0.18
	(0.10)	(0.10)	(0.14)	(0.13)	(0.11)	(0.17)
Adj. R^2	0.06	0.06	0.04			
Num. obs.	2,408	2,408	2,408	2,408	2,408	2,408

Notes: The dependent variables are the natural logarithm of cities' total GHS population (in 2015) and VIIRS nighttime light emissions (in 2016). The data sets A, B, and C differ in their spatial delineation of cities as GHS agglomerations, functional urban areas, and ADM 2 regions respectively. The estimations only include cities which intersect with the shape of China in 1820 CE as reported by the special cross-sectional CHGIS data set on that year. The explanatory indicator variables are mutually exclusive, marking the modern city according to the highest ranked administrative settlement that it hosted. I.e. modern cities are only treated as historic market town, if they neither hosted a prefecture seat nor a county seat in 1820 CE. Heteroskedasticity-robust standard errors are in parentheses (***$p < 0.01$; **$p < 0.05$; *$p < 0.1$). The results remain statistically significant at the 1% level when repeating the OLS estimations using Bartlett kernel Conley standard errors with a 150 km and a 500 km radius. Most results remain statistically significant at the thresholds denoted by the asterisks when repeating the OLS estimations using Bartlett kernel Conley standard errors with 150 km and 500 km radii. Exceptions are the market town estimate in specification one of data set C and which is significant at the 5% level when using a 500 km radius, and the county and market town estimates in specification three of data set C which are significant at the 5% level when using a 150 km radius and insignificant when using a 500 km radius.