Geography and State Fragmentation*

Shuhei Kitamura[†] Nils-Petter Lagerlöf[‡]

October 18, 2017

Abstract: Some of the richest places in the world have very high historical border presence, and are often located in particular geographic environments. Here we compile grid-cell level data on borders between sovereign states in Europe and surrounding areas from 1500 until today. We find that state borders tend to be located in rugged and mountainous terrain, by rivers, and where it rains a lot. Moreover, modern development, as measured by night lights and population density, shows positive correlation with border presence, even when controlling for geography. However, cells that have more borders than immediate neighbors are less developed than those neighbors. These patterns are consistent with a theory in which state competition benefits long-run development, but these benefits accrue more to the center than the periphery of countries.

*We received helpful input from Scott Abramson, Tasso Adamopoulos, Philippe Aghion, Matilde Bombardini, Andrew Dickens, Viktoria Hnatkovska, Kris Inwood, Ken Jackson, Mark Koyama, and attendants at presentations at the IIES, Kyoto University, St Louis Fed, University of Guelph, York University, the 2015 Growth Theory Conference at Brown University, the Econometric Society Meetings in Toulouse in 2015, and a workshop at the University of Guelph in 2017. Kitamura thanks Handelsbanken's Research Foundations for financial support, and Lagerlöf thanks the Social Sciences and Humanities Research Council of Canada for financial support. All errors are ours. An earlier version of this paper carried the title Natural Borders.

†Osaka School of International Public Policy, Osaka University. Email: kitamura@osipp.osaka-u.ac.jp.

‡Department of Economics, York University, Toronto. Email: lagerlof@econ.yorku.ca.

1 Introduction

Many scholars have emphasized the benefits of state fragmentation for preindustrial economic and institutional development. For example, interstate competition can spur elites to build state and fiscal capacity (Tilly 1992), invest in education (Aghion et al. 2015), and democratize (Halperin 2004, Ch. 5).

A somewhat separate body of literature discusses the relationship between geography and state fragmentation. Diamond (1997) argues that Western Europe became more fragmented than, e.g., China because of its mountains, rivers, and indented coastline. Others examine what constitutes "natural" borders between states and peoples (e.g., Pounds 1972).

Taken together, these two ideas—that geography affects state fragmentation, and that state fragmentation affects development—seem to suggest a relatively unexplored mechanism through which differences in geography across world regions can create divergent trajectories of long-run economic development.

Here we want to explore this further. We start with a set of historical maps over sovereign states covering Europe, Western Asia, and North Africa. We focus on a region with statehood present from 1500 to 2000 and divide this into grid cells. For each cell, we compute what we call border frequency—the fraction years in which a cell had a border, defined as more than one state entering the cell.

Our first task is to examine how various measures of geography correlate with border frequency. We find that high border frequency is associated with mountainous, rugged, and rainy terrain, and with rivers.

We also want to know how local these correlations are. To that end, we compute differences in border frequency between each cell and its closest neighboring cells, and correlate these differences with the corresponding differences in our geography variables. Many of the correlations hold at the local level: cells with more rivers, mountains, ruggedness, and rainfall than their neighboring cells have more borders than those neighbors. In that sense, these features of the landscape seem to constitute some sort of natural borders. By extension, it suggests that political fragmentation may be partly caused by an abundance of such natural borders.

Some other geography variables have a less intuitive effect on borders. For example, how suitable the land is for agriculture seems to matter, but differently depending on the type of agriculture: suitability for rainfed agriculture is associated with more borders, and suitability for irrigated agriculture with fewer borders. We propose an interpretation that relates to Wittfogel (1957).

We do not claim that geography was the only factor determining the shape of states. Our favorite regression specification generates an R-squared of about 12%. At the same time, geography can have effects through interaction with non-geographical factors. For example, we find that borders are less likely to survive if they separate small states from large, but this size-difference effect is mitigated by, e.g., rugged terrain.

We then explore the correlation between borders and modern development. We use two contemporary outcome variables—population density and night lights—both of which are higher in cells with higher border frequency. This reflects that some of the most developed locations according to these measures are clustered in historically border dense areas, in particular around the Alps and southern Germany, while more unified areas, e.g., in today's Turkey and Russia, are less developed.

These correlations hold when controlling for geography, suggesting that borders are not merely capturing a direct effect from geography on development, and are also robust to many other controls. However, we do find weaker effects when treating the Holy Roman Empire as unified, which is not too surprising, since this includes some of the most developed areas of the world today.

We then compute differences in night lights and population density between each cell and its adjacent neighboring cells, to correlate with the corresponding deviation in border frequency. Perhaps surprisingly, we find that cells with higher border frequency than neighboring cells emit *less* night lights, and have *lower* population densities, than those neighbors. In other words, the benefits of state fragmentation seem to be a macro phenomenon, capturing variation across regions, while the costs of fragmentation seem to appear locally precisely at the borders. We present some tentative evidence that this may relate to war and conflict. For example, we find that instability of borders has a negative effect on modern development when controlling for overall border frequency.

To make sense of some of these correlations between modern development and historical border frequency we propose a spatial model with inter-state resource competition. The idea is that more fragmentation induces more investment in technology for the purpose of resource competition, while the economic benefits of these investments accrue disproportionately to the center of each state. Border locations thus become less developed than the states' interiors, while regions with more state fragmentation, and higher border frequency, may still be on average more developed than more unified regions.

This paper relates to a debate about whether geography caused Europe's high fragmentation compared to other regions, such as China, and the effects this has had on development (e.g., Diamond 1997, Lagerlöf 2014, Ko et al. 2017). Our data do not cover China *per se*, but several other politically unified regions, such as Russia and the Ottoman Empire, so we may say something about the deeper state formation factors at play across the world.

Several influential papers study ethno-linguistic fractionalization (e.g., Ahlerup and Olsson 2012, Michalopoulos 2012, Ashraf and Galor 2013), often using data on contemporary locations of language groups. Studying historical state borders is related, but starts from a different angle. For example, the process of state formation and unification can influence ethnic diversity through genocide and ethnic cleansing, especially in the region we study.

Because the absence of borders may partly reflect early state development, in particular around the Middle East, our analysis may also add to a literature documenting that regions which developed statehood early tend to be poorer today than those with intermediate state antiquity (e.g., Borcan et al. 2015).

Another strand of the literature analyzes the size, shape, and composition of countries, often with a focus on the determinants of the optimal and/or equilibrium state structure (e.g., Friedman 1977, Bolton and Roland 1997, Alesina and Spolaore 2003, Gancia et al. 2016). Alesina et al. (2011) measure how artificial, or non-natural, modern borders are and look at correlations with economic outcomes. Abramson and Carter (2016) study the effects of historical borders on contemporary disputes between countries. None of these links both geography to fragmentation, and fragmentation to development.

Finally, our suggested deep-rooted link from geography to development relates to several papers on the interactions between geography, institutions, and development, but here working through state fragmentation, and applying to the Old World, rather than post-Columbian migration and the New World disease environment (e.g., Sokoloff and Engerman 2000; Acemoglu et al. 2001, 2002).

The rest of this paper is organized as follows. In Section 2, we summarize some of the discussion on the links between geography, state fragmentation, and development. In Section 3 we set up a simple model of state competition to help us interpret some of the results. Section 4 describes the data. Section 5 presents our empirical results, first regressing borders on geography (Section 5.2) and then modern outcomes on borders (Section 5.3). Section 6 concludes.

2 Background

2.1 State fragmentation and development

There are many indications that fragmentation may promote economic development. Perhaps most obviously, many small European states are among the world's richest (Alesina 2003, Alesina et al. 2005), and Western Europe as a whole is more fragmented and richer than most of the rest of Eurasia.

Explanations often center on interstate competition. This created incentives for ruling elites to build state and fiscal capacity (Tilly 1992, Besley and Persson 2011, Dincecco and Prado 2012), invest in education (Aghion et al. 2015), pursue democratic reform (Halperin 2004, Ch. 5; Ticchi and Vindigni 2008), and embrace technological innovation for military purposes (Jones 1981, McNeill 1982, Lagerlöf 2014, Hoffman 2015). Interstate conflict can also facilitate the emergence of credit markets to fund wars (Ferguson 2009, Gennaioli and Voth 2015).

State fragmentation can also spread risks, making widespread destructive consequences less likely if the ruler of one state makes a bad decision. For example, China's large 15th-century overseas explorations ended in the wake of infighting within the Chinese court, while Christopher Columbus could solicit funding for his voyages from different European monarchs (e.g., Diamond 1997, pp. 412-413; Landes 1998, pp. 93-98).

The size of countries can also affect, e.g., culture and migration opportunities. Guiso et al. (2016) find that Italian cities which experienced more independence in the Middle Ages have higher levels of social capital today than other Italian cities, according to various indicators. Data on 411 European authors living 1660-1961 suggest that they were more likely to flee from small countries (Potrafke and Vaubel 2014).[1]

Access to public goods can also decline with the distance from the geographical center of countries, where capitals are often located; for examples, see Alesina and Spolaore (2003, pp. 34-35). This may also affect economic development.

2.2 State fragmentation and geography

Authors linking geography and state (and linguistic) fragmentation often compare Europe and China. Diamond (1997, pp. 414-415) lists several differences. First, he points to Europe's indented coastline, i.e., its many peninsulas and islands, like Iberia, Italy, Scandinavia, Britain, and Ireland (see also Cosandey 1997, Ch. 6). Second, he argues that Europe is particularly disconnected by mountain chains, such as the Alps and Pyrenees. (These two points are critiqued by Hoffman 2015, Ch. 4.) Third, Diamond suggests that rivers in Europe are particularly likely to separate states and peoples because they run north-to-south, thus connecting regions which are climatically different.

Exactly how geography shapes borders has been discussed at least since the early 20th century (e.g., Lord Curzon of Keddleston 1907, Holdich 1916, Brigham 1919; for an overview, see Pounds 1972). Mountains tend to be natural borders because they are easy to defend militarily (Pounds 1972, pp. 86-88). Many famous defensive fortifications were located in mountainous and rugged areas, such as the Great Wall of China and Hadrian's Wall in Britain. Rivers, by contrast, often had a unifying character in pre-state times; people have always lived, traveled, and traded along rivers (Pounds 1972, Ch. 11). The fragmenting role

[1] Karayalçin (2008) argues that despots who care about the size of their populations have an incentive to reform to curb emigration. See also Mokyr (2006, 2007) for further arguments that migration was one mechanism behind the benefits of state fragmentation.

of rivers seems to be a by-product of state formation.[2]

Wittfogel (1957) argued that large-scale irrigation projects tended to make societies more despotic; see also Bentzen el al. (2016) for a test of this hypothesis. This may have affected state fragmentation, if strong and despotic states and rulers were prone to, and capable of, spatial expansion. The Middle Eastern regions where states and irrigated agriculture first evolved were also centers of several of the largest empires in the world (Taagepera 1978). The rise of the first central states in the region has been linked to centralized control of the water supply (see, e.g., Nissen and Heine 2009, Ch. 5). By contrast, where agriculture is mostly rainfed it may be more difficult for any single state to dominate others by monopolizing water supply.

Rainfall itself may play a direct role for fragmentation, independently of how it affects suitability for rainfed agriculture. First, it exhibits positive (albeit weak) correlation with ruggedness and elevation, especially at low elevations, and may thus proxy for a mountainous landscape.[3] Second, a dry climate can facilitate storage, which in turn may have promoted the early emergence of states, possibly territorially large states too (Mayshar et al. 2015, Elis et al. 2017). Third, drier climates may be more suitable for pastoral farming, which can also affect state formation (Barfield 1989, Umesao 2003, Turchin 2009, Ko et al. 2017). Because such factors can impact both borders and development, we choose to include rainfall as a control in our benchmark regressions.

3 A model

As hinted already, and will be shown later, the correlations between borders and development outcomes seem to differ at the "global" and the "local" levels, i.e., depending on whether we compare all cells, or only cells close to each other. This section proposes a model to help us make sense of these patterns.

Let a region consist of N identically sized states, or countries, located on the interval $[0, 1]$, each of size $s = 1/N$. Changing N captures global correlations, and moving across locations within one region, holding N constant, captures local correlations.

To focus specifically on development outcomes, we treat N as exogenous. However, we may think of these states as separated by "natural" borders, determined exogenously by

[2] In the words of Lord Curzon of Keddleston (1907, p. 8): "As States developed and considerable armies were required for their defense, the military value of rivers, in delaying the enemy, and in concentrating defensive action at certain bridges, or fords, or posts, became apparent." Consistent with this interpretation, studies looking at ethno-linguistic diversity find little effect from rivers (Michalopoulos 2012, p. 1525 and Footnote 10).

[3] Rainfall tends to increase with elevation until about 1000-2000 meters above sea level, after which it declines with elevation (Miller 1953, p. 39). In our data, it is maximized around 1,300 meters.

features of the landscape, such as mountains. As will be seen in Section 5.2, several measures of geography show robust correlation with borders in the data.

Decisions in each country are made by an elite, who collect taxes and invest in a public good, which we call technology. Technology is an input in both production and resource competition, which allows the model to simultaneously capture a scale effect, through which larger countries can accumulate more technology, as well as a technology-promoting role of state competition.[4] Before we study the elite's optimal behavior, we explain production.

3.1 Production

Production takes place across the whole space $[0, 1]$. Building on the idea that technology (or other productivity-enhancing public goods, such as law enforcement) is concentrated at the center of each country, we assume that productivity decreases with the distance from that center. This assumption is inspired by Alesina and Spolaore (2003), although here applied to productivity rather than preferences. Specifically, output at distance $d \in [0, s/2]$ from the center is denoted $\widetilde{Y}_{i,t}(d)$, where $i \in \{1, ..., N\}$ indicates the country and t the period.[5] We let

$$\widetilde{Y}_{i,t}(d) = \widetilde{Z}(d) A_{i,t}^{\alpha} R_{i,t}^{1-\alpha}, \tag{1}$$

where $\alpha \in (0, 1)$; $\widetilde{Z}(d)$ is a location dependent productivity factor, such that $\widetilde{Z}(0) = 1$ and $\widetilde{Z}'(d) < 0$; $A_{i,t}$ is country i's level of technology, which is non-rivalrous and thus not location specific; and $R_{i,t}$ is the resource input at each location. We thus assume that resources are uniformly distributed across locations, which can be rationalized with a Leontief-type aggregation of resource inputs across space.[6]

To derive closed-form solutions, we let

$$\widetilde{Z}(d) = 1 - 4\gamma d, \tag{2}$$

where γ measures the cost of distance. We assume that $\gamma \leq 1/2$ to ensure non-negative productivity in a fully unified region ($s = 1$) at the maximum distance from the center ($d = 1/2$).

[4] See Lagerlöf (2014) for a related but non-spatial framework, where both manpower and technology are inputs in resource competition.

[5] To be precise, country i's borders are located at $s(i-1)$ and si, and its center at $c_i = si - (s/2)$. The distance from a location $l \in [0, 1]$ to the closest center (the one in the country to which l belongs) equals $d = \min_{i \in \{1,...,N\}} |l - c_i|$.

[6] Alternatively, we could let output be given by $\widetilde{Y}_{i,t}(d) = [\widetilde{Z}(d) A_{i,t}]^{\alpha} \left[\widetilde{R}_{i,t}(d)\right]^{1-\alpha}$, where $\widetilde{R}_{i,t}(d)$ is the amount of resources allocated to locations at distance d from the center. This generates a reduced-form production function very similar to that in (1). See Section 1 of the Online Appendix.

Since each country has two symmetric halves of size $s/2$, average output per location in country i equals $Y_{i,t} = (2/s) \int_0^{s/2} \widetilde{Y}_{i,t}(x)dx$. Using (1) and (2), noting that $\int_0^{s/2} \widetilde{Z}(x)dx = (s/2)(1-\gamma s)$, we can write output per location as

$$Y_{i,t} = \frac{2}{s}\left[\int_0^{s/2} \widetilde{Z}(x)dx\right] A_{i,t}^{\alpha} R_{i,t}^{1-\alpha} = (1-\gamma s)\, A_{i,t}^{\alpha} R_{i,t}^{1-\alpha}. \tag{3}$$

Note also that $sY_{i,t}$ is country i's total output.

3.2 Resource competition

All locations are endowed with a mass of resources, $\overline{R}$, but each country controls only a fraction $1-\lambda$ of the resources on its territory. The reminder is allocated among the N countries in proportion to their technology levels. This is a convenient way to model inter-state competition, while keeping the borders themselves fixed.

Country i's resources per location, $R_{i,t}$, are thus given by

$$R_{i,t} = (1-\lambda)\overline{R} + \frac{1}{s}\left(\frac{A_{i,t}}{\sum_{j=1}^{N} A_{j,t}}\right)\lambda\overline{R}, \tag{4}$$

where (recall) $A_{i,t}$ is country i's technology, and $A_{i,t}/\sum_{j=1}^{N} A_{j,t}$ its share of the pool of contested resources, of total size $\lambda\overline{R}$; the conquered resources are distributed uniformly across a territory of size $s = 1/N$. In a symmetric equilibrium, where all countries have the same $A_{i,t}$, it can be seen that $R_{i,t} = \overline{R}$.

3.3 The elite's decisions

The elite tax the country's total output, $sY_{i,t}$, at an exogenous rate τ. Tax revenues are used for elite consumption, denoted $C_{i,t}$, and investment in the next period's technology, $A_{i,t+1}$, at a cost of one unit of consumption per unit of technology, i.e.,

$$C_{i,t} = \tau s Y_{i,t} - A_{i,t+1}. \tag{5}$$

The elite care about their own current consumption, $C_{i,t}$, and the tax revenue of the next generation of the elite, $\tau s Y_{i,t+1}$. Utility is logarithmic, with relative weight $\beta \in (0,1)$ on next period's tax revenue, i.e.,

$$U_{i,t} = (1-\beta)\ln(C_{i,t}) + \beta \ln(\tau s Y_{i,t+1}). \tag{6}$$

Next period's technology, $A_{i,t+1}$, is set to maximize utility in (6), subject to (5), and (3) and (4) forwarded to period $t+1$. The first-order condition can be written

$$(1-\beta)\left[C_{i,t}\right]^{-1} = \beta\left\{\alpha + (1-\alpha)\left(\frac{\partial R_{i,t+1}}{\partial A_{i,t+1}}\frac{A_{i,t+1}}{R_{i,t+1}}\right)\right\}\left[A_{i,t+1}\right]^{-1}. \tag{7}$$

3.4 Symmetric equilibrium and steady state

In a symmetric equilibrium (denoted by dropping the subindex i) all countries are identical in each period, and have resources $\overline{R}$ per location. As shown in Section A of the Appendix, each country sets technology investment for the next period to $A_{t+1} = \tau s Y_t \beta F(s, \lambda)/[1 - \beta + \beta F(s, \lambda)]$, where

$$F(s, \lambda) = \alpha + (1 - \alpha)\lambda\,(1 - s)\,. \tag{8}$$

The factor $\lambda\,(1 - s)$ in (8) is the elasticity of each country's resources with respect to its technology level, $(\partial R_{i,t+1}/\partial A_{i,t+1})\,(A_{i,t+1}/R_{i,t+1})$, appearing in (7), evaluated in a symmetric equilibrium.

Next, we can use (3), and the dynamics of A_t, to derive an dynamic equation for Y_t; see (29) in the Appendix. Letting Y^* denote the steady state level of Y_t (derived by setting $Y_{t+1} = Y_t = Y^*$), some algebra shows that

$$Y^* = [s^{\alpha}\,(1 - \gamma s)]^{\frac{1}{1-\alpha}} \left[\frac{\tau\beta F(s, \lambda)}{1 - \beta + \beta F(s, \lambda)}\right]^{\frac{\alpha}{1-\alpha}} \overline{R}. \tag{9}$$

Now (8) and (9) define a relationship between steady-state output per location (Y^*) and fragmentation ($s = 1/N$) across regions. Increasing fragmentation (lowering s) has three different effects. Two factors benefit fragmented (small-s) regions. First, in more fragmented regions the average location is closer to the center of its country. This effect is stronger if γ is large. Second, more fragmentation implies stronger competition-driven incentives to invest in technology, which also affects output. This effect is stronger if λ is large, working through $F(s, \lambda)$. The third factor is a scale effect, which works against fragmented regions, since each country's non-rivalrous technology gets applied over a smaller area. This effect tends to be stronger if α is large.[7]

To summarize the model's predictions we can compare a fully unified region to one with two countries. For ease of exposition, first let

$$Q = \left(\frac{1}{2}\right)^{\frac{1+\alpha}{\alpha}} \left(\frac{2 - \gamma}{1 - \gamma}\right)^{\frac{1}{\alpha}}, \tag{10}$$

which is increasing in γ. Intuitively, a high Q is associated with a high cost of distance (a high γ). It can be seen that $Q \geq 1/2$, since $\gamma \geq 0$. Using (8) and (9), the following result is derived in Section A of the Appendix:

Result 1 *Consider two regions which are both in steady state, one fully unified ($s = 1$) and one with two states ($s = 1/2$).*

[7] For example, if we close down the first two effects (setting $\gamma = \lambda = 0$), then $Y^* = [s\tau\beta\alpha/(1 - \beta + \beta\alpha)]^{\alpha/(1-\alpha)}\,\overline{R}$, which is clearly increasing in s (as long as $\alpha > 0$), and if we close down the scale effect (setting $\alpha = 0$), then $Y^* = (1 - \gamma s)\,\overline{R}$, which is decreasing in s.

(a) If $Q \geq 1$, then output per location (Y^*) is always (weakly) higher in the fragmented region.
(b) If $Q \leq \left(\frac{\alpha}{1+\alpha}\right)\left(\frac{2-\beta+\alpha\beta}{1-\beta+\alpha\beta}\right)$, then output per location (Y^*) is always (weakly) higher in the unified region.
(c) If $\left(\frac{\alpha}{1+\alpha}\right)\left(\frac{2-\beta+\alpha\beta}{1-\beta+\alpha\beta}\right) < Q < 1$, then output per location (Y^*) is (weakly) higher in the fragmented region if, and only if,

$$\lambda \geq \frac{2\alpha}{1-\alpha}\left[\frac{(1-Q)(1-\beta+\alpha\beta)}{Q(1-\beta+\alpha\beta)-\alpha\beta}\right]. \tag{11}$$

Result 1 is illustrated in Figure 1. Part (a) describes the case when the cost of distance is so high that fragmented regions are always richer, even without technology-inducing resource competition ($\lambda = 0$). Part (b) describes the case when the cost of distance is so low that fragmented regions are always poorer, even with maximum resource competition ($\lambda = 1$). Part (c) describes the case with intermediate costs of distance. Then the fragmented region is richer if λ is large enough.

Recall that Y^* measures output when averaging across locations. Within each country, locations farther from the center always have lower output than those closer to the center. To see this, let steady state output at a location at distance d from the center of its country be denoted $\widetilde{Y}^*(d)$, derived by imposing steady state on (1). Some algebra shows that this can be written as

$$\widetilde{Y}^*(d) = \frac{\widetilde{Z}(d)}{1-\gamma s}Y^* = \frac{1-4\gamma d}{1-\gamma s}Y^*, \tag{12}$$

where we use (2). Intuitively, the mean of $\widetilde{Y}^*(d)$ across locations equals $\left[(2/s)\int_0^{s/2}\widetilde{Z}(x)dx\right] \times [Y^*/(1-\gamma s)] = Y^*$. That is, Y^* can be written as average output across locations in one half of the country. It is evident from (12) that $\widetilde{Y}^*(d)$ is decreasing in d. We can thus conclude the following:

Result 2 *Regardless of how fragmented a region is, locations farther from the center of a country have lower output.*

In sum, Result 2 states that border locations are unambiguously poorer than neighboring locations farther from the border, but when comparing regions with different degrees of fragmentation, Result 1 states that the average location in a fragmented region can be more or less developed, depending on parametric assumptions.

Figure 2 illustrates the levels of $\widetilde{Y}^*(d)$ and Y^* for two otherwise identical regions, one with $N = 8$ and one with $N = 2$, for a numerical example. Locations closer to a border tend to be poorer in both regions, but in this case the more fragmented region (the one with more borders) has higher output on average.

3.5 Discussion

The mechanisms driving both these results can be nested in more realistic frameworks. For example, technology could represent just about any public good that serves as input in both production and inter-state competition. One example could be education (cf. Aghion et al. 2015).

Letting resource competition have destructive effects through war, or resource waste, would countervail the benefits of fragmentation, but the net effects can still be positive (Lagerlöf 2014). In our highly stylized setting, we may loosely think of the costs of resource competition as incorporated in the cost of distance, the idea being that territory is more difficult to defend in the periphery than the center of a country (see Section 5.3.3 below).

There are also less destructive forms of inter-state competition than war. In a model with migration, states could compete for (skilled) labor rather than resources. In such a model, investments in technology can raise production directly, as well as indirectly by attracting labor. The results would be the same as in our current model.

In this model we compare different regions with identically sized states. A more challenging modelling task would be to allow for heterogeneity in state size within a given region, and allow some states to absorb others over time. However, the basic logic captured here would still hold if larger states face weaker incentives than smaller ones to invest in technology or public goods for defensive purposes.

At any rate, our model offers one prism through which we can interpret some of the correlations between borders and development in the data.

4 Data

4.1 Sample and border variables

This section provides a brief description of our data (see Section B of the Appendix for more details). The unit of observation is a cell with sides 0.5 degrees latitude and longitude, which corresponds to about 55.5 kilometers at the equator. This cell sized has been used by, e.g., Berman et al. (2017).

Our source for borders is Euratlas (Nüssli 2010), which supplies data on sovereign state territories at the turn of the centuries 1500-2000 (i.e., for six different years), over a large region centered on Europe. For each year we construct a border dummy equal one if more than one sovereign state was present in a cell in that year, and zero otherwise.

We restrict attention to cells covered by at least one sovereign state in all years 1500-2000. Dropping non-state areas makes sense, since we are interested in borders and state fragmentation, rather than statehood as such.[8] In other words, a border is here defined as the presence of at least two states, conditional on at least one being present.

For related reasons, while the Euratlas maps go back to 1 CE, we use 1500 as a start year. In earlier years central government capacities were often limited, and it is unclear if (sovereign) statehood meant the same as it does today. Investments in technology for the purpose of interstate resource competition, as captured by the model in Section 3, may also have been less common before 1500.

This results in a dataset of 5202 cells, with border dummies for six different years (1500 to 2000). Figure 3 shows where the 5202 cells with full statehood are located.

Some cells are partly covered by sea, what we call coastal cells, and a cell's land area also varies by latitude. This could potentially affect the cell's probability of having a border. We address this by controlling for log land area in the regressions, and often also 0.5-degree latitude fixed effects. Focusing on cells with at least one state should also mitigate these concerns.[9]

4.2 Modern outcomes

Output in our model may correspond to any disaggregate measure of economic activity. We will look at night lights and population density.

Night lights data come from the National Oceanic and Atmospheric Administration (NOAA), and refer to nighttime lights measured from satellites circling Earth several times a day. We use the log of (one plus) the average across pixels in each cell, and across all years currently available, 1992-2013, and relevant satellites.[10]

Night lights are often used as a proxy for GDP/capita, in particular where official data are missing, e.g., at the grid cell level. High correlation between night lights and official sub-national GDP/capita measures is documented by Hodler and Raschky (2014, Appendix B), and in the Online Appendix to Michalopoulos and Papaioannou (2014). Other applications using these data include, e.g., Bleakley and Lin (2012), Henderson et al. (2012), Michalopoulos and Papaioannou (2013), Berman et al. (2017), and Dickens (2017).

Population density seems like a useful outcome variable given our historical focus, and since output would translate to population density in a Malthusian setting. We use data from the Gridded Population of the World at SEDAC at NASA. To get population density we divide the cell's total population (averaged over the period 2000-2015) by the land area

[8] What determines the spread of statehood is an interesting topic in itself, and studied by others using the same Euratlas data, e.g., Harish and Paik (2016).

[9] More precisely, we estimate the probability of a cell having at least two states, conditional on having at least one. It is not obvious that this conditional probability must mechanically depend on land area.

[10] Alternative log transformations of night lights are considered in Section 7 of the Online Appendix.

of the cell. To avoid losing cells reporting zero population, we take the logarithm of one plus population density. One cell completely lacks population data, leaving us with 5201 cells when using this outcome variable.[11]

4.3 Geography variables

Our selection of geography variables is motivated by the literature discussed in Section 2.2.

We use a couple of different measures of how mountainous a territory is. Log ruggedness is the log of (one plus) the standard deviation in elevation.[12] We do not use mean elevation, since ruggedness and elevation have correlation coefficient of 0.8, and thus essentially capture the same variation. Instead, we use what we call mountain dummies, indicating whether the mean elevation of the cell exceeds 1000 or 2000 meters, respectively.

We define river and coastline density as the length of the river, or coastline, divided by the land area of the cell. Most variation in these density variables is between cells with positive and zero density, but we choose not to define them as zero-one dummy variables, to mitigate any concern about the effect that a cell's land area could have on the results.[13]

Log distance to the coast is the logarithm of (one plus) the distance in kilometers to the nearest coast.

Data on agricultural suitability come from the Global Agro-Ecological Zones (GAEZ) project and measure agricultural output when using intermediate levels of input, relative to the maximum attainable with the same inputs, under perfect environmental conditions. These data are available separately for rainfed and irrigated agriculture, and for various crops. We use the average of the most common crops—wheat, barley, oats and rye—and normalize the variables to fall between zero and one.

Rainfall data are also from GAEZ.

5 Empirical results

5.1 Descriptive statistics

From the summary statistics in Table 1, we learn that the fraction cells with a border declines monotonically from 18% in 1500 to 9% 1900, and then increases to 16% in 2000. This mirrors the overall trends in the number of sovereign states (e.g., Alesina and Spolaore 2003, Gancia

[11] This cell is located by the Black Sea.

[12] While related, ours is not exactly the same definition of ruggedness as that in Nunn and Puga (2012), but rather similar to what Michalopoulos (2012) calls variation in elevation.

[13] The results do not change qualitatively when using river and coast dummies instead. See Section 3 of the Online Appendix.

et al. 2016). The map in Figure 3 illustrates the decline between 1500 and 1900; note, e.g., the unifications of Germany and Italy.

To analyze the cross-sectional correlation between borders, geography, and modern outcomes, we construct a variable which we refer to interchangeably as the border index, or border frequency. It is simply the fraction of the six years (1500 to 2000) in which a cell had a border. Averaging across centuries in this way should alleviate any concerns about measurement error and the changing roles of state borders. Letting $b_{i,t}$ be the border dummy, indicating if a border was present in cell i and year t, this index can be written

$$B_i = \frac{1}{6}\sum\nolimits_{t=1500}^{2000} b_{i,t}. \tag{13}$$

As reported in Table 1, B_i shows positive correlation with each of the year-specific border dummies ($b_{i,t}$).[14] In that sense, our border frequency index is not capturing the distribution of borders in any particular year, or set of years.

5.2 Regressing borders on geography

Before we study how border frequency correlates with modern outcomes, we explore if the geographical factors discussed in Section 2.2 influence border locations. Documenting that geography correlates with borders also helps making sense of the model in Section 3, where borders were treated as exogenous to agents' choices.

The map in Figure 4 shows that locations with mountains and rivers often coincide with those with high border frequency.

Figure 5 shows the means of some geography variables for different levels of the border index. (The unconditional correlation coefficients are reported in Table 1.) Among cells with a border present in all six years ($B_i = 1$), more than 8% have mountains above 2000 meters, and the corresponding number for cells with no borders in any year ($B_i = 0$) is about 1%. River density, log ruggedness, and rainfall also increase with border frequency.

In our regression analysis, the baseline regression specification is

$$B_i = \alpha + \mathbf{G}_i\boldsymbol{\beta} + \varepsilon_i, \tag{14}$$

where i indicates the cell, B_i is the border index in (13), $\mathbf{G}_i$ a vector containing different geography variables, $\boldsymbol{\beta}$ a vector of the coefficients of interest, α a constant, and ε_i the error term.

Table 2 shows the results when estimating variations of (14) with ordinary least squares. Standard errors are adjusted to control for spatial correlation using the Conley (1999) method, assuming correlation among cells located within 1.45 degrees of each other. This

[14] The border dummies are also positively correlated with each other; see Section 2 in the Online Appendix.

captures two neighboring cells in all eight directions (north, south, east, west, and four diagonal directions).

Recall that cells have different land areas, so all specifications control for the (log) land area of the cell.

Columns (1)-(3) in Table 2 confirm that borders are more frequent in cells with mountains and rugged terrain. The two mountain dummies and log ruggedness to some extent measure the same variation, but mountains over 2000 meters and log ruggedness both come out as positive and significant when entered together; see column (3).

Column (4) confirms that cells with higher river density also have more borders.

Columns (5)-(7) show that areas suitable for rainfed agriculture have more borders, while those more suitable for irrigated agriculture have fewer. This holds when these are entered both separately in columns (5) and (6), and together in column (7). To help us understand why, the map in Figure 6 shows that, e.g., the Iberian peninsula has both below-median rainfed suitability, and above-median irrigated suitability. It also has relatively few borders, despite being mountainous and having several rivers (cf. Figure 4), so its agricultural suitability profile partly accounts for its unification. As discussed in Section 2.2, this is reminiscent of the theories of Wittfogel (1957) about the role of irrigation for the rise of despotic states. Translated to our context, smaller states may have survived more easily when external powers could not control water supply.

When controlling for rainfall in column (8) the estimated coefficient on suitability for rainfed agriculture shrinks in size, but stays precisely estimated, while rainfall itself comes out as positive and significant. While the two are obviously connected, as discussed in Section 2.2, suitability for rainfed agriculture and rainfall do measure slightly different things. Moreover, the coefficient on irrigated suitability is almost unchanged, and all three coefficients are precisely estimated when entered together in column (8).

Column (9) adds coastline density and (log) distance to the coast. While the associated coefficients do not come out as significant in this specification, the negative sign on distance to the coast, and the positive sign on coastline density, seem to capture that eastern inland areas are more unified. Notably, the estimates are more significant with latitude fixed effects in Table 3 below. This is interesting, because our definition of borders (two or more sovereign states entering the same cell) allows us to capture mostly land borders. For example, most of Britain's coast is not a border, except for cells by the English Channel containing both France and England. The effects of coastline density and distance to the coast may partly capture that regions that are more cut off by the sea, like peninsulas, constitute natural states, making their land connections more suitable as borders. The Iberian peninsula could be an example.

In what follows, we shall refer to the independent variables in column (9) as our benchmark set of geography controls.

5.2.1 Quantifying the effects

To get a sense of magnitudes, consider, e.g., ruggedness. The standard deviations in log ruggedness and border frequency equal 1.28 and 0.24, respectively (see Table 1). Using the coefficient on log ruggedness in column (9) of Table 2 (0.009) the associated standardized (or beta) coefficient becomes $0.009 \times 1.28/0.24 \approx 0.048$. That is, a one standard deviation increase in log ruggedness raises border frequency by 0.048 standard deviations. The corresponding values for the other variables in column (9) can be computed as follows (standardized coefficients in parentheses): 2000 meter mountain dummy (0.070); river density (0.113); suitability for rainfed agriculture (0.102); suitability for irrigated agriculture (-0.132); rainfall (0.241); log distance to coast (-0.058); coastline density (0.029).[15]

Some of these numbers may seem small, but each of them refers to a partial change, holding constant all the other benchmark geography controls. Suppose instead that we simultaneously change each of these eight geography variables by one standard deviation in the direction that raises border frequency. Then border frequency increases by 0.794 standard deviations (i.e., the sum of the eight standardized coefficients in absolute terms). In other words, the fragmenting effects of geography seem to work along multiple dimensions, rather than just a few variables.

We also note that geography alone generates an R-squared of at most 12% in column (9), so it is hardly the *only* factor determining state fragmentation, at least by the measures of geography and fragmentation used here. However, its explanatory power is not negligible either.

5.2.2 Predicting borders

Figure 7 illustrates where geography predicts borders to be located. Out of 5202 cells in our data, 70 had borders in all six years 1500-2000 ($B_i = 1$). The map in Panel (a) shows the locations of these cells, as well as the 70 cells predicted to have the highest border frequency based on the regression in column (9) of Table 2. Panel (b) does the same for 454 cells with borders in at least four of the six years ($B_i \geq 2/3$).

Geography successfully predicts high border frequency around the Alps and in the Pyrenees, but the Caucasus region has been more unified than geography predicts. One reason could be its relative vicinity to territorially expansive states, such as the Ottoman Empire and Russia.

[15] We here ignore log land area, which we interpret as a control rather than a measure of geography.

5.2.3 Robustness checks

Table 3 explores further what drives the correlations between borders and geography. Column (1) is identical to column (9) in Table 2, but adds latitude fixed effects, i.e., a full set of dummies for each of the half-degree latitudes in our sample. As mentioned above, coastline density and distance to the coast now come out as slightly more significant, as those correlations capture variation along latitudes. We also note that the coefficient on rainfed agriculture, which is somewhat unstable across regressions, loses significance with latitude fixed effects.

In column (2) we use the 1000-meter mountain dummy, instead of the 2000-meter one, with qualitatively similar results. This is useful to note for later comparisons.

In column (3) we construct an alternative measure of borders, by letting the Holy Roman Empire (HRE) be defined as one single country in the years it existed (1500-1800); see Section B.1 in the Appendix. The results are surprisingly similar to those in column (1), using the same specification without HRE adjustment. Intuitively, the changes in the border index when HRE adjusting are not strongly correlated with most of the geography variables.[16] By contrast, when regressing modern outcomes on borders in Section 5.3, HRE adjustment does affect the results.

Column (4) reverts to the baseline border measure (without HRE adjustment), but drops all coastal cells. With this restricted sample, land area only varies with latitude, and since we include latitude fixed effects, the land area control is automatically omitted, and so is coastline density. The results for the remaining variables are mostly unchanged.

Columns (5) and (6) extend the specifications in columns (1) and (2) to include longitude, and well as latitude, fixed effects (i.e., dummies for each half-degree row and each half-degree column in our grid). Again, the results are broadly unchanged, although distance to the coast again comes out as insignificant. Intuitively, the longitude fixed effects absorb unification in the eastern, more inland, sections of our sample. We also note that the estimated coefficient on rainfed agricultural suitability is here once again positive and significant.

The Online Appendix explores the links between geography and borders further. We measure borders in other ways, using, e.g., ethno-linguistic borders, and state borders from another source (Abramson 2017). We explore alternative geography variables, in particular other measures of agricultural suitability, e.g., the Caloric Suitability Index by Galor and Özak (2016), and suitability for potato used by Nunn and Qian (2011). We also present results with other measures of state fragmentation than borders, and explore alternative ways to deal with spatial correlation. The results discussed above are broadly robust, at least where we would expect them to be.

[16] See Section 2 of the Online Appendix.

5.2.4 Local deviations in border frequency

One possible concern is that some geography variables, such as mountains and rivers, tend to cluster (or form connected lines), as do borders. Therefore it is hard to know if borders are present in a cell because of the geography in that cell itself, or because of the geography and/or border presence in neighboring cells. Latitude and longitude fixed effects may address this, if such clustering is mostly a north-south and/or east-west phenomenon.

Next, we propose a different approach, by regressing local deviations in border density on local deviations in our geography variables. Each cell in our dataset can have at most eight adjacent neighboring cells that are also in the data (to the east, west, north, south, and in one of the four diagonal directions); fewer than eight if some neighbors are sea cells or stateless. Let $\mathcal{N}_i$ be the set of indices of the (eight or fewer) adjacent neighbors of cell i in the data (i.e., excluding sea cells or non-state cells), and let $\#(\mathcal{N}_i) \leq 8$ be the number of such cells. Furthermore, let $G_{k,j}$ be the value of some geography variable k in cell j. Then

$$G_{k,-i} = \frac{1}{\#(\mathcal{N}_i)} \sum\nolimits_{j \in \mathcal{N}_i} G_{k,,j} \tag{15}$$

is the mean of geography variable k among cell i's neighbors. Similarly, we let

$$B_{-i} = \frac{1}{\#(\mathcal{N}_i)} \sum\nolimits_{j \in \mathcal{N}_i} B_j \tag{16}$$

be mean border frequency among the same neighboring cells. Now, using (15) and (16), we can define

$$\begin{aligned} \Delta B_i &= B_i - B_{-i}, \\ \Delta G_{k,i} &= G_{k,i} - G_{k,-i}, \end{aligned} \tag{17}$$

which we shall refer to as the local deviations in the border index, and geography variable k, respectively, both for cell i. The regression equation can be written

$$\Delta B_i = \alpha^{\Delta} + \Delta \mathbf{G}_i \boldsymbol{\beta}^{\Delta} + \varepsilon_i^{\Delta}, \tag{18}$$

where $\Delta \mathbf{G}_i$ is a vector with elements $\Delta G_{k,i}$, α^{Δ} and $\boldsymbol{\beta}^{\Delta}$ are vectors of coefficients, and ε_i^{Δ} is an error term.

The results from this regression are shown in Table 4, using the same independent geography variables as in our benchmark specifications in columns (1) and (2) of Table 3. We show results without any fixed effects in columns (1)-(2), while latitude fixed effects are included in columns (3)-(7). We treat the Holy Roman Empire as unified in column (4), drop coastal cells in column (5), and include longitude fixed effects in column (6).

To interpret the results in Table 4, consider first the estimated coefficient on log ruggedness. In most columns, this is positive and precisely estimated. This means that places that were more rugged than their surroundings also had more borders than those surroundings. Put differently, borders seem to gravitate locally towards more rugged locations.

Mountains also tend to attract borders at the local level, although the results now hold only for mountains (i.e., mean elevation) above 1000 meters, not 2000 meters; see columns (1) and (2). The map in Figure 4 may help us understand why. Relatively few mountains exceed 2000 meters, some in the border-dense Alps, and others in the more unified Caucasus; the Alps drive the positive effect at the global level. Our local-deviation analysis allows smaller mountains to have an effect, since these can attract borders in regions that are otherwise relatively flat and unified, such as the Pyrenees, but not in regions where they are surrounded by other mountains, such as the Caucasus.

Distance to the coast comes out as negative, but mostly insignificant, and coastline density as positive and mostly significant. This also broadly mirrors the results in Tables 2 and 3.

Because these geography variables—ruggedness, mountains over 1000 meters, rivers, rainfall, and distance to the coast—carry the same sign in terms of local deviations as in the "global" analysis, it appears that these features of the landscape constitute some sort of natural state borders. By extension, it also suggests that political fragmentation may have partly resulted from an abundance of such natural borders.

However, cells with higher suitability for rainfed agriculture than neighboring cells have *fewer* borders than those neighbors, even without latitude fixed effects; see columns (1) and (2). This is the opposite of what we found in the "global" analysis. (For suitability for irrigated agriculture, we find no significant local effects.) The pattern at the global level may thus reflect macro factors: states are smaller in regions which are overall more suitable for rainfed agriculture, essentially Western Europe, where rulers have lacked the tools that come with irrigation infrastructure to dominate their populations. At the same time, larger states in Western Europe seem to have formed in areas more suitable for rainfed agriculture. For example, south-eastern England has higher suitability for rainfed agriculture, and fewer borders, than Scotland; cf. Figure 6.

5.2.5 Geography and border dynamics

In this section, we utilize the time variation in our border dummies. Recall that the declining trend in border frequency from 1500 to 1900 was reversed between 1900 and 2000, possibly reflecting factors unique to the 20th century and the region we study. Here we consider the period from 1500 to 1900.

Column (1) of Table 5 regresses the border dummy in one century on the same dummy in the previous century, and the benchmark set of geography controls and latitude and century fixed effects. The positive and significant coefficient on the lagged border dummy shows that borders are highly persistent from one century to the next.

Column (2) adds cell fixed effects (and drops all controls absorbed by the cell fixed effect). Now the coefficient on the lagged border dummy comes out as negative and significant. Intuitively, this amounts to regressing the change in borders on the lagged change in borders, and these changes tend to go in different directions; after gaining or losing a border the change can only go in the opposite direction.

Column (3) adds the log of the ratio of the largest state's size over the smallest state's size in the cell. The negative sign on this coefficient means that borders separating larger states from small are less likely to survive. This is interesting, because it suggests that initial unification can spur further unification. As one state expands by absorbing neighboring states, holding constant the size of the non-absorbed states, the size gap increases among its border cells, making the expanding state more likely to absorb more neighbors. Thus, geographical factors which induce earlier transitions to statehood should be associated with less fragmentation, all else equal.

Column (4) includes an interaction term between the log size gap and log ruggedness, which comes out as positive and significant, meaning that ruggedness reduces the magnitude of the direct effect from the state size gap. Evaluated at the sample mean of log ruggedness of 4.12 (see Table 1), the marginal effect from the log size gap is $-0.089+0.015\times4.12 \approx -0.027$, and becomes zero in cells with log ruggedness of about 6, around the 95th percentile of our sample. Thus, the most rugged locations are just about fully insulated from the unifying effects of having a much larger neighbor. Indicatively, some of the world's smallest countries today are located in the mountains between large states, such as Andorra in the Pyrenees between France and Spain. The CIA World Factbook (CIA 2013) describes its terrain as "rugged mountains dissected by narrow valleys" and reports a mean elevation of almost 2000 meters.

Column (5) reports the interaction effects between the log size gap and rainfall. Just like rugged terrain, ample rainfall reduces the magnitude of the direct negative effect of the state size gap. As discussed, this need not be due to rainfall itself, but could capture other features of the landscape that correlate with rainfall.

5.3 Regressing modern outcomes on borders

The results so far support the notion that geography exerts some effect on borders. This is conceptually consistent with how we treated borders as exogenous in our model in Section 3. Of course, borders in the data are not constant over time, and states are never symmetric, or identical in size, but we can think of certain broad regions as being more fragmented than others partly because of differences in geography.

Next we examine how the correlations between borders and economic development in the data compare to the predictions of our model. We use night lights and population density to measure modern development outcomes. As discussed in Section 4.2, night lights often proxy for GDP per capita, or economic activity more broadly. Alternatively, both measures might represent population density in preindustrial times, which in a Malthusian environment should be positively related to total output and productivity.[17]

This section studies the global correlations, i.e., correlations across all cells, which can be compared to Result 1 in the model. Section 5.3.2 below looks at the correlations in terms of local deviations, which can be compared to Result 2.

The bar graphs in Figure 8 suggest a positive relationship between border frequency and modern development outcomes. Figure 9 shows a map of the locations of the 454 cells which had a border in at least four of the six years 1500-2000 ($B_i \geq 2/3$), and the 454 cells that recorded the highest night lights. While it is clear that border frequency cannot explain all spatial variation in night lights, there is some suggestive overlap, especially around Germany and the Alps.

Letting $\ln Y_i$ denote the logarithm of (one plus) either night lights or population density, we can write the baseline regression specification as

$$\ln Y_i = \gamma + \delta B_i + \mathbf{G}_i \boldsymbol{\lambda} + \epsilon_i, \tag{19}$$

where i indicates the cell, B_i is the border index, $\mathbf{G}_i$ and $\boldsymbol{\lambda}$ are vectors with geography controls and coefficients, γ is a constant, ϵ_i is an error term, and δ is the main coefficient of interest. Controlling for geography should mitigate concerns that borders pick up variation in night lights and population density caused directly by geography. For example, Figure 9 shows that night lights are high in coastal areas, so it makes sense to control for coastline density.[18]

Panel A of Table 6 uses the log of (one plus) night lights as the dependent variable, and Panel B the log of (one plus) population density.

Consider first the results for night lights. Column (1) shows a positive and significant effect of border frequency, without any controls. Column (2) adds benchmark geography controls used when regressing borders on geography (including log land area); see column (9) in Table 2. The positive correlation holds, suggesting that the results are not due to any direct effects of geography.

[17] An alternative approach is to use night lights as the dependent variable, while controlling for population density, thus estimating the effects of border frequency on night lights per capita (see Section 7.1 of the Online Appendix). Here we rather think of population density and night lights as two different measures of long-run economic development.

[18] It is well known that coastal areas have more economic activity. For the case of the United States, see Rappaport and Sachs (2003).

Column (3) adds latitude fixed effects. The correlation stays positive and significant, although the size of the coefficient shrinks a little. The results are thus not driven only by north-south variation in night lights and borders.

Column (4) shows that the coefficient on the HRE adjusted border frequency (i.e., treating the Holy Roman Empire as unified) is positive but not significant, at least with geography controls and latitude fixed effects. This is not too surprising, since some of the most developed areas in the world today belonged to the HRE, including Germany, Luxembourg, and Switzerland. Arguably, if we believe the mechanism is about state competition, it makes sense not to treat the HRE as unified, since states and cities within the HRE fought and competed with one another (see, e.g., Huning and Wahl 2017). However, we note that the positive effect of historical fragmentation on modern development is sensitive to how we interpret the HRE.

Columns (5) and (6) revert back to the benchmark border frequency index, keeping geography controls and latitude fixed effects as in column (3). In column (5) we drop coastal cells, finding that border frequency still has a significant effect on night lights. Column (6) adds longitude fixed effects (keeping coastal cells). This renders the correlation insignificant again, which illustrates that the positive correlation is (at least partly) driven by east-west variation in borders and night lights.

Panel B of Table 6 shows results from the same regression specifications, but with (log) population density as the dependent variable. The results resemble those for log night lights, but are marginally stronger.

5.3.1 Quantifying the effects

Using the point estimate in column (2) of Panel A of Table 6, the predicted gap in (non-logged) night lights between two cells with maximum and minimum border frequency (one and zero, respectively) is about $\exp(0.563) \approx 1.76$. That is, a cell that has a border at the turn of every century 1500-2000 is predicted to emit 76% more night lights than one that is unified over the same period, holding geography constant. Using the estimate in column (2) of Panel B, population density varies between the same two cells by a factor of $\exp(0.856) \approx 2.35$.

These seem like large numbers, but the overall variation in the data is even larger; as a reference point, non-logged night lights across the whole sample of cells vary by a factor of approximately 60. Moreover, these comparisons apply the maximum difference in border frequency. Using the point estimate in column (2) of Panel A of Table 6, and the standard deviations in log night lights and border frequency in Table 1 (0.94 and 0.24, respectively), we can calculate that an increase in border frequency by one standard deviation raises log night lights by 0.14 standard deviations ($0.563 \times 0.24/0.94 \approx 0.14$). The corresponding effect on population density is roughly the same in size ($0.856 \times 0.24/1.48 \approx 0.14$).[19] We do not suggest that border frequency is the *only* factor explaining gaps in development, but these effects are not negligible either.

5.3.2 Local deviations

In Table 7 we report results from regressions similar to those in Table 6, but now regressing local deviations in log night lights and log population density on local deviations in the border frequency, and latitude fixed effects. The structure is otherwise identical to that in Table 6.

In columns (1)-(6) of Panel A, the dependent variable is the difference between log night lights in one cell and the same variable in its eight closest neighboring cells. Column (1) shows the results without any controls, column (2) adds controls for local deviations in the benchmark geography variables, column (3) adds latitude fixed effects (not in local deviations), column (4) uses local deviations in HRE adjusted border frequency, column (5) drops coastal cells, and column (6) adds longitude fixed effects (not in local deviations).

Interestingly, in all specifications we find that if a cell has a higher border frequency compared to its eight closest neighbors, it tends to emit *less* light. That is, the local pattern is the reverse of what we saw at the global level in Table 6.

Panel B of Table 7 shows the corresponding regression results when using local deviations in log population density as the dependent variable. Again, the negative correlation holds in all specifications.

One might think that these results simply reflect the underlying geography. For example, high mountains have more borders than flat areas, and may also have less night lights and lower population densities due to, e.g., higher transport costs in such terrain. However, the negative correlation holds when we control for local deviations in mountains and ruggedness, and the other benchmark geography variables.

In sum, cells in fragmented regions are more developed than unified ones, but conditional on a cell being located in a fragmented region, there is no benefit to being right at the border, rather than one cell removed. This is consistent with the model in Section 3, where output per location can be higher or lower in more fragmented regions (Result 1), while the richest locations are always at the center of each state (Result 2).[20]

[19] Expressed differently, a one standard deviation increase in border frequency raises (non-logged) night lights by a factor of $\exp(0.24 \times 0.563) \approx 1.14$ and (non-logged) population density by a factor of $\exp(0.24 \times 0.856) \approx 1.23$, i.e., by 14% and 23%, respectively.

[20] The Online Appendix shows that the same regression results hold when using local deviations in the cell's historical distance to a border as the independent variable, instead of the cells' border frequency. That is, cells that have been closer to borders 1500-2000 have worse modern outcomes. This may be more in line with the model, where output decreases continuously with distance from the center.

The positive effect at the global level seems to capture regional differences between Western Europe on the one hand and more unified regions, like today's Russia, Ukraine, and Middle East, on the other; cf. Figure 9. To understand the negative local correlation, note that some big capital cities within the rich and fragmented Western Europe—such as Madrid and Paris in the interior of Spain and France, respectively—have higher levels of night lights than many border regions.

5.3.3 Border stability

Recall from Result 1 in our model that fragmented regions are more likely to be more developed when there is more competition between states (higher λ). Of course, state competition has also carried costs. If state competition did play a role, we should thus expect to see better outcomes where those negative effects of state competition were milder.

Below we explore this further by regressing modern outcomes on border stability. The idea is that more stable borders, all else equal, should be associated with less negative effects from state competition, e.g., less warfare and foreign occupation.[21] First we define average border change, C_i, as the frequency with which cell i has switched between having a border, and not having one, from 1500 to 2000, i.e.,

$$C_i = \frac{1}{5} \sum\nolimits_{t=1600}^{2000} (b_{i,t} - b_{i,t-1})^2. \tag{20}$$

Table 8 shows the results from regressing log night lights and log population density on the border frequency index used previously, and defined in (13), and the new border change variable in (20). Because the way these are defined from the same border dummies, they have an inversely U-shaped relationship in the data. That is, border change takes its minimum value of zero both in fully unified cells (with no borders) and in cells that always have borders.[22]

Column (1) of Table 8 replicates the result in column (2) of Table 6, showing a positive and significant effect from border frequency on log night lights when controlling for our benchmark set of geography variables. In column (2), we run the same regression but with border change replacing border frequency. We find a positive but insignificant effect from border change. However, when entering both independent variables together in column (3), border change has a negative effect, and border frequency a positive effect, both highly significant. Also, the coefficient on border frequency is larger than when entered alone in column (1). This holds also when adding latitude fixed effects in column (4).

[21] This obviously does not fit with a literal interpretation of our model, where borders are exogenous and constant over time. However, we could possibly think of the cost-of-distance parameter (γ) as related to border change, or other negative effects of state fragmentation.

[22] The results are very similar when using the variance in borders; see Section 5 of the Online Appendix.

Columns (5)-(8) show the results when using log population density as the dependent variable. The results are similar to those for night lights, although border change now comes out as positive and significant on its own in column (6).

To see what drives these results, note that the best predicted development outcomes can be found in cells that have had a border all years 1500-2000, i.e., with a border frequency of one and border change of zero. As shown in Figure 10, these are located, e.g., along the borders of Switzerland, along Germany's borders to Austria and the Czech republic, and in the Pyrenees separating France and Spain. These are all places with relatively high levels of development. The least stable border areas can be found in poorer areas, e.g., Moldova and eastern Poland. This is consistent with the idea that border change proxies for the costs of fragmentation, seemingly linked to military conflict, and/or the rise and fall of a few empire-like states, such as the Austria-Hungarian and Ottoman Empires, and Russia/USSR.[23]

For example, one of the highest levels of border change is found in the cell containing the city of Tiraspol at 47 degrees latitude and 29.5 degrees longitude. According to our data, this cell was unified in 1700 and 1900 (under the Ottoman Empire and Russia, respectively), and a border cell all other years. It also lies in the breakaway territory of Transdniestria, which fought an independence war against Moldova in 1992 (The Economist 2014).

5.3.4 Urbanization

Levels of night lights and population density are higher in cities than rural areas; as mentioned, capitals are often located at the center of countries. This suggests that the effect of borders on modern development, in particular in terms of local deviations, may relate to urbanization.

To explore this, we utilize the History Database of the Global Environment (HYDE), which provides spatially disaggregated data on historical urban and rural population (Klein Goldewijk et al., 2010, 2011). We calculate the urbanization rate as urban population over total population. Because we have spatial and temporal variation in both urbanization rates and border dummies we are able to utilize the panel structure of our data.

Table 9 shows the results when regressing the urbanization rate on the previous century's urbanization rate and the lagged and current border dummy. Columns (1)-(3) control for century and latitude fixed effects, and the benchmark set of geography variables, while

[23] This interpretation might ostensibly contrast with the findings of Dincecco and Onorato (2016) that conflicts in European history have often promoted city growth through a so-called safe-harbor effect. However, they find weaker effects on city growth from sieges than battles occurring outside cities; the former had more destructive effects. Moreover, they study the period 800-1800. The following two centuries probably saw more devastating warfare.

columns (4)-(6) enter cell fixed effects (and drop latitude and geography controls), which amounts to a difference-in-differences estimator.

The coefficient on the current border dummy is negative and significant in all specifications, while the coefficient on the lagged border dummy is positive but much smaller and less significant. This indicates a negative simultaneous effect on city growth from border change, but little effect from being initially endowed with a border. The simultaneous effect seems to be driven by variation among cell-years with initial borders present: cells that keep borders urbanize less than those that become unified. One example is the unification of Germany between 1800 and 1900, at a time of faster-than-average urbanization in this region. The direction of causality is not obvious, but one interpretation is that such unifications led to city growth in the interior, e.g., by lowering trade barriers within the new state (cf. Gancia et al. 2016).

5.3.5 Alternative theories: culture or institutions

The correlations we have documented seem suggestive of a causal relationship running from geography to state fragmentation, and on to modern development. Some regions have more natural borders, and thus more states, which in turn impacts long-run development. Section 3 tried to capture this in a simple model of state competition.

What competing theories could explain these patterns? Perhaps some other factor than geography, such as inclusive institutions, a growth-promoting culture, or initial conditions—which we cannot measure well at the cell level—might hinder the expansion of empires. For example, rulers in Western Europe may have been more institutionally constrained than rulers elsewhere in terms of pursuing territorial conquests, and those same institutions might have promoted development. That could explain the positive global correlation between border frequency and modern development.

However, if culture or institutions were the fundamental cause of state fragmentation, it is not clear why border frequency, and other measures of fragmentation, should be correlated with geography at all. Another argument against a theory based on culture or institutions as the fundamental cause of fragmentation is that whatever factors made Europe itself politically fragmented did not prevent it from building empires elsewhere.

We do not think that culture and institutions are irrelevant, but rather that these factors are best thought of as the endogenous outcomes of state fragmentation, since state competition can induce elites to undertake, e.g., political and educational reform. While our proposed model in Section 3 interpreted competition through investments in "technology" this may just be shorthand for a whole vector of growth promoting actions.

6 Conclusion

Compared to many other regions of the world, Europe has historically been highly politically fragmented and shaped by continuous interstate conflict and competition. Some suggest that this has contributed to Europe's unique long-run development path. Others propose that Europe's high degree of state fragmentation is fundamentally caused by its geography.

In this paper, we document a few patterns that we think can shed light on this. We utilize a dataset of 5202 grid cells, covering a region encompassing Europe, Western Asia, and North Africa, to explore two issues: how geography correlates with the location of borders, and how population density and night lights in modern times correlate with those border locations.

Several geography variables show strong and robust correlation with borders, in particular mountains, ruggedness, rainfall, and rivers. Borders separating large and small states are less likely to survive over time, in particular in non-rugged terrain, suggesting that initial unification can spur more unification.

We also document that historical border presence has positive correlation with modern outcomes. This seems to be due to the macro-level benefits of state competition. In particular, the positive correlation is reversed at the local level: cells with more historical borders than their neighboring cells are rather less developed today. In other words, conditional on being located in a fragmented region, like Western Europe, there are no benefits to being right at a border.

We also present a simple model predicting that regions with more state fragmentation may be richer even though border locations are always poorer than locations at the center of a country.

As a final note, we emphasize that we are not trying to put forward geography as the *only* factor determining the shape of states, or state fragmentation as the *only* factor behind Europe's unique development path. The explanatory power is not too large in any of our regressions, and the quantitative implications not huge either. At the same time, we may be underestimating some effects. For example, we essentially measure only land borders, not fully capturing the effects of Europe's indented coastline. Similarly, state fragmentation can be measured in other ways than just border frequency. Our Online Appendix shows the results when using several alternative measures, such as distance to a border or the size of the state(s) intersecting the cell, but a more systematic analysis is left for future research.

References

1. Abramson, S.F. (2017). "The Economic Origins of the Territorial State." *International Organization*, 71, 97-130.

2. Abramson, S. F., and D. B. Carter (2016). "The Historical Origins of Territorial Disputes." *American Political Science Review*, forthcoming.

3. Acemoglu, D., S. Johnson, and J. A. Robinson (2001). "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*, 91, 1369-1401.

4. Acemoglu, D., S. Johnson, and J. A. Robinson (2002). "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics*, 117, 1231-1294.

5. Aghion, P., X. Jaravel, T. Persson, and D. Rouzet (2015). "Education and Military Rivalry." *Mimeo.*

6. Ahlerup, P. and O. Olsson (2012). "The Roots of Ethnic Diversity." *Journal of Economic Growth*, 17, 71-102.

7. Alesina, A., (2003). "The Size of Countries: Does it Matter?" *Journal of the European Economic Association*, 1, 301-316.

8. Alesina, A., W. Easterly, and J. Matuszeski (2011). "Artificial States." *Journal of the European Economic Association*, 9, 246-277.

9. Alesina, A. and E. Spolaore (2003). *The Size of Nations.* Cambridge: The MIT Press.

10. Alesina, A., E. Spolaore, and R. Wacziarg (2005). "Trade, Growth and the Size of Countries." In Philippe Aghion and Steven Durlauf (eds.), *Handbook of Economic Growth*, Vol. 1, Part 2, Chapter 23, 499-1542. Amsterdam: North Holland.

11. Ashraf, Q. and O. Galor (2013) "The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development." *American Economic Review*, 103, 1-46.

12. Barfield, T.J. (1989). *The Perilous Frontier: Nomadic Empires and China.* Cambridge: Basil Blackwell Inc.

13. Bentzen, J. S., N. Kaarsen, and A. M. Wingender (2016). "Irrigation and Autocracy." *Journal of the European Economic Association*, forthcoming.

14. Berman, N., M. Couttenier, D. Rohner, M. Thoenig (2017). "This Mine Is Mine! How Minerals Fuel Conflicts in Africa." *American Economic Review*, 107, 1564–1610.

15. Besley, T. and T. Persson (2011). *Pillars of Prosperity: The Political Economics of Development Clusters.* Princeton: Princeton University Press.

16. Bleakley, H., and J. Lin (2012). "Portage and Path Dependence." *Quarterly Journal of Economics*, 127, 587–644.

17. Bolton, P. and G. Roland (1997). "The Breakup of Nations: A Political Economy Analysis." *Quarterly Journal of Economics*, 112, 1057-1090.

18. Borcan, O., O. Olsson, and L. Putterman (2015). "State History and Economic Development: Evidence from Six Millennia." *Mimeo.* Brown University and University of Gothenburg.

19. Brigham, A. P. (1919). "Principles in the Determination of Boundaries." *Geographical Review*, 7, 201-219.

20. CIA (2013). *The World Factbook 2013-14.* Washington, DC: Central Intelligence Agency.

21. Conley, T. G. (1999). "GMM Estimation With Cross Sectional Dependence." *Journal of Econometrics*, 92, 1-45.

22. Cosandey, D. (1997). *Le Secret de l'Occident.* Paris: Arléa.

23. Curzon of Keddleston, Lord (1907). *Frontiers* (Romanes Lecture). Oxford: Clarendon Press.
https://www.dur.ac.uk/resources/ibru/resources/links/curzon.pdf

24. Diamond, J. M. (1997). *Guns, Germs, and Steel: The Fates of Human Societies.* New York and London: W.W. Norton & Company.

25. Dickens, A. (2017). "Ethnolinguistic Favoritism in African Politics." *Mimeo.* Brock University.

26. Dincecco, M. and M. Prado (2012). "Warfare, Fiscal Capacity, and Performance." *Journal of Economic Growth*, 17, 171-203.

27. Dincecco, M. and M.G. Onorato (2016). "Military Conflict and the Rise of Urban Europe." *Journal of Economic Growth*, 21, 259-282.

28. Elis, R., S. Haber, and J. Horrillo (2017). "Climate, Geography, and the Evolution of Economic and Political Systems." *Mimeo.* Stanford University.

29. Ferguson, N. (2009). *The Ascent of Money: A Financial History of the World.* New York: Penguin Books.

30. Friedman, N. (1977). "A Theory of the Size and Shape of Nations." *Journal of Political Economy*, 85, 59-77.

31. Galor, O. and Ö. Özak (2016). "The Agricultural Origins of Time Preference." *American Economic Review*, 106, 3064-3103.

32. Gancia, G., G. A. M. Ponzetto, and J. Ventura (2016). "Globalization and Political Structure." *NBER Working Paper*, 22046.

33. Gennaioli, N. and H.-J. Voth (2015). "State Capacity and Military Conflict." *Review of Economic Studies*, 82, 1409-1448.

34. Guiso, L., P. Sapienza, and L. Zingales (2016). "Long-term Persistence." *Journal of the European Economic Association*, 14, 1401-1436.

35. Halperin, S., (2004). *War and Social Change in Modern Europe: The Great Transformation Revisited*. Cambridge: Cambridge University Press.

36. Harish, S.P., and C. Paik (2016). "State and Development: A Historical Study of Europe from 1 AD to 2000 AD." *Mimeo.* New York University and NYU Abu Dhabi.

37. Henderson, J.V., A. Storeygard, D.N. Weil (2012). "Measuring Economic Growth from Outer Space." *American Economic Review*, 102, 994-1028.

38. Hodler, R., and P. A. Raschky (2014). "Regional Favoritism." *Quarterly Journal of Economics* 129 (2), 995-1033.

39. Hoffman, P.T. (2015). *Why Did Europe Conquer the World?* Princeton and Oxford: Princeton University Press.

40. Holdich, T. H. (1916). *Political Frontiers and Boundary Making*. London: MacMillan & Co.

41. Huning, T.R., and F. Wahl (2017). "You Reap What You Know: Origins and Dynamics of State Capacity." *Mimeo.* Humboldt-Universität zu Berlin and University of Hohenheim.

42. Jones, E. (1981). *The European Miracle: Environments, Economies and Geopolitics in the History of Europe and Asia*, 3rd ed. Cambridge: Cambridge University Press.

43. Karayalçin, C. (2008). "Divided We Stand, United We Fall: The Hume-North-Jones Mechanism for the Rise of Europe." *International Economic Review*, 49, 973-997.

44. Klein Goldewijk, K. , A. Beusen, M. de Vos and G. van Drecht (2011). "The HYDE 3.1 Spatially Explicit Database of Human Induced Land Use Change Over the Past 12,000 Years." *Global Ecology and Biogeography*, 20, 73-86.

45. Klein Goldewijk, K. , A. Beusen, and P. Janssen (2010). "Long Term Dynamic Modeling of Global Population and Built-Up Area in a Spatially Explicit Way: HYDE 3.1." *The Holocene*, 20, 565-573.

46. Ko, C. Y., M. Koyama, and T.-H. Sng (2017). "Unified China and Divided Europe." *International Economic Review*, forthcoming.

47. Lagerlöf, N.-P. (2014). "Population, Technology and Fragmentation: The European Miracle Revisited." *Journal of Development Economics*, 108, 87-105.

48. Landes, D. S. (1998). *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor.* New York and London: W. W. Norton & Company.

49. Lowe, M. (2014). "Night Lights and ArcGIS: A Brief Guide." *Mimeo.* MIT. http://economics.mit.edu/files/8945

50. Mayshar, J., O. Moav, Z. Neeman, and L. Pascali (2015). "Cereals, Appropriability and Hierarchy." *Mimeo.* Hebrew University of Jerusalem, University of Warwick, and Tel Aviv University.

51. McNeill, W. H. (1982). *The Pursuit of Power: Technology, Armed Force, and Society Since A.D. 1000.* Chicago: The University of Chicago Press.

52. Michalopoulos, S. (2012). "The Origins of Ethnolinguistic Diversity." *American Economic Review*, 102, 1508-1539.

53. Michalopoulos, S., and E. Papaioannou (2013). "Pre-Colonial Ethnic iInstitutions and Contemporary African Development." *Econometrica* 81, 113–152.

54. Michalopoulos, S., and E. Papaioannou (2014). "National Institutions and Subnational Development in Africa." *Quarterly Journal of Economics*, 129, 151–213.

55. Miller, A. A. (1953). *Climatology.* London: Methuen and Co., Ltd.

56. Mokyr, J. (2006). "Mobility, Creativity, and Technological Development: David Hume, Immanuel Kant and the Economic Development of Europe." In G. Abel (Ed.) *Kolloquiumsband of the XX. Deutschen Kongresses für Philosophie*, 1131-1161.

57. Mokyr, J. (2007). "The Market for Ideas and the Origins of Economic Growth in Eighteenth Century Europe." *Tijdschrift voor Sociale en Economische Geschiedenis*, 4 (1), 3-38.

58. Nissen, H. J. and P. Heine (2009). *From Mesopotamia to Iraq: A Concise History*. Chicago: The University of Chicago Press.

59. Nunn, N., and N. Qian (2011). "The Potato's Contribution to Population and Urbanization: Evidence from a Historical Experiment." *Quarterly Journal of Economics*, 126, 593–650.

60. Nunn, N. and D. Puga (2012). "Ruggedness: The Blessing of Bad Geography in Africa." *Review of Economics and Statistics*, 94, 20-36.

61. Nüssli, C. (2010). "Euratlas Georeferenced Vector Data Description." *Mimeo*. http://www.euratlas.com/about.html

62. Pounds, N. J. G. (1972). *Political Geography*, 2nd ed. New York: McGraw-Hill Inc.

63. Potrafke, N. and R. Vaubel (2014). "The Interjurisdictional Migration of European Authors of Liberty, 1660-1961: A Quantitative Analysis." *Mimeo*.

64. Rappaport, J., and J.D. Sachs (2003). "The United States as a Coastal Nation." *Journal of Economic Growth*, 8, 5-46.

65. Sokoloff, K. L. and S. L. Engerman (2000). "History lessons: Institutions, Factors Endowments, and Paths of Development in the New World." *Journal of Economic Perspectives*, 14, 217-232.

66. Taagepera, R. (1978). "Size and Duration of Empires: Systematics of Size." *Social Science Research*, 7, 108-127.

67. The Economist (2014). "Transdniestria—Beyond the Tiver." The Economist April 26th, 2014.

68. Ticchi, D., and A. Vindigni (2008). "War and Endogenous Democracy." *IZA Discussion Paper*, No. 3397.

69. Tilly, C. (1992). *Coercion, Capital and European States, A.D. 990-1992*. Malden: Blackwell.

70. Turchin, P. (2009). "A Theory of Formation of Large Empires." *Journal of Global History*, 4, 191-217.

71. Umesao, T. (2003). *An Ecological View of History: Japanese Civilization in the World Context.* Melbourne, Australia: Trans Pacific Press.

72. Wittfogel, K. A. (1957). *Oriental Despotism: A Comparative Study of Total Power.* New Haven: Yale University Press.

Appendix

A The model

A.1 The dynamics of Y_t

Let $\overline{A}_{i,t+1}$ be the mean of $A_{j,t+1}$ across all $N-1$ countries, excluding country i, i.e.,

$$\overline{A}_{i,t+1} = \frac{\sum_{j \neq i} A_{j,t+1}}{N-1}, \tag{21}$$

Using (21), we can write (4) forwarded one period as

$$R_{i,t+1} = (1-\lambda)\overline{R} + \frac{1}{s}\left(\frac{A_{i,t+1}}{A_{i,t+1} + (N-1)\overline{A}_{i,t+1}}\right)\lambda\overline{R}\,. \tag{22}$$

Differentiating (22) with respect to $A_{i,t+1}$, holding constant $\overline{A}_{i,t+1}$, gives

$$\frac{\partial R_{i,t+1}}{\partial A_{i,t+1}} = \frac{1}{s}\left(\frac{[N-1]\,\overline{A}_{i,t+1}}{\left[A_{i,t+1} + (N-1)\overline{A}_{i,t+1}\right]^2}\right)\lambda\overline{R}. \tag{23}$$

Imposing $A_{i,t+1} = \overline{A}_{i,t+1}$ and $s = 1/N$, on (22) and (23) gives

$$R_{i,t+1} = \overline{R}, \tag{24}$$

and

$$\frac{\partial R_{i,t+1}}{\partial A_{i,t+1}} = \frac{1}{s}\left(\frac{N-1}{N^2}\right)\left(\frac{1}{A_{i,t+1}}\right)\lambda\overline{R} = (1-s)\left(\frac{1}{A_{i,t+1}}\right)\lambda\overline{R}. \tag{25}$$

where we note (from $s = 1/N$) that $(1/s)\,(N-1)\,/N^2 = 1-s$. Now (24) and (25) give us an expression for the elasticity of resources with respect to technology in a symmetric equilibrium

$$\frac{\partial R_{i,t+1}}{\partial A_{i,t+1}}\frac{A_{i,t+1}}{R_{i,t+1}} = (1-s)\,\lambda. \tag{26}$$

Using (26), the first-order condition in (7) can be written

$$(1-\beta)\,[C_{i,t}]^{-1} = \beta F(s,\lambda)\,[A_{i,t+1}]^{-1}\,, \tag{27}$$

where $F(s,\lambda)$ is defined in (8). Using (5) and (27), and dropping all i subscripts, we can solve for A_{t+1} as

$$A_{t+1} = \left[\frac{\beta F(s,\lambda)}{1-\beta+\beta F(s,\lambda)}\right]\tau s Y_t. \tag{28}$$

Forwarding (3), using (28), and suppressing the i subscripts, we get a dynamic equation for output per location in each country:

$$Y_{t+1} = (1-\gamma s)\,A_{t+1}^{\alpha}\overline{R}^{1-\alpha} = s^{\alpha}\,(1-\gamma s)\left[\frac{\beta F(s,\lambda)}{1-\beta+\beta F(s,\lambda)}\right]^{\alpha}\overline{R}^{1-\alpha}\,(\tau Y_t)^{\alpha}\,. \tag{29}$$

A.1.1 Proof of Result 1

First rewrite (9) as

$$Y^* = [s^\alpha (1-\gamma s)]^{\frac{1}{1-\alpha}} \left[\frac{\tau\beta F(s,\lambda)}{1-\beta+\beta F(s,\lambda)}\right]^{\frac{\alpha}{1-\alpha}} \overline{R} \equiv \widehat{Y}(s). \tag{30}$$

The task is to find conditions under which $\widehat{Y}(\frac{1}{2}) \gtreqless \widehat{Y}(1)$. Using (30), and the expression for $F(s,\lambda)$ in (8), gives

$$\begin{aligned} \frac{\widehat{Y}(\frac{1}{2})}{\widehat{Y}(1)} &= \left\{\left[\frac{\left(\frac{1}{2}\right)^{1+\alpha}(2-\gamma)}{1-\gamma}\right]\left[\left(\frac{1-\beta+\beta F(1,\lambda)}{\beta F(1,\lambda)}\right)\left(\frac{\beta F(1/2,\lambda)}{1-\beta+\beta F(1/2,\lambda)}\right)\right]^\alpha\right\}^{\frac{1}{1-\alpha}} \\ &= \left\{\left[\left(\tfrac{1}{2}\right)^{1+\alpha}\left(\frac{2-\gamma}{1-\gamma}\right)\right]^{\frac{1}{\alpha}}\left(\frac{1-\beta+\beta\alpha}{\beta\alpha}\right)\left(\frac{\beta[\alpha+(1-\alpha)\lambda/2]}{1-\beta+\beta[\alpha+(1-\alpha)\lambda/2]}\right)\right\}^{\frac{\alpha}{1-\alpha}} \\ &= \left\{Q\left(\frac{1-\beta+\beta\alpha}{\beta\alpha}\right)H(\lambda)\right\}^{\frac{\alpha}{1-\alpha}}, \end{aligned} \tag{31}$$

where we have used the definition of Q in (10), and where we let

$$H(\lambda) = \frac{\beta\left[\alpha+(1-\alpha)\lambda/2\right]}{1-\beta+\beta\left[\alpha+(1-\alpha)\lambda/2\right]}. \tag{32}$$

Note that $H'(\lambda) > 0$, and recall that $\lambda \in [0,1]$. Then some algebra demonstrates that

$$\begin{aligned} H(0) &= \frac{\alpha\beta}{1-\beta+\beta\alpha}, \\ H(1) &= \frac{\beta\left(\frac{1+\alpha}{2}\right)}{1-\beta+\beta\left(\frac{1+\alpha}{2}\right)} = \frac{\beta(1+\alpha)}{2-\beta+\alpha\beta}. \end{aligned} \tag{33}$$

We can now show the following:

If $Q \geq 1$, then $\widehat{Y}(\frac{1}{2})/\widehat{Y}(1) \geq 1$ always holds. To see this, use (31) and (33) to note that it holds even when $\lambda = 0$, and since $\widehat{Y}(\frac{1}{2})/\widehat{Y}(1)$ is increasing in λ, this proves part (a) of the result.

If $Q \leq \left(\frac{\alpha}{1+\alpha}\right)\left(\frac{2-\beta+\alpha\beta}{1-\beta+\alpha\beta}\right)$, then $\widehat{Y}(\frac{1}{2})/\widehat{Y}(1) \leq 1$ always holds. To see this, use (33), and some algebra, to note that the given condition on Q is equivalent to $Q\left(\frac{1-\beta+\beta\alpha}{\alpha\beta}\right)H(1) \leq 1$, which implies that $\widehat{Y}(\frac{1}{2})/\widehat{Y}(1) \leq 1$ holds even when $\lambda = 1$; then recall again that $\widehat{Y}(\frac{1}{2})/\widehat{Y}(1)$ is increasing in λ, so it must hold also for $\lambda < 1$. This proves part (b) of the result.

Finally, if $Q \in \left(\left(\frac{\alpha}{1+\alpha}\right)\left(\frac{2-\beta+\alpha\beta}{1-\beta+\alpha\beta}\right), 1\right)$, then $\widehat{Y}(\frac{1}{2})/\widehat{Y}(1) \geq 1$ holds if, and only if,

$$Q\left(\frac{1-\beta+\beta\alpha}{\alpha\beta}\right)H(\lambda) \geq 1. \tag{34}$$

Using (32), some algebra shows that this inequality can be written as in (11). This proves part (c) of the result.

B Data

B.1 Border variables

Euratlas data The border data were purchased from Euratlas (www.euratlas.com), © 2010 Christos Nüssli. For each turn of the century from 1 to 2000 CE, the Euratlas data contain shapefiles for different political formations in Europe and its surroundings. The ones used in our benchmark analysis refer to *sovereign* states, defined by Euratlas as states with an authority, ruling over a territory and a population, and where "this authority is sovereign, i.e. not subject to any other power or state" (Nüssli 2010).

The shapefiles span from -19.25 to 51.5 degrees longitude, and from 19.25 to 60.25 degrees latitude. One degree is about 111 kilometers at the equator; the exact length depends on where on earth it is measured.

The Holy Roman Empire In some regressions we treat the Holy Roman Empire (HRE) as unified. The polygons of the HRE are also from Euratlas. To create the HRE adjusted border dummies, we overwrite the polygons of the sovereign states using the HRE polygons, and then follow the same procedures as in the baseline setting, i.e., we let a border cell be a cell with two or more states. This is done for each of the four years when the HRE existed according to the Euratlas data (i.e., for 1500, 1600, 1700, and 1800), applying the `Update` tool in ArcGIS. We then create the HRE adjusted border frequency index by averaging from 1500 to 2000 for each cell.

B.2 Geography variables

Data on elevation (topography) are from the National Geophysical Data Centre (NGDC) (www.ngdc.noaa.gov/mgg/topo/globe.html) at the National Oceanic and Atmospheric Administration (NOAA). These data refer to land areas only and are provided at the 30-arc-second (about 0.0083 degrees) level. For each cell (of size 0.5×0.5 degrees), there are thus at most 3600 elevation points, depending on how much of it is covered by land (see below). From these elevation data we calculate ruggedness as the standard deviation in elevation points in each cell.

The mountain dummies take the value one if the average elevation in the cell exceed 1000 or 2000 meters, respectively, and zero otherwise.

Using the 1:10 million scale land polygons, available from Natural Earth (NE) (www.naturalearthdata.com), we define sea cells as those cells for which either the NGDC elevation data, or the NE's land data, are completely missing. Cells that are not sea cells are land cells (i.e., they do not miss NGDC or NE data). Coastline density is then defined as the length of the coastline from NE intersecting the cell, divided by the cell's land area, also from NE.

Distance to the coast is calculated from the centroid of each cell to the nearest point on any coastline using the `Near` tool in ArcGIS.

We also construct river density from NE (see above). First, we use a shapefile called Rivers and Lake Centerlines at the scale 1:10 million, to map both rivers and lake centerlines. We then use a different shapefile called Lakes and Reservoirs, also at the scale 1:10 million, to remove the lake centerlines. We are left with river lines, and can calculate the length of these in each cell. River density is the length of the river line intersecting the cell, divided by the size of the land area of the cell. Land area is also from NE.

The source for the agricultural suitability data is the Global Agro-Ecological Zones (GAEZ) website (www.fao.org/nr/gaez), sponsored by the Food and Agriculture Organization of the United Nations (FAO). These are provided at the 5-arc-minute (about 0.0833 degrees) level, implying a maximum of 36 observations per cell. This suitability index measures agricultural output of some given crop and level of input, relative to the output level for the same crop and the same level of inputs, but under perfect environmental conditions, and based on the climatic conditions 1961-90.

The variables we construct refer to the average of wheat, barley, oats and rye, under intermediate inputs. For these crops (although not for all other crops), GAEZ also distinguishes between two types of water supply, rain and irrigation. This allows us to construct two suitability variables, one for each of the two water-supply categories. We normalize these two variables to run from 0 to 1 (the original scale runs from 0 to 10,000).

Rainfall data are also from GAEZ. We use total annual millimeter rainfall (precipitation), averaged over the period 1961-90, and compute millimeters per day by dividing by 365.

B.3 Modern outcomes

Night lights data are from National Oceanic and Atmospheric Administration (NOAA), at https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html. These are cleaned by NOAA to remove observations affected by northern lights, forest fires, clouds, and more, in order to measure human-made lights only. We remove gas flares, following the guidelines constructed by Lowe (2014). The resulting data are reported on a scale from 0 to 63 at the pixel level. We then average across pixels in each cell, and over the years 1992-2013.

Modern population data is from the Gridded Population of the World, v4, at SEDAC at NASA, linked to here: http://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-count-adjusted-to-2015-unwpp-country-totals. To get population density we divide the cell's total population (averaged over the period 2000-2015) by the land area of the cell, calculated from the same data source as rivers and coastlines (Natural Earth).

Table 1: Summary statistics.

	Mean	Standard deviation	Min	Max	Median	Number of cells	Correlation with border frequency
Border dummy 1500	0.18	0.38	0	1	0	5202	0.72
Border dummy 1600	0.15	0.35	0	1	0	5202	0.72
Border dummy 1700	0.15	0.35	0	1	0	5202	0.81
Border dummy 1800	0.10	0.30	0	1	0	5202	0.72
Border dummy 1900	0.09	0.28	0	1	0	5202	0.60
Border dummy 2000	0.16	0.36	0	1	0	5202	0.60
Border frequency 1500-2000	0.14	0.24	0	1	0	5202	1.00
Border change 1500-2000	0.11	0.18	0	1	0	5202	0.64
Mountain >2000m	0.02	0.14	0	1	0	5202	0.08
Mountain >1000m	0.12	0.32	0	1	0	5202	0.09
Log ruggedness	4.12	1.28	0	7.07	4.01	5202	0.15
River density	0.01	0.01	0	0.60	0	5202	0.11
Ag. suit. rainfed	0.47	0.27	0	1	0.49	5202	0.11
Ag. suit. irrig.	0.35	0.29	0	1	0.31	5202	−0.04
Rainfall	1.70	0.84	0	6.37	1.66	5202	0.25
Log dist. to coast	0.13	0.13	0	0.69	0.09	5202	−0.04
Coastline density	0.06	0.30	0	9.90	0	5202	−0.07
Log land area	7.31	0.99	−1.41	7.97	7.60	5202	0.10
Log night lights	1.51	0.94	0	4.11	1.47	5202	0.17
Log population density	3.76	1.48	0	9.10	3.86	5201	0.19

Notes: Summary statistics for some of the more important variables across 5202 cells. Border frequency 1500-2000 is the fraction of the years in which a cell had a border 1500-2000.

Table 2: Geography and border frequency.

	Dependent variable: Border frequency 1500-2000								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mountain >2000m	0.136***		0.097**	0.095**	0.106**	0.091*	0.103**	0.118***	0.122***
	(0.049)		(0.047)	(0.047)	(0.047)	(0.046)	(0.047)	(0.042)	(0.042)
Mountain >1000m		0.052**							
		(0.023)							
Log ruggedness			0.022***	0.024***	0.027***	0.023***	0.026***	0.011**	0.009*
			(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
River density				2.023***	1.836***	2.141***	1.989***	1.866***	1.884***
				(0.723)	(0.669)	(0.775)	(0.727)	(0.682)	(0.682)
Ag. suit. rainfed					0.087***		0.152***	0.079***	0.090***
					(0.019)		(0.023)	(0.027)	(0.028)
Ag. suit. irrig.						−0.047**	−0.116***	−0.107***	−0.108***
						(0.020)	(0.023)	(0.022)	(0.022)
Rainfall								0.069***	0.068***
								(0.011)	(0.011)
Log dist. to coast									−0.069
									(0.047)
Coastline density									0.023*
									(0.014)
Log land area	0.024***	0.022***	0.018***	0.015***	0.007**	0.018***	0.009**	0.027***	0.035***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)
R^2	0.02	0.02	0.03	0.04	0.05	0.05	0.07	0.12	0.12
Number of obs.	5202	5202	5202	5202	5202	5202	5202	5202	5202

Notes: Ordinary least squares regressions with Conley standard errors in parentheses assuming spatial autocorrelation among observations within 1.45 degrees of each other. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 3: Geography and border frequency: robustness.

	Dependent variable: Border frequency 1500-2000					
	(1)	(2)	(3)	(4)	(5)	(6)
Mountain >2000m	0.127***		0.131***	0.108***	0.153***	
	(0.040)		(0.037)	(0.039)	(0.033)	
Mountain >1000m		0.052***				0.074***
		(0.019)				(0.018)
Log ruggedness	0.027***	0.025***	0.017***	0.029***	0.019***	0.015***
	(0.007)	(0.007)	(0.006)	(0.009)	(0.006)	(0.006)
River density	1.806***	1.794***	1.146**	2.579***	1.382**	1.365**
	(0.644)	(0.645)	(0.453)	(0.380)	(0.658)	(0.660)
Ag. suit. rainfed	0.004	−0.000	0.040	−0.067*	0.084***	0.081***
	(0.030)	(0.030)	(0.025)	(0.036)	(0.029)	(0.029)
Ag. suit. irrig.	−0.104***	−0.098***	−0.041**	−0.124***	−0.101***	−0.091***
	(0.022)	(0.023)	(0.017)	(0.024)	(0.019)	(0.019)
Rainfall	0.040***	0.040***	0.030***	0.095***	0.039***	0.040***
	(0.013)	(0.013)	(0.010)	(0.020)	(0.013)	(0.013)
Log dist. to coast	−0.116**	−0.117**	−0.066	−0.147**	−0.026	−0.032
	(0.052)	(0.052)	(0.042)	(0.060)	(0.053)	(0.053)
Coastline density	0.023**	0.020*	0.018**		0.026**	0.021**
	(0.010)	(0.010)	(0.009)		(0.010)	(0.010)
Log land area	0.028***	0.026***	0.018***		0.028***	0.026***
	(0.006)	(0.006)	(0.005)		(0.006)	(0.005)
R^2	0.19	0.19	0.14	0.26	0.33	0.33
Number of obs.	5202	5202	5202	3869	5202	5202
Fixed effects	Latitude	Latitude	Latitude	Latitude	Lat./Long.	Lat./Long.
HRE adjustment	No	No	Yes	No	No	No
Drop coastal cells	No	No	No	Yes	No	No

Notes: Ordinary least squares regressions with Conley standard errors in parentheses assuming spatial autocorrelation among observations within 1.45 degrees of each other. Latitude fixed effects contain of a full set of dummies for each half-degree latitude, i.e., one dummy for each row in the grid. Similarly, longitude fixed effects contain dummies for each half-degree longitude, i.e., one dummy for each column in the grid. HRE adjustment indicates that the dependent variable is redefined treating the Holy Roman Empire as unified (see text). * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 4: Geography and border frequency: local deviations.

	Dependent variable: Δ Border frequency 1500-2000					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Mountain >2000m	0.050					
	(0.037)					
Δ Mountain >1000m		0.069***	0.070***	0.073***	0.076***	0.071***
		(0.018)	(0.018)	(0.018)	(0.019)	(0.019)
Δ Log ruggedness	0.021***	0.020***	0.019***	0.016***	0.021***	0.018***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.005)
Δ River density	0.716***	0.731***	0.717***	0.726***	1.069***	0.678***
	(0.233)	(0.238)	(0.240)	(0.249)	(0.272)	(0.250)
Δ Ag. suit. rainfed	−0.104***	−0.099***	−0.109***	−0.101***	−0.133***	−0.099***
	(0.039)	(0.037)	(0.036)	(0.036)	(0.046)	(0.035)
Δ Ag. suit. irrig.	0.013	0.020	0.022	0.028	0.005	0.024
	(0.019)	(0.019)	(0.018)	(0.019)	(0.021)	(0.018)
Δ Rainfall	0.064**	0.064**	0.065**	0.061**	0.085*	0.073**
	(0.028)	(0.028)	(0.028)	(0.028)	(0.044)	(0.029)
Δ Log dist. to coast	−0.168	−0.300	−0.352	−0.273	−0.099	−0.327
	(0.281)	(0.274)	(0.275)	(0.276)	(0.333)	(0.278)
Δ Coastline density	0.018**	0.016**	0.015*	0.014*		0.013
	(0.008)	(0.008)	(0.008)	(0.007)		(0.009)
Δ Log land area	0.016***	0.015***	0.015***	0.014***		0.014***
	(0.005)	(0.005)	(0.005)	(0.004)		(0.005)
R^2	0.02	0.03	0.04	0.04	0.05	0.06
Number of obs.	5202	5202	5202	5202	3869	5202
Fixed effects	None	None	Latitude	Latitude	Latitude	Lat./Long.
HRE adjustment	No	No	No	Yes	No	No
Drop coastal cells	No	No	No	No	Yes	No

Notes: Ordinary least squares regressions with Conley standard errors in parentheses assuming spatial autocorrelation among observations within 1.45 degrees of each other. The dependent variable is the local deviation in border frequency between each cell and its eight closest neighboring cells. Column (4) uses the local deviation in HRE adjusted border frequency; see notes to Table 3. The independent variables of interest are the corresponding local deviations in each of the geography variables, as indicated. Latitude and longitude fixed effects mean the same as in Table 3 (not measured as local deviations). * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 5: Borders and state-size gaps: panel regressions.

	Dependent variable: Border dummy$_t$				
	(1)	(2)	(3)	(4)	(5)
Border dummy$_{t-1}$	0.449***	-0.063^{***}	-0.031^{*}	-0.015	-0.014
	(0.010)	(0.013)	(0.018)	(0.018)	(0.018)
Log state size gap$_{t-1}$			-0.016^{***}	-0.089^{***}	-0.062^{***}
			(0.006)	(0.019)	(0.015)
$\times$ Log ruggedness				0.015***	
				(0.004)	
$\times$ Rainfall					0.019***
					(0.005)
R^2	0.31	0.58	0.58	0.58	0.58
Number of obs.	20808	20808	20808	20808	20808
Geography controls	Yes	No	No	No	No
Fixed effects	Latitude/century	Century/cell	Century/cell	Century/cell	Century/cell

Notes: Ordinary least squares regressions with robust standard errors in parentheses, based on a panel with five centuries (1500-1900) and 5202 cells. All specifications include both century (year) and cell fixed effects, except column (1) which includes no cell fixed effects, but century and latitude fixed effects, and the benchmark set of geography controls used in column (9) of Table 2. The log state size gap is the logarithm of the ratio of the size of largest state intersecting the cell over the size of the smallest state. The log state size gap thus equals zero for non-border cells, since they have only one state. * indicates $p<0.10$, ** $p<0.05$, and *** $p<0.01$.

Table 6: Borders and modern outcomes.

Panel A	Dependent variable: Log night lights					
	(1)	(2)	(3)	(4)	(5)	(6)
Border frequency 1500-2000	0.680***	0.563***	0.371***	−0.003	0.324***	0.048
	(0.097)	(0.096)	(0.084)	(0.080)	(0.089)	(0.071)
R^2	0.03	0.28	0.35	0.34	0.39	0.46
Number of obs.	5202	5202	5202	5202	3869	5202
Panel B	**Dependent variable: Log population density**					
Border frequency 1500-2000	1.157***	0.856***	0.540***	0.103	0.510***	0.184*
	(0.134)	(0.139)	(0.120)	(0.115)	(0.129)	(0.112)
R^2	0.03	0.28	0.39	0.38	0.45	0.44
Number of obs.	5201	5201	5201	5201	3869	5201
Geography controls	No	Yes	Yes	Yes	Yes	Yes
Fixed effects	None	None	Latitude	Latitude	Latitude	Lat./Long.
HRE adjustment	No	No	No	Yes	No	No
Drop coastal cells	No	No	No	No	Yes	No

Notes: Ordinary least squares regressions with Conley standard errors in parentheses assuming spatial autocorrelation among observations within 1.45 degrees of each other. The dependent variable is log night lights in Panel A, and log population density in Panel B. Geography controls are those used in column (9) of Table 2. Latitude and longitude fixed effects, and HRE adjustment, mean the same as in Table 3. * indicates p <0.10, ** p <0.05, and *** p <0.01.

Table 7: Borders and modern outcomes: local deviations.

Panel A	Dependent variable: Δ Log night lights					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Border frequency 1500-2000	−0.165***	−0.183***	−0.179***	−0.169***	−0.283***	−0.190***
	(0.054)	(0.053)	(0.053)	(0.052)	(0.049)	(0.052)
R^2	0.00	0.11	0.13	0.13	0.16	0.15
Number of obs.	5202	5202	5202	5202	3869	5202
Panel B	**Dependent variable: Δ Log population density**					
Δ Border frequency 1500-2000	−0.286***	−0.291***	−0.280***	−0.267***	−0.437***	−0.294***
	(0.078)	(0.076)	(0.076)	(0.073)	(0.072)	(0.075)
R^2	0.00	0.06	0.08	0.08	0.10	0.10
Number of obs.	5201	5201	5201	5201	3869	5201
Geography controls	No	Yes	Yes	Yes	Yes	Yes
Fixed effects	None	None	Latitude	Latitude	Latitude	Lat./Long.
HRE adjustment	No	No	No	Yes	No	No
Drop coastal cells	No	No	No	No	Yes	No

Notes: Ordinary least squares regressions with Conley standard errors in parentheses assuming spatial autocorrelation among observations within 1.45 degrees of each other. The dependent variables in Panels A and B are the local deviations in log night lights and log population density, respectively, between each cell and its eight closest neighboring cells. The independent variable of interest is the corresponding local deviation in border frequency, HRE adjusted in column (4). Geography controls are local deviations in the independent variables used in column (9) of Table 2. Latitude and longitude fixed effects, and HRE adjustment, mean the same as in Table 3. * indicates p <0.10, ** p <0.05, and *** p <0.01.

Table 8: Borders and modern outcomes: the role of border stability.

	Dependent variable:							
	Log night lights				Log population density			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Border frequency 1500-2000	0.563***		0.866***	0.659***	0.856***		0.974***	0.690***
	(0.096)		(0.114)	(0.096)	(0.139)		(0.176)	(0.145)
Border change 1500-2000		0.078	−0.612***	−0.582***		0.536***	−0.240	−0.303**
		(0.095)	(0.115)	(0.100)		(0.136)	(0.171)	(0.145)
R^2	0.28	0.26	0.29	0.35	0.28	0.26	0.28	0.39
Number of obs.	5202	5202	5202	5202	5201	5201	5201	5201
Geography controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	None	None	None	Latitude	None	None	None	Latitude

Notes: Ordinary least squares regressions with Conley standard errors in parentheses assuming spatial autocorrelation among observations within 1.45 degrees of each other. Geography controls are the variables used in column (9) in Table 2. The latitude fixed effects contain dummies for each half-degree latitude, as in Table 3. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 9: Borders and urbanization: panel regressions.

	Dependent variable: Fraction $urban_t$					
	(1)	(2)	(3)	(4)	(5)	(6)
Fraction $urban_{t-1}$	1.024^{***}	1.024^{***}	1.024^{***}	0.802^{***}	0.804^{***}	0.802^{***}
	(0.007)	(0.007)	(0.007)	(0.012)	(0.012)	(0.012)
Border $dummy_t$	-0.005^{**}		-0.007^{***}	-0.011^{***}		-0.011^{***}
	(0.002)		(0.003)	(0.003)		(0.003)
Border $dummy_{t-1}$		0.002	0.005^{**}		0.002	0.002
		(0.002)	(0.002)		(0.003)	(0.003)
R^2	0.81	0.81	0.81	0.85	0.85	0.85
Number of obs.	25178	25178	25178	25178	25178	25178
Geography controls	Yes	Yes	Yes	No	No	No
Fixed effects	Century/ latitude	Century/ latitude	Century/ latitude	Century/ cell	Century/ cell	Century/ cell

Notes: Ordinary least squares regressions with robust standard errors in parentheses, based on a panel with six centuries (1500-2000) and 5202 cells. Columns (4)-(6) include both century (year) and cell fixed effects, while columns (1)-(3) include no cell fixed effects, but century and latitude fixed effects, and the geography controls used in column (9) of Table 2. * indicates $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Figure 1: Illustration of Result 1. The figure is drawn for $Q \geq 1/2$, and under the assumption that $\left(\frac{\alpha}{1+\alpha}\right)\left(\frac{2-\beta+\alpha\beta}{1-\beta+\alpha\beta}\right) > \frac{1}{2}$, which holds for large enough β.

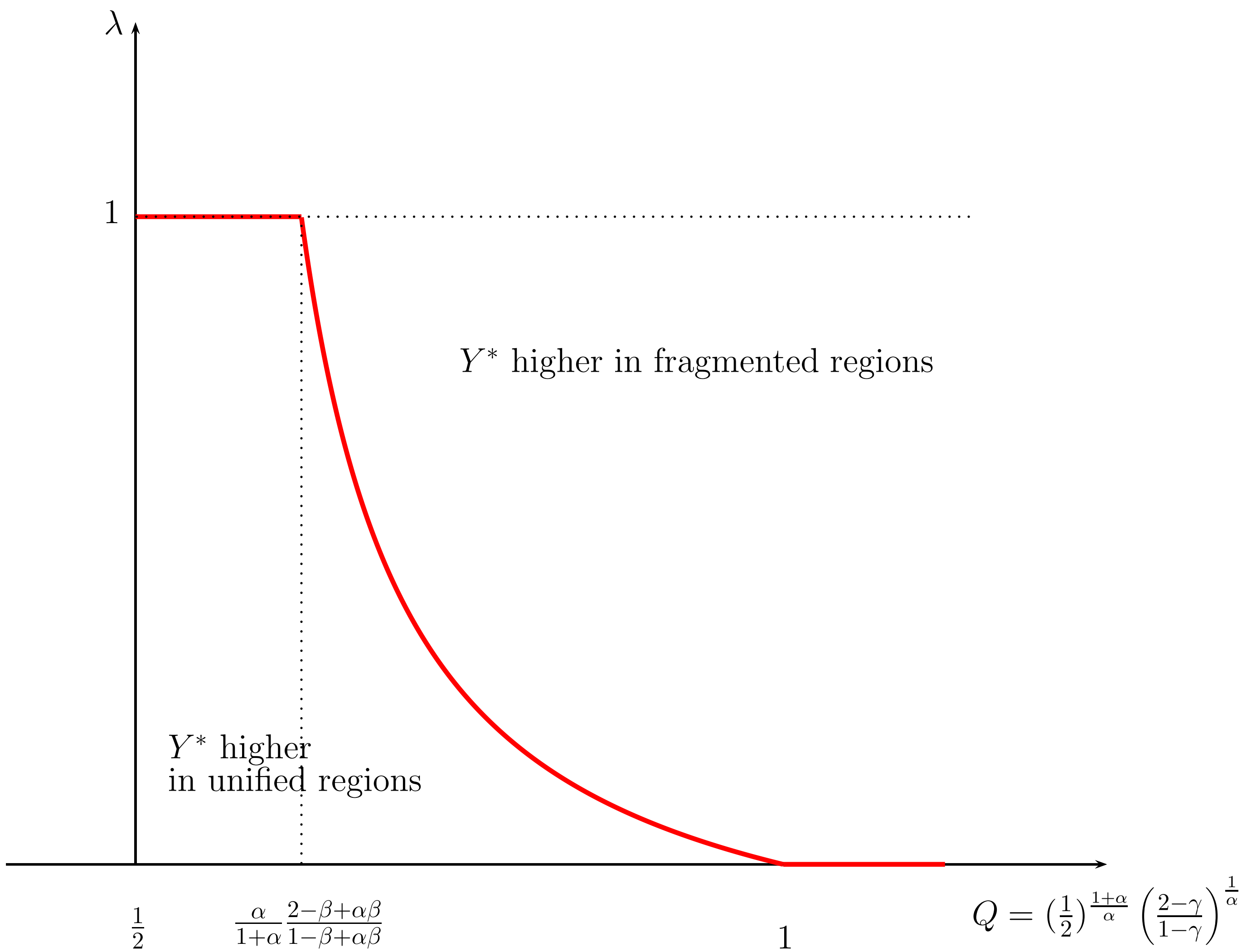


Figure 2: How location specific steady-state output, $\widetilde{Y}^*(d)$, varies across locations in two regions with different numbers of countries, $N=2$ and $N=8$, for a numerical example ($\tau=1$, $\alpha=0.1$, $\beta=\gamma=\lambda=0.5$, and $\overline{R}=100$).

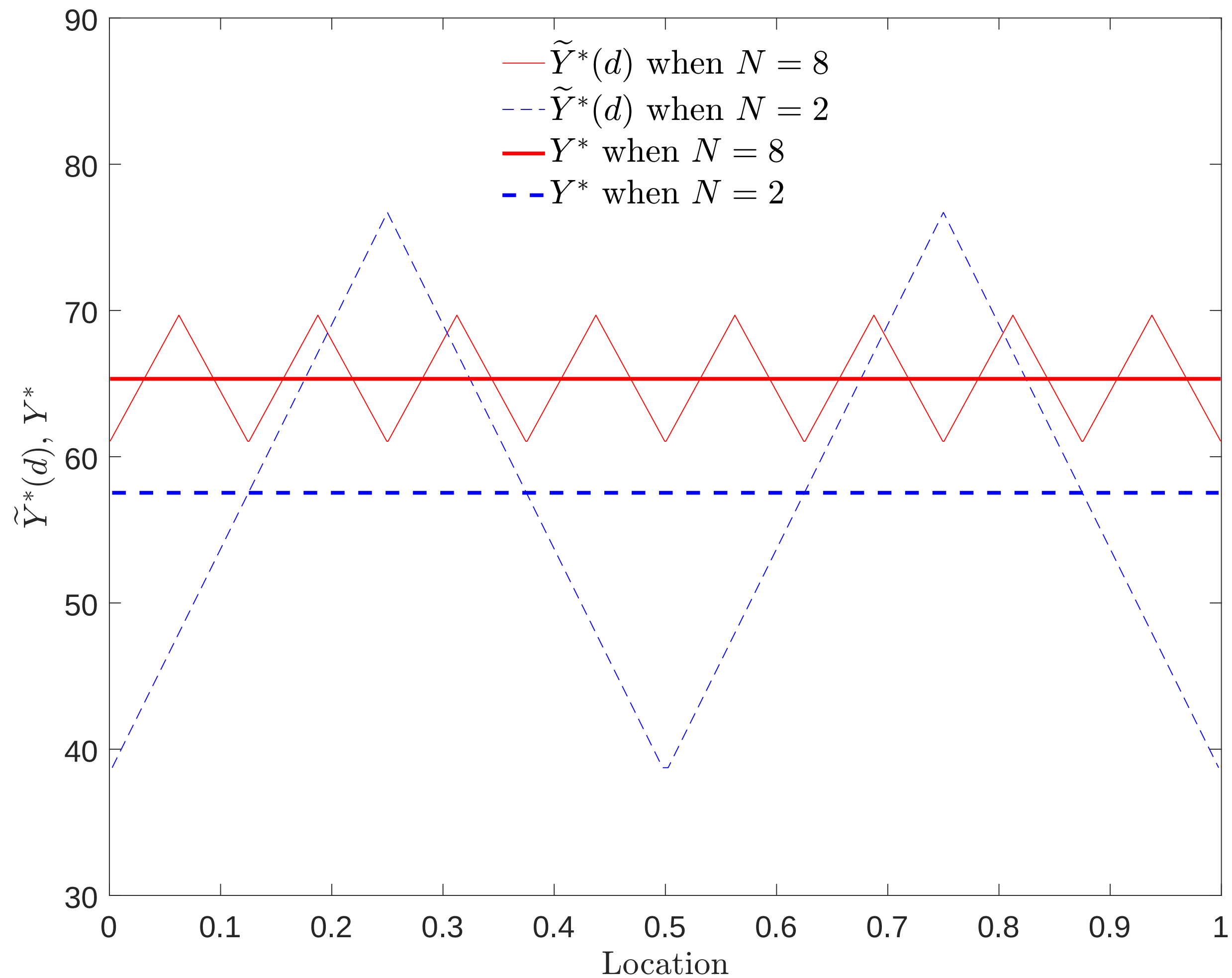


Figure 3: Map showing coastal cells, cells with state presence 1500-2000, and border cells in 1500 and 1900, respectively.

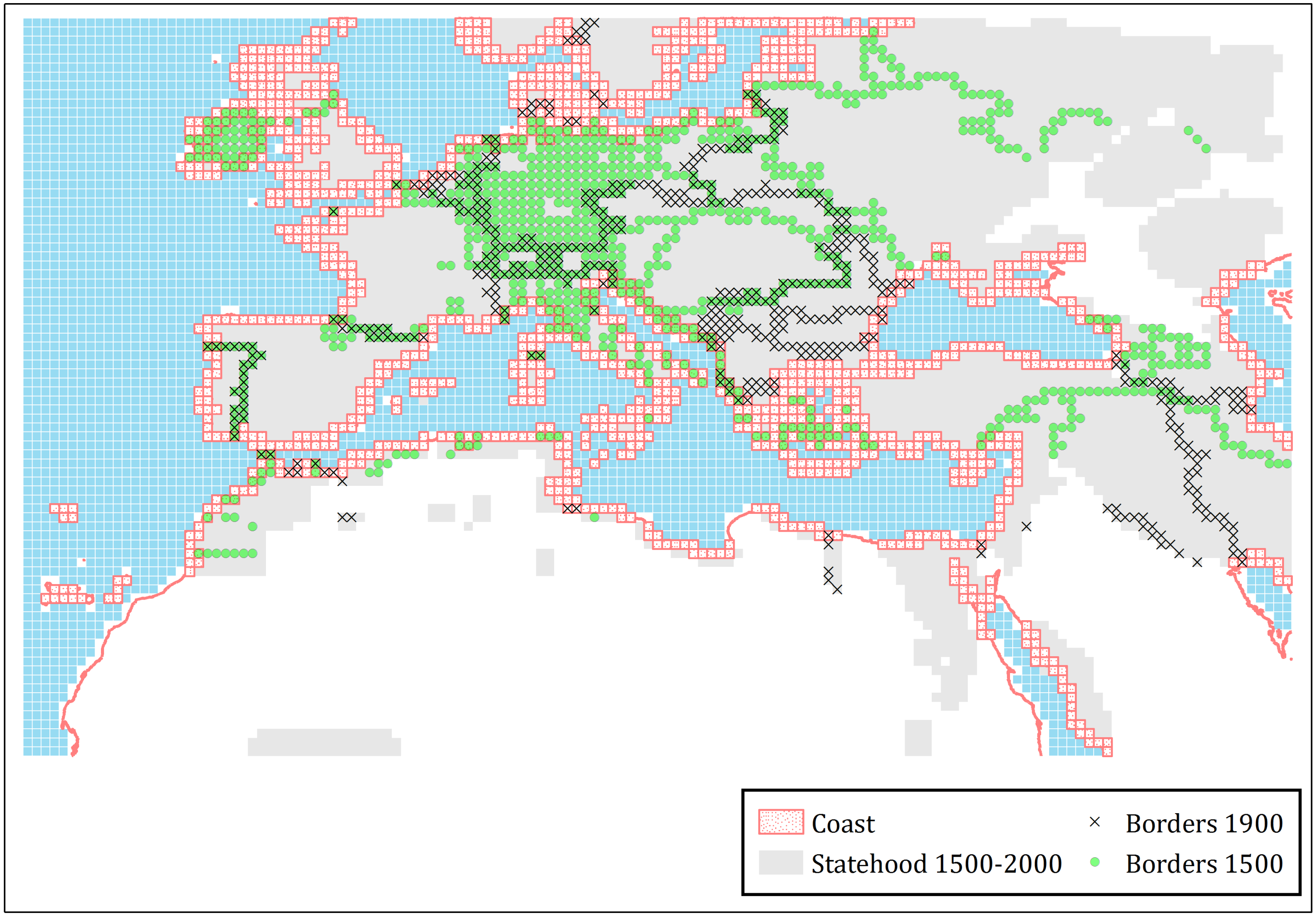


Figure 4: Map showing cells with borders at least five of the six years 1500-2000 ($B_i \geq 5/6$), mountains over 1000 and 2000 meters, and rivers.

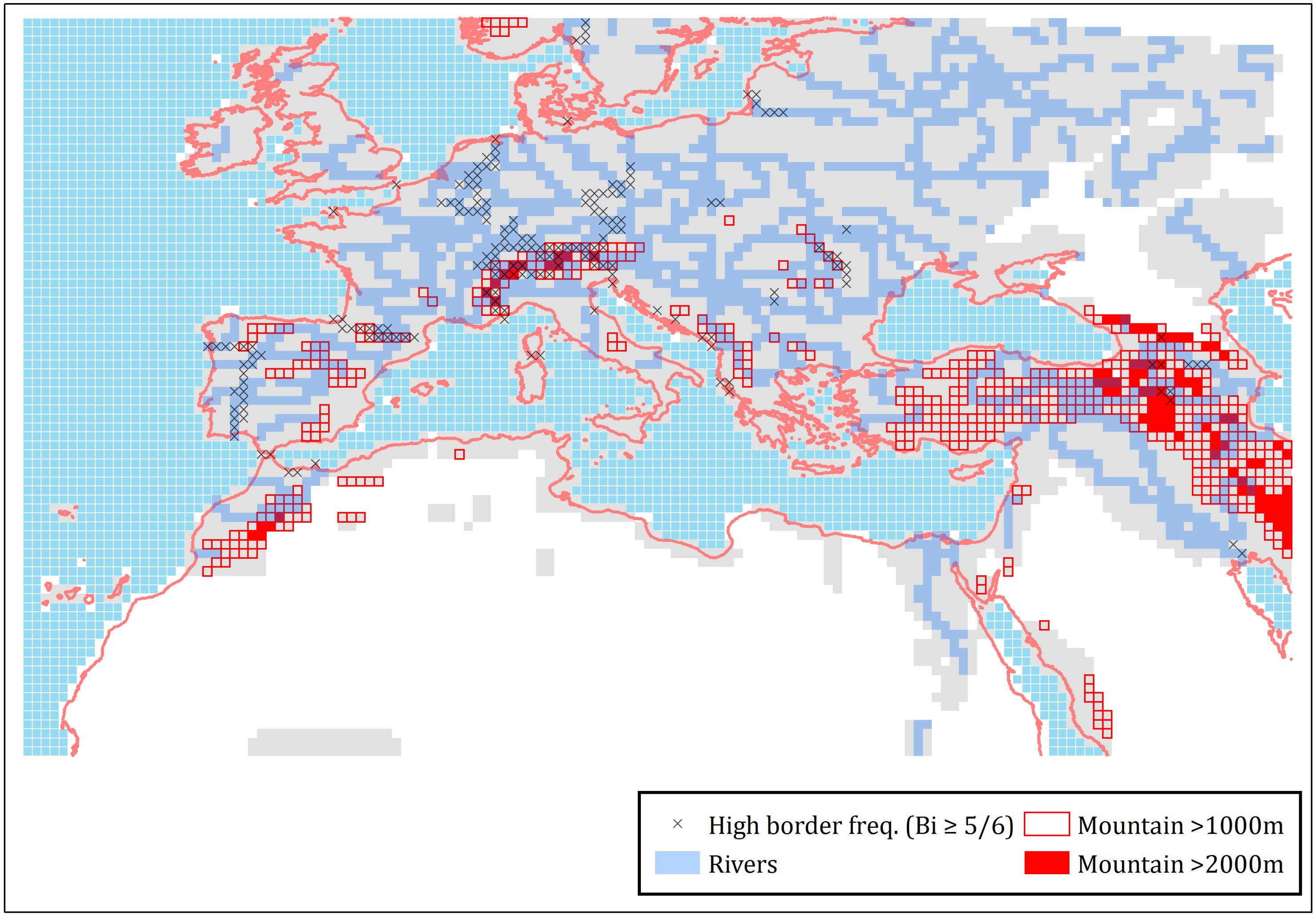


Figure 5: Border frequency and different measures of geography.

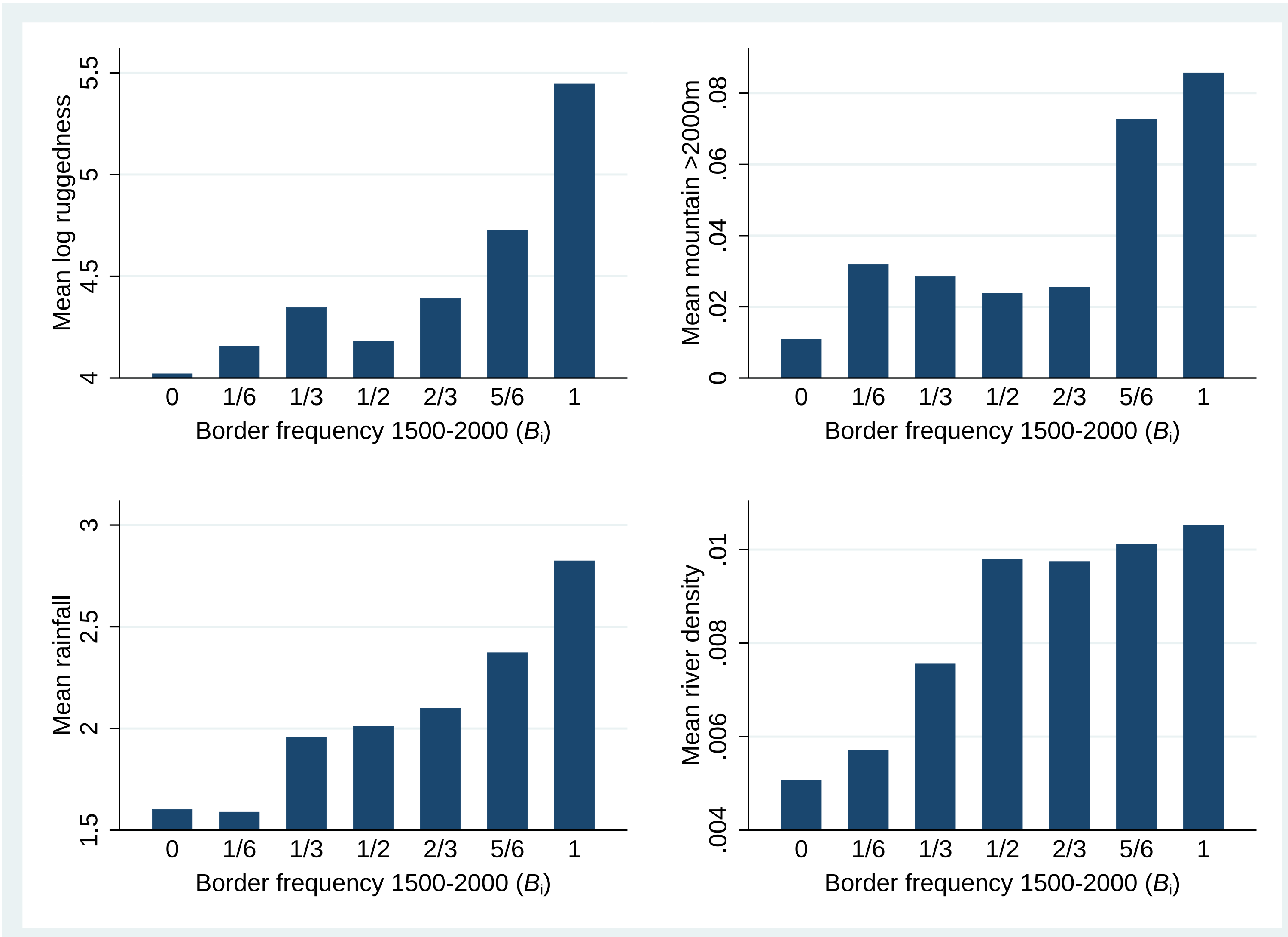


Figure 6: Map showing cells suitable for rainfed and irrigated agriculture.

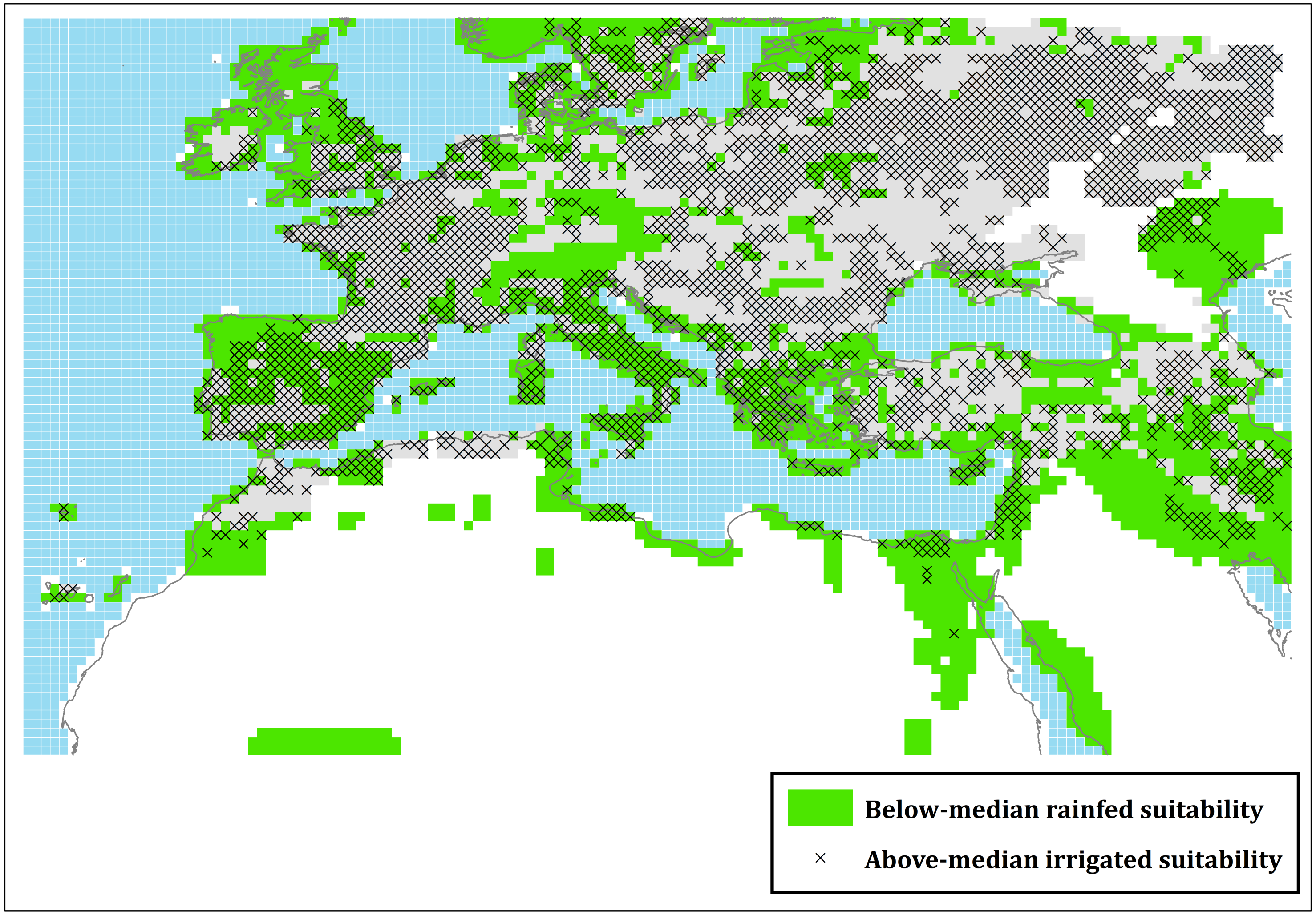
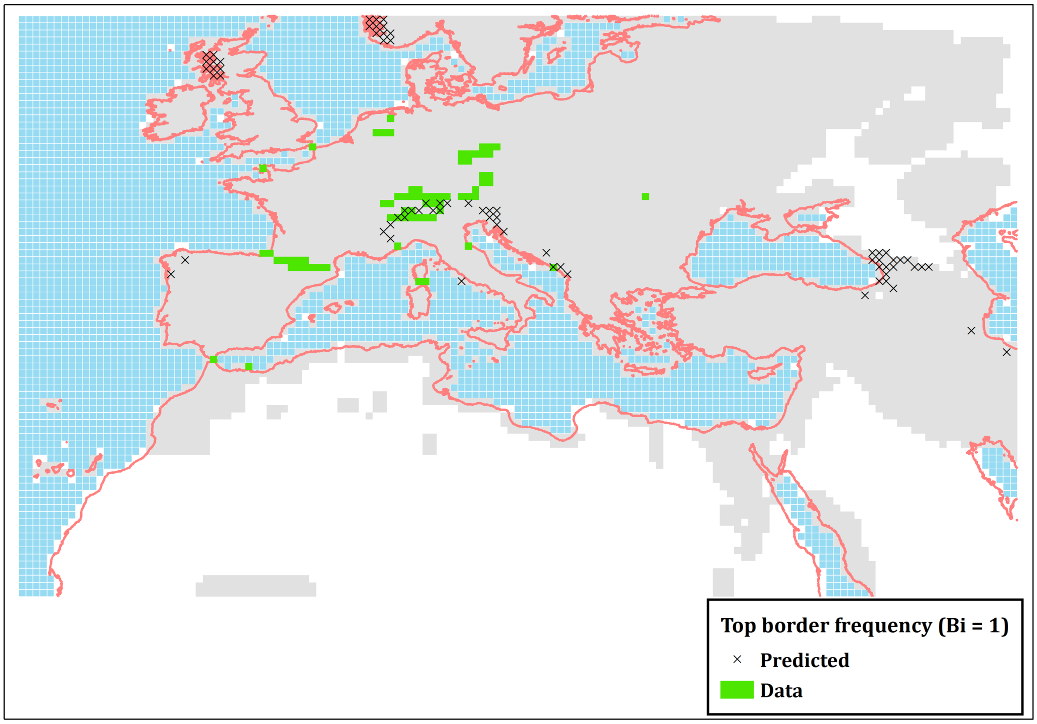
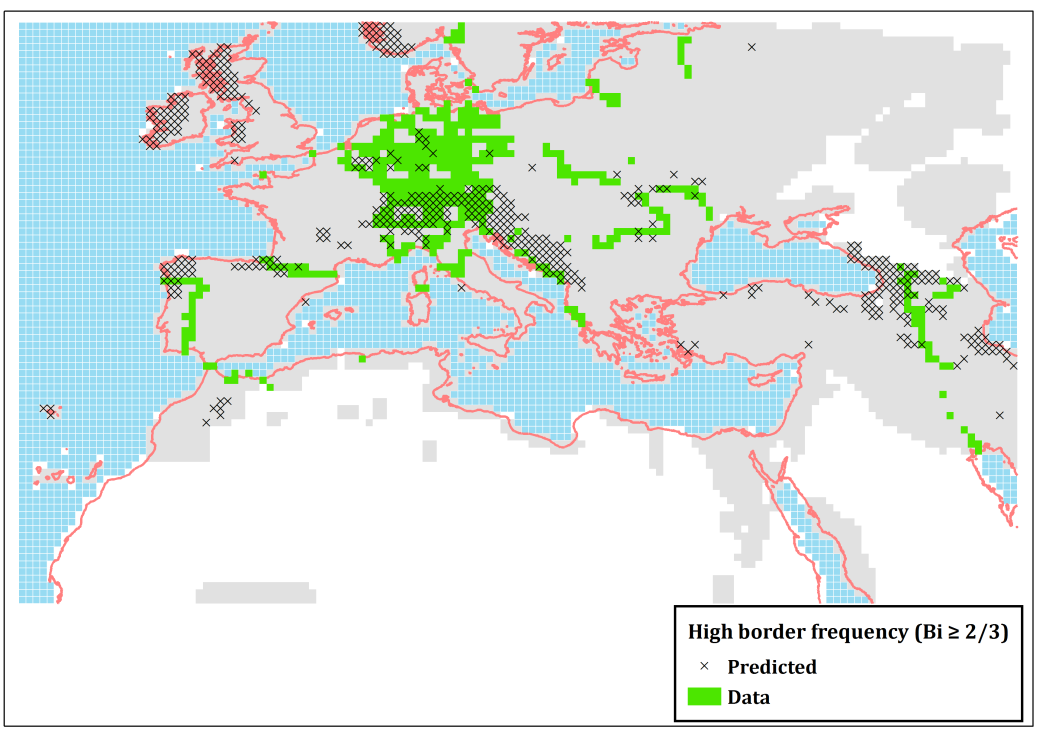


Figure 7: Maps of actual and predicted border locations.



(a) Borders all years 1500-2000 ($B_i = 1$)



(b) Borders at least 2/3 of the years 1500-2000 ($B_i \geq 2/3$)

Notes: Panel (a) shows the locations of the 70 cells with borders in all years 1500-2000 in the data, and the 70 cells with the highest predicted border frequency. Panel (b) does the same for cells with borders in at least four of the six years (454 cells in total).

Figure 8: Border frequency and modern outcomes.

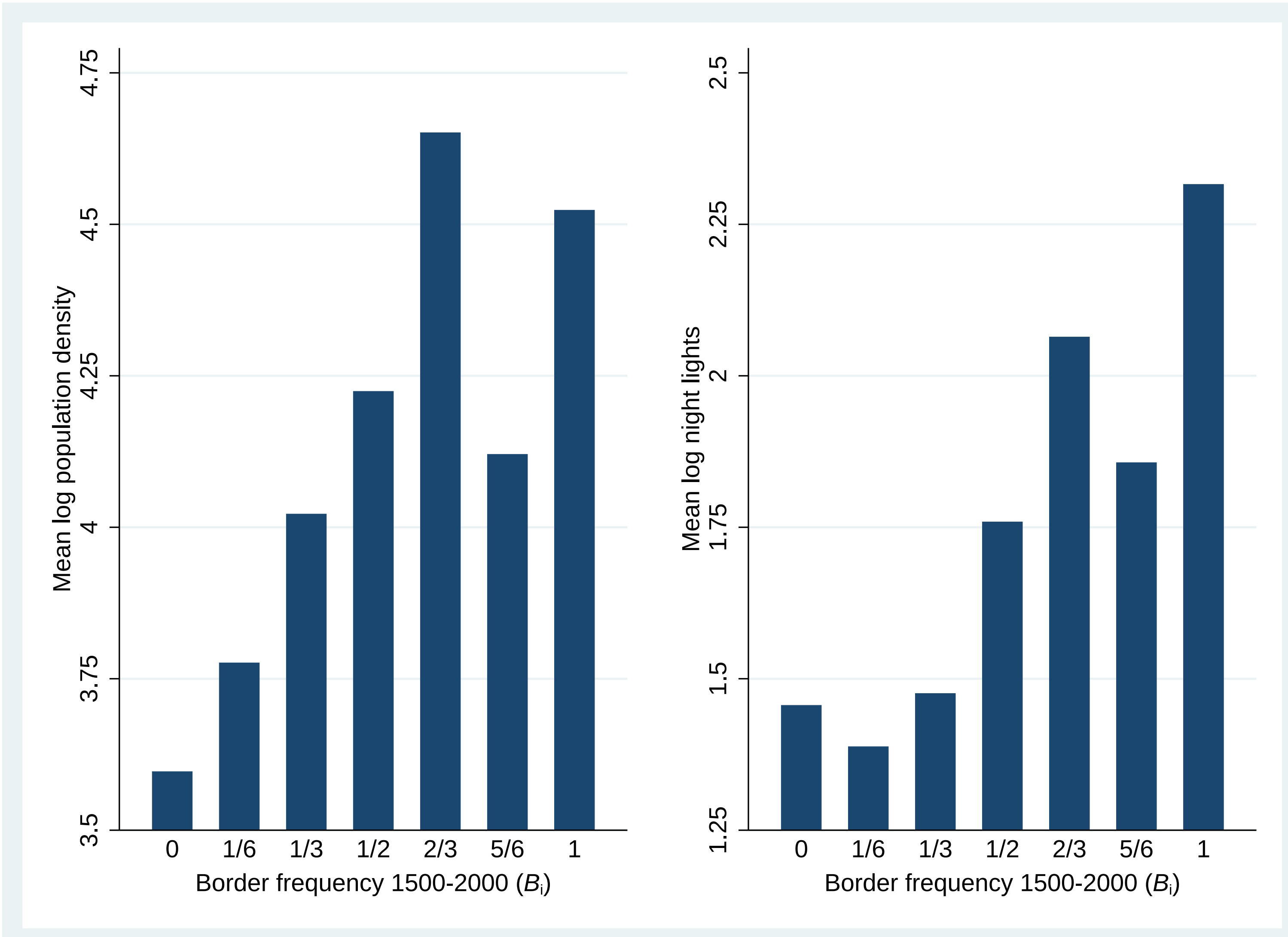


Figure 9: Map showing the locations of cells with borders in at least four of the six years (454 cells in total), and the same number of cells emitting the most night lights.

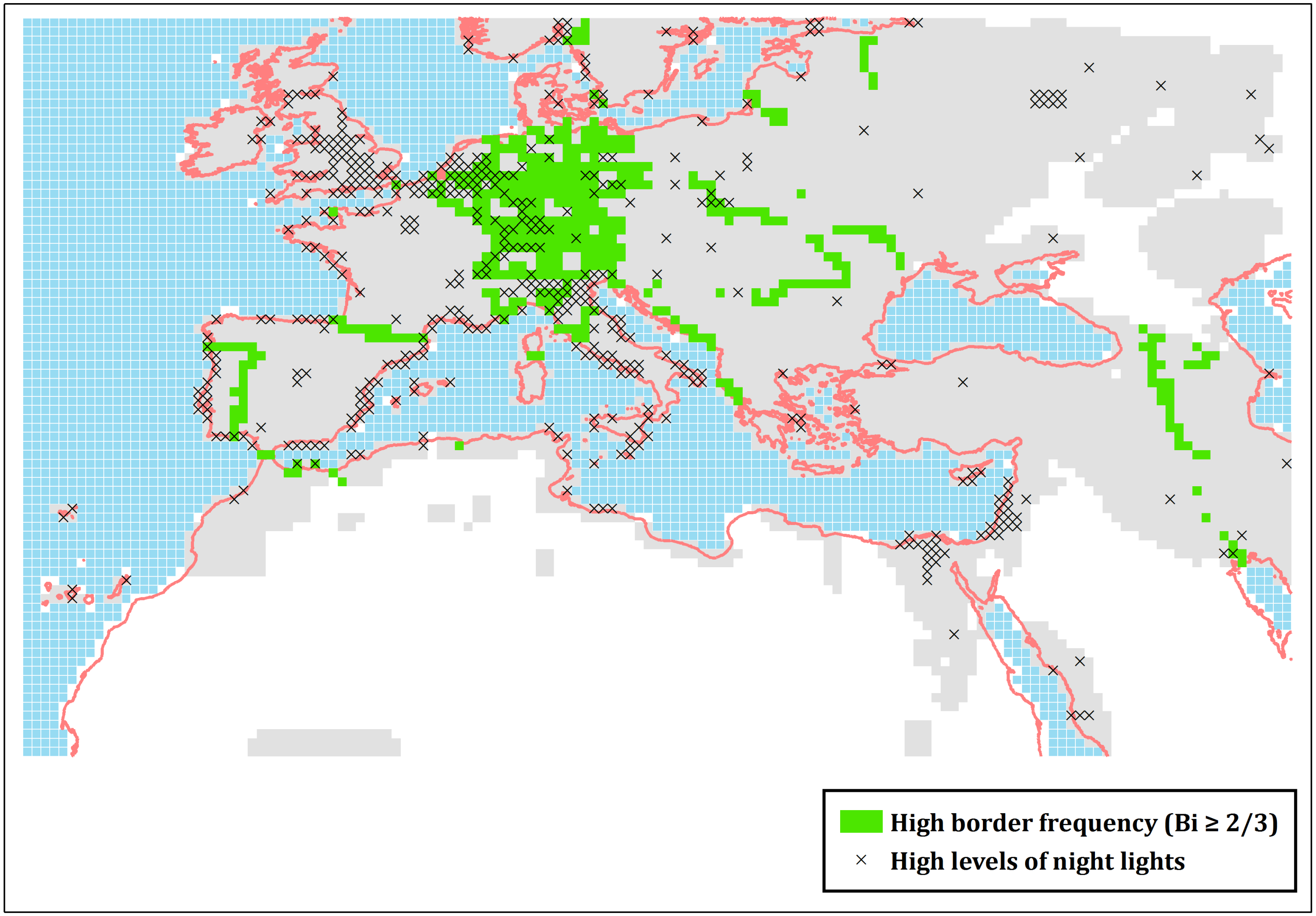


Figure 10: Map showing locations of cells with high and low border change (C_i).