**Geography, Economics and Political Systems: A Bird’s Eye View**

Elise S. Brezis¹ and Thierry Verdier²

**Introduction**

As early on as in the writings of Montesquieu and Adam Smith, geography plays a preponderant role in explaining disparities in the development of nation states. Smith placed an emphasis on topography, and especially on the role of coasts and rivers in the development of regions. For Montesquieu, climate was an essential element to economic development. His theory asserted that climate may substantially influence the nature and development of human societies.

For a long time, economists did not pay attention to geography, and this field was left to sociologists, historians, and geographers. Economists have only recently started to analyze the impact of geography, and this article will examine the various avenues of research that have been taken, wherein the notion of geography mostly encompasses matters related to location, soils, and topography; but also to climate and epidemiology.

The first part of the article focuses on the relationship between geography and economics, while the second part relates geography to political systems and public economics.

**Geography and economics**

The field of economic geography focuses on the study of the spatial aspects of economic activities, and examines the location, distribution, and spatial organization of economic activities, taking into consideration economies of agglomeration and the effects of distances and transportation.

Among the early economic geographers, the main researchers into spatial decisions were Hotelling, Weber, von Thünen, and Christaller. Hotelling emphasized how space matters for competition between firms, Alfred Weber embodies the precursor of location theory; Christaller developed the central-place theory analyzing the centripetal forces in play between the periphery and the city; and von Thünen’s cone theory focused mainly on the centrifugal forces linking the city and its outskirts, analyzing the efficient outcome of the development of the city at the core, and agriculture on its periphery.

These geographical arguments were not incorporated into mainstream economic research until the work of Krugman (1991a, 1991b), who found a way to analytically combine all of these previous elements into a single general equilibrium framework: the core-periphery model.

This core-periphery model examines the uneven development of regions and the emergence of clusters. On the one hand, there are some benefits to spatial concentration: firms have an incentive to locate in the larger market to exploit economies of scale in production, which can be due either to increasing returns or to scale externalities, including spill over effects like the backward and forward linkages. These are the centripetal forces that explain the dynamics of clustering together.

On the other hand, there are also centrifugal forces, which mainly encompass the negative effect of competition between firms clustering in the same city, congestion, and the higher prices of land and rents. To close the framework, the remaining elements affecting the spatial concentration of industry and development of regions are determined by distance and transportation costs. Depending upon distance and size of transportation costs, equilibrium will either be of a mono-core, or

---

¹ Bar-Ilan University.
² PSE-ENPC.
of two or more centers, with the agriculture sector on their periphery.\textsuperscript{4}

Related to this main core-periphery framework, other subjects of research have been developed. In connection with the remarks of Montesquieu, research has shown that tropical areas prone to infectious disease are poorer than temperate zone areas due to health effects leading to lower human capital and labor productivity, as well as lower output (Gallup, Sachs and Mellinger 1999; Bloom and Sachs 1998).

As far as Smith’s assertion that topography affects development is concerned, cross-country empirical research shows that GDP per capita is positively correlated with coastal proximity (Mellinger, Sachs and Gallup 2000). Moreover, 57 percent of income is generated in areas lying within 80km of the coast, while these areas account for only 13 percent of land mass (Rappaport and Sachs 2003).

The development of cities and regions is also dynamic over time. Cities first developed mostly in highly fertile regions like those near river valleys, at a time when international trade was not yet developed. Subsequently, due to trade, cities developed near the Atlantic, and in proximity to coal mines during the Industrial Revolution. Today, cities with high-quality universities are magnets for highly-skilled workers and high-tech industries.

Yet the new technologies have reduced the importance of geographic strategic position, as well as the disadvantage of residing in an interior region. The initial reason for a city’s development, such as the presence of a port or suitable topography, nevertheless leads to such concentration and linkages that, despite the disappearance of this initial advantage, the city tends to remain a center, as in the case of New York and London, as shown in Table 1.\textsuperscript{5} The stability of the big poles does not apply to medium-sized cities, whose concentration forces still rely on one or two sectors, as in the case of automobile industry based in Detroit. Historical evidence suggests that life cycles are an important aspect of urban history, and centers of population shift with economic activity, as shown in Table 1.\textsuperscript{6}

\textsuperscript{4} See Fujita and Thisse (2009) for a review of this theory.

\textsuperscript{5} On geography, economic activity, and path dependence, see also Davis and Weinstein (2002), Duranton (2007), Glaeser (2008) and Bleakley and Lin (2012).

\textsuperscript{6} See also Bairoch (1988). On the dynamics of the relative density of rural and urban areas, see Michaels et al. (2012).
This life cycle in the development of cities is related to economic activity, since the very success of an urban center in a traditional technology may put it at a disadvantage in terms of the implementation of a newer, ultimately more productive technology, as is shown in Brezis and Krugman (1997).

They highlight that when a new technology is introduced, for which this accumulated experience is irrelevant, older centers favor staying with a technology in which they are more efficient. New centers, however, turn to the new technology and are competitive despite the raw state of that technology, because their distance from the big center means that they offer lower land rents and wages. Over time, as the new technology matures, the established cities are overtaken. Therefore, during times of major technological change, medium-size cities are often overtaken by upstart metropolitan areas.

In conclusion, the new economic geography has analyzed the dynamics of clustering based on two main elements: first, the production function that displays increasing returns or knowledge externalities, and second, distance and transport costs. The new economic geography has not dealt with political variables. The following section of this article analyzes the relationship between geography and political systems.

**Geography and political systems**

**Geography and state formation**

The first question raised relating geography to political regimes was at the origin of the state itself: what causes groups of people and small communities to aggregate into political units as states, and how is this aggregation related to geographical factors?

The “hydraulic hypothesis” of Wittfogel was one of the first theories relating geography to the origin of the state. He claimed that small communities in a small geographic concentration cannot develop large-scale irrigation, and therefore several communities are compelled to merge. More generally, this theory emphasizes that small autonomous political units in a given geographic area cannot carry out the construction of infrastructure with high fixed costs, thereby leading to the formation of states. This Hydraulic Society thesis has run into difficulties since countries like China and Mexico became states long before the advent of large scale irrigation.

War and force have also been related to state formation. As Jenks (1900, 73) stated: “Historically speaking, there is not the slightest difficulty in proving that all political communities of the modern type owe their existence to successful warfare”. War lies at the root of the state. However, Carneiro (1970, 1978) asserted that if warfare is a necessary condition for the rise of the state, it is not a sufficient one, and geographic elements must be added.

Carneiro highlights the importance of conditions of territorial “circumscription” in the formation of primary states. This circumscription theory claims that the development of a state will start in areas circumscribed agricultural land, with a geographical border that does not permit inhabitants to flee rapidly in case of war and not to pay taxes to the victor. The coercive coordination due to geographic circumscription is therefore necessary for the creation of surplus, and is a causal factor in the origins of the state.

Carneiro opposes two regions with contrasting ecologies: one with circumscribed agricultural land, like the coastal valleys of Peru from which Andean coastal populations could not escape; and the other a region with extensive and unlimited land like the Amazon basin, where Amazonian populations could always disperse. Therefore, in the Amazon basin, there was no tendency for villages to combine into larger political units.7

The social circumscription theory, which claims that it is optimal to cluster, for purposes of both attack and defense, was adapted to this theory. It is more difficult for inhabitants in the nuclear area to escape attack than it is in the periphery, where ability to move is greater (Chagnon 1968).

As far as circumscription is concerned, Ashraf, Galor and Ozak (2010) exploited cross-country variation in the degree of geographical isolation, prior to the advent of seafaring and airborne transportation technologies, to examine its impact on the course of economic development across the globe. Their empirical investigation establishes that prehistoric geographical isolation has generated a persistent beneficial effect of cultural attributes that are conducive to innovation and thus contributed to contemporary variation in the standard of living across countries.

7 See also Diamond (1997) who proposes the “geographical fragmentation” as an explanation for technological innovation.
Geography, institutions and state capacity

The literature linking economic growth, institutions and state capacity has mainly focused on redistributive conflicts between social classes, arguing that the power of elites and elite interconnections influence the types of institutions chosen which affects economic growth. However, a new body of literature has arisen focusing on geography as an exogenous factor, which affects institutions and state capacity.

Historians such as Blockmans (1978, 1997) highlight the significant effect of geographic scale and compactness on the type of representative institutions of pre-modern European polities. Bairoch (1997) also stresses that the compactness of small states with strong regional autonomy in their political systems leads to early economic development. Along those lines, Stasavage (2010) argues that geographically compact polities can more easily sustain intensive forms of political representation. Stasavage’s empirical analysis using a dataset on representative institutions in medieval and early modern Europe suggests a strong effect of geographic scale on the formation of political representation. Diamond (1997) stresses that climatic, topographic and environmental characteristics, which favored the early development of stable agricultural societies, ultimately led to immunity to diseases endemic in agricultural animals, and the development of powerful and organized states.

One line of research into the role of geography, and its interactions with historical events, and into the evolution of institution-building has focused more specifically on Africa. Nunn and Puga (2012) and Nunn and Wantchenkon (2011) consider the historical case of the slave trade and its interaction with geographic topography and distance in the African context. Nunn and Puga (2012) look at terrain ruggedness and show, as expected, that rugged terrain generally has a negative direct effect on income, and economic development as it hinders trade and most productive activities. However, in the African context, this effect is more than compensated for by an indirect positive effect associated with the protection that rugged terrain provided to those being raided during the slave trades. As income is positively associated with institution building, this suggests that ruggedness of terrain could have been a positive contributing factor for institutional development in Africa.

Nunn and Wantchenkon (2010) investigate how the historical origins of civil mistrust within Africa may be associated with the exposure to slave trade. Combining contemporary household survey data with historic data on slave shipments, they show that individuals whose ancestors were heavily raided during the slave trade exhibit less trust in neighbors, relatives, and their local government. As an instrument for exposure to slave trade, they use the historic distance from the coast of a respondent’s ancestors, while controlling for the individual’s current distance from the coast. These results emphasize the role of geography in the building up of social capital and civic culture in Africa and via that channel, its long-term implications for building up legitimacy and trust in public and state institutions.

In the avenue of genetic research, Ashraf and Galor (2013) established that migratory distance from Africa and therefore genetic diversity is an underlying cause of a broad spectrum of manifestations of ethnic and cultural fragmentation, and their effect on social cohesion and civil conflicts.

Acemoglu, Johnson and Robinson (2001) examine the effects of colonialism mostly in Africa, and show that geography can also affect state capacity through the legacy of colonialism, especially corruption. In colonies with an inhospitable climate and topography, very few Europeans settled, and the institutions therein were created in extractive forms, giving way to corruption and embezzlement after the end of colonialism. In colonies with more hospitable climates, large number of Europeans settled, and the institutions therein were created in extractive forms, giving way to corruption and embezzlement after the end of colonialism. When such colonized countries gained independence, the Europeans left them more developed and less corrupt.

Focusing on the US, Galor, Moav and Vollrath (2009) establish that geographical characteristics that were complementary to large plantations generated a concentration of landownership that adversely affected the emergence of human-capital promoting institutions (for example, public schooling), and thus the pace of the transition from an agricultural to an industrial economy. This, in turn, contributed to the emergence of the great divergence in income per capita across countries. The prediction of the theory regarding the adverse effect of the concentration of landownership on education ex-
penditure is established empirically based on evidence from the beginning of the 20th century in the US.

As far as state capacity and efficiency is concerned, Alesina and Spolaore (2003) discuss the size of nations and raised the importance of distance on political integration decisions. Distance matters because of how it affects the trade-off between the efficiency of provision of public goods and the differentiated preferences of agents over such goods. They emphasize that the optimal size of nations results from a basic trade-off between economies of scale and congestion. Indeed, on the one hand, there are benefits of scale associated with the provision of public goods, which are cheaper when more taxpayers pay for them and larger countries can also better internalize cross-regional externalities, or natural shocks. On the other hand, however, there is congestion and a higher heterogeneity of preferences across citizens.

As far as fiscal state capacity is concerned, Mayshar, Moav and Neeman (2013) examine geographical factors that can affect extractive institutions and state capacity, and like Wittfogel, they stress the importance of water irrigation. They claim that differences in institutions found between Upper Mesopotamia and Ancient Egypt can be explained by the degree of informal transparency in the supply of water, which affects the ability to appropriate revenues. In particular, the homogeneity of farmland and irrigation systems increased the transparency of farming, thereby increasing appropriability. It is this difference in the system of water irrigation that led to differences in the power of the state, state institutions and political systems in the Ancient world.

Finally, the geographic topography may also interact with the incentives for taxation and redistribution in societies dominated by elites with limited state capacity. As Bourguignon and Verdier (2012) emphasize, the incentives of elites to tax, redistribute, and increase state capacity are affected by the complementarities or substitutability in the production process between the factors controlled by the elite and other social groups. These technological properties of the production process are themselves naturally dependent on the geographic and climatic context in which these societies evolve. As such, geography plays a role in the way elites are induced to design policies and institutions that best fit their interests, whether they are in congruence or not with the interests of the other groups in society.

Geography, conflicts and the transition process

A well-established body of literature on international relations argues that climate, topography, and location are important determinants of state behavior. Boulding’s (1962) research on conflict argues that the lower the costs of the transportation of goods, or of violence, the less likely the parties in a conflict are to have unconditional viability. Moreover, Sprout and Sprout’s (1965) environmental probabilism claimed that environment and climate affects the choices and actions of players in foreign policy, which in turn affects the international political realm.

With its focus on the microeconomics of the geography of political repression, Gates (2002) develops a model showing how geography interacts with ethnicity and ideology to determine military success and to shape rebels’ recruitment activities. He analyzes how this interaction affects the distance between insurgents and regimes; and how the distance between the insurgent group and the seat of government enables an insurgent group to expand.

Among the first empirical studies to include variables on topography, Fearon and Laitin (2003) argue that rough terrain (area poorly served by roads, which is measured by the proportion of the country that is mountainous) is an essential opportunity for rebellion, and for translating grievance and discontent into insurgencies or civil conflict. Collier and Hoeffler (2004) added some more indicators of both forests and mountainous terrain as causal variables explaining civil wars. Their findings are weakly supporting the expected positive mountain-war correlation, whereas forests fail to generate significant effects.

The transition process in Eastern Europe constitutes another interesting example of the impact of geography and space on conflicts and the evolution of political regimes (Fearon and Laitin 1999). More specifically, the relationship between geography and conflict can be applied to the analysis of the timing of the transition process of Eastern European countries.

It is interesting to note that the sequence of political liberalizations and economic transitions observed across Eastern Europe in the late 1980s actually began in countries located further from the “free world” frontier. Indeed, although the nearest country to the West was East Germany, Poland and Hungary were the first countries to undergo a change of regime, with
the process then spreading eastward and westward (see Figure 1).

As a matter of fact, one might reasonably expect that closer proximity to the West would facilitate support for dissidence against the old regime and therefore render a shift of regime more likely. The elements that explain this “non-linearity in the process of transition” are the effect of distance and topography on the effectiveness of the army in suppressing uprisings.

In the aforementioned literature relating geography and state capacity, geography was usually introduced through its effect on the provision of public goods. In this section, geographic distance matters through its effect on the effectiveness of the army and the police in dissuading dissidents from starting a revolution. At the heart of the analysis lies the assertion that the repressive capacity of an incumbent political elite is affected by the spatial ability of dissidents to escape the regime.

Brezis and Verdier (2012) assume on the one hand that geographic distance and topography affect the costs of escaping from a country: the closer the country to the “free world”, the easier it is for dissidents to escape the authoritarian regime run by the incumbent elite. All else being equal, this reduces the expected costs of dissidence, thereby stimulating stronger incentives for counter-regime activities, implying that power shifts are more likely in political regimes located closer to the “free world” frontier. This link therefore provides a first direct channel through which geography and space circumscribe the nature of political power and tend to favor democratization.

On the other hand, distance also endogenously affects the scale of repression, and this provides a second induced channel through which geography interacts with conflicts and politics. Indeed, the size of the repressive forces in a given authoritarian regime is not exogenous, but rather actually results from the choice made by the incumbent elite to remain in power. This feature implies in particular that the size of the repressive apparatus will be a function of the geographical characteristics of the country. As a matter of fact, countries with easier access to “safe havens” will face stronger repressive efforts from their elites than countries where such access is more difficult.

The rationale underlying this relationship is simple. Geographic distance and policing are substitutes from the standpoint of the expected cost of sanction as perceived by dissidents in a given country. The closer the country is to safe havens, the more likely its dissidents are to decide to flee whenever an uprising is repressed; and so the more profitable, in terms of perceived sanctions, it is for the regime to police those trying to flee. It follows that there should be more intensive policing and repressive apparatus in countries closer to the free world frontier than in countries located further from such frontier. This second induced effect of geography enters with opposite sign to the first direct channel of distance.

These conflicting effects of geography on political power shifts lie at the source of these intricate patterns of democratization processes. They may therefore explain the sequence of political liberalizations and economic transitions observed across Eastern Europe in the late 1980s, which took the form of a U-shaped function of distance, as shown in Figure 1.

In conclusion, this recent line of research into geography and political systems has emphasized the importance of topography and proximity on two fundamental dimensions of politics: the technologies of conflicts and control, and the degree of heterogeneity of preferences. Through such channels, geographic factors happen to be constraining factors for the emergence or sustainability of specific political regimes.
To date, however, the aforementioned analyses have essentially been confined to static spatial dimensions. A future consideration of geography might include dynamical spatial effects. Important questions for future research along these lines include the implications of geography for states’ competition and cooperation dynamics in international relations and how conflict or political systems diffuse from neighboring or regional states.10

**Conclusion**

This paper shows that there are many paths via which geography has shaped economic development and the formation of institutions, conflicts, and political systems. Yet, considering the fact that geography is one of the only elements that is an exogenous variable *par excellence*, it is quite striking that macroeconomics and political economy have not focused more intensely on its effects. This could change soon: with the further reduction of transportation costs (leading also to increasing costs of saturation, and to congestion), we could face the greater effects of geography on our economic and political systems. Therefore greater scrutiny on this exogenous variable will become necessary.

However, to characterize geography as an exogenous element may not hold anymore in the near future. Indeed, the new environmental economics emphasize that economic and political decisions influence our climate, our oceans, and Earth’s topography; geography becomes affected by economic and political systems. As a result, geography will no longer be exogenous: geography affects political regimes, which, in turn, affects geography.

Interestingly this co-evolution process between geography and social systems may give rise to important reinforcing complementarities and generate the possibility of multiple long-run geographic-institutional equilibria, highlighting the crucial role of history and initial conditions on the dynamics of political systems in the world. Ultimately, geography cannot be disentangled from history.

---

10 This diffusion effect is analyzed in Brezis and Verdier (2003) but without a geographical dimension.


