Geospatial Information Systems

Noel D. Johnson

At the 1969 meeting of the International Geographical Union Swiss geographer Waldo Tobler proposed what has come to be known as the First Law of Geography: "everything is related to everything else, but near things are more related than distant things" (Tobler 1970). While seemingly banal, Tobler's insight, when combined with modern geospatial information systems (GIS) software and the massive increases in desktop computing power that occurred in the 1990's, has led to a revolution in empirical social science. It's impact within economics has been influential in many of the more fundamentally spatial sub-fields such as urban and trade. However, arguably no field has been more affected by the GIS Revolution than economic history. This influence has manifested itself in two ways: by opening up new sources of data and by allowing for new avenues of analysis.

GIS and New Sources of Data

The combination of GIS software with powerful desktop computing has allowed economic historians to access and create both more and, in many cases, better data. These new sources of data have tended to be of two different types. First, GIS allows modern data sets to be effectively matched with historical units of analysis. One situation in which such matching can be appropriate is when the spatial variable of interest is relatively time-invariant and thus can be legitimately used to explain events in the past. For example, Nunn and Qian (2011) employ a modern measure of the soil's suitability to grow potatoes from the Food and Agriculture Organization (FAO) in order to evaluate the historical hypothesis (due originally to Crosby (2003) and others) that the introduction of the potato to the Old World after opening of the Columbian Exchange led to historical population growth. Nunn and Qian effectively leverage the highly disaggregated nature of the FAO data – – it covers roughly the entire world at a resolution of 0.5 degrees by 0.5 degrees (about 56 by 56 kilometers at the equator) – – in order to make their argument. In addition to

studies using modern soil suitability measures (e.g. Michalopoulos and Naghavi, 2014) other examples of modern GIS data that have been effectively matched to historical sources include measures of topography (Nunn and Puga, 2012) and climate (Anderson et al., 2017; Fenske and Kala, 2015; Grosfeld et al., 2017).

Another situation in which modern GIS measures have been effectively exploited is in studying the very long-run effect of some historical variable on contemporary outcomes. For example, Michalopoulos and Papaioannou (2013) is interested in the question, "what was the effect of pre-colonial state capacity in Africa on current economic development?" A major obstacle to studying this question, however, is that pre-colonial political boundaries in Africa differed a great deal from their present configuration. To solve this problem, Michalopoulos and Papaioannou exploit modern data on night light intensity captured from space using satellite observations. Since these night light data are both highly spatially disaggregated and correlated with economic activity, they can be used to test the authors' hypotheses. Other prominent examples of such persistence papers that use modern GIS data include Alesina et al. (2013), who look at how a region's soil suitability for either hoe or plow agricultural impacts present day gender preferences, and Galor and Ozak (2016), who look at a region's soil suitability for certain crops and measures of present-day time preferences.

A second major innovation in data creation brought about by GIS has been to facilitate the use of existing historical data sets in manner that allows for empirical analysis. For example data on Plague recurrence in Europe from Biraben (1975) have been matched to cities by Dittmar and Meisenzahl (2017) and historical maps of rail networks have been used to test theories of urban agglomeration and trade (e.g. Donaldson and Hornbeck, 2016; Donaldson, 2018; Jedwab et al. 2017). One effect of the ability to map historical data using GIS has been a general shift towards disaggregation. Whereas the natural unit of analysis used to be nations or states, in the last decade there has been a wave of papers focusing on cities as the unit of analysis (e.g. Bosker et al., 2013; Jha, 2013; Dincecco and Onorato, 2016). One notable early example of this approach is Dittmar (2011), who matches a database of the locations of early printing presses with historical city growth to investigate the effect of print technology on economic development.

To illustrate how GIS technology can be leveraged to analyze existing historical databases, consider recent work on the causes of witchcraft trials in Europe. For decades historians have been compiling lists on where and when individuals were tried as witches in early-modern Europe. Despite the fact that many historians proposed that witch trial frequency was associated with political fragmentation, there was not much that could be done in the way of consistent empirical work with these data until the advent of GIS. Consider Figure 1 which is adapted from Johnson and Koyama (2018) and maps out witch trial locations between 1300 and 1850 compiled by Leeson and Russ (2015) as well as political boundaries in Europe in 1500. These trials were collected from numerous secondary sources, but only when they have been geo-coded and mapped do obvious patterns begin to emerge. It is clear, for example, that the majority of the trials occurred in regions that were politically fractionalized (i.e., states were small). It should also be noted that an additional benefit of mapping the historical data is that shortcomings or important caveats in any hypotheses are also easier to spot. For example, in Figure 1, there are hardly any trials in the Iberian Peninsula, a strong clue that the dynamic of persecution of heretical belief was different there than in other parts of Europe.

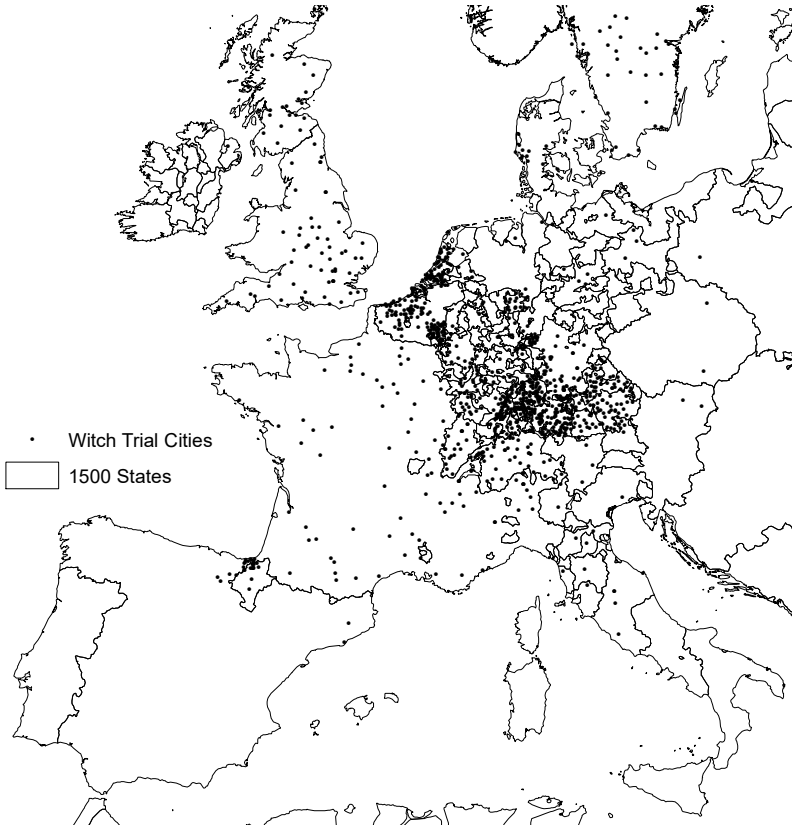

Figure 1: Locations of witch trials and executions, 1300-1850 and political borders in 1500.

GIS and New Avenues of Analysis

The GIS Revolution has not just been about the creation of new data sets. It has also introduced new options for the analysis of data. Many of these methods involve exploiting Tobler's First Law. For example, if things that are close to each other spatially are more related, then it is often possible to extrapolate the unobserved characteristics of a nearby object based on the observed characteristics of its neighbors. For example, Jedwab et al. (2017) observe Black Death mortality rates for a subset of cities in Europe between 1347 and 1352 but wish to predict the impact of the disease on the likelihood that Jews are persecuted in a non-overlapping set of cities. They thus use spatial interpolation techniques to predict the mortality rate for all cities based on their nearest neighbors.

Another way in which Tobler's First Law has been used is in the analysis of spatial networks. One of the ways in which GIS technology has facilitated these approaches is that it allows for the calculation of the least cost travel paths between locations. As an illustration, consider Figure 2. Adapted from Johnson and Koyama (2018), it illustrates the network of four different travel technologies in Europe during the early-modern period: seas, rivers, roads, and "none" or portage. Using historical sources it is possible to attach costs of traveling along these different routes. Then, one can apply an algorithm which determines what the least cost path is between two locations on the map as well as the cumulative cost of taking this path. In Figure 2 the least cost path between London and Rome is illustrated as an example. "Optimal path" measures such as these have played a prominent role in recent research. For example, Dell (2015) studies how Mexican drug trading networks reconfigured themselves after enforcement cut off certain routes.

Another highly fruitful area of research has been to investigate the effect of historical changes in infrastructure on market access. Market access is a measure of a location's potential for trade that combines information on surrounding market sizes

(usually populations) with the least cost travel cost of reaching those markets. A major benefit of doing analysis in terms of market access is that trade theorists have shown that under a set of assumptions it is a sufficient statistic to capture all of the direct and indirect impacts of changing costs on trade. Recent work focusing on market access includes Donaldson (2018), who looks at the effect of railroad construction in nineteenth century India and Donaldson and Hornbeck (2016) who look railroad construction in the US.

Figure 2: Travel technologies and the least cost path between London and Rome. Darkers colors correspond to higher cost technologies.

Tobler's Law has also been exploited in the creation of instrumental variables. An example of this in economic history is Dittmar (2011) who uses distance to Mainz – the origin city of Johannes Gutenberg's printing press – as an instrument for printing press presence in other European cities. Of course, a major concern with using geographic distances as IV's is that Tobler's Law might hold for more than a single characteristic, thus invalidating the exclusion restriction. One way researchers have attempted to

minimize this possibility is by testing placebo locations (e.g. does Wittenberg also explain printing press presence?). Another, more recent innovation, is to exploit both "push" in addition to "pull" forces in the IV. An example of this is found in Johnson and Koyama (2017). They are interested in predicting whether a city's population growth is affected by the presence of a Jewish community between 1400 and 1850. As an instrument for Jewish presence in a city, they exploit the fact that Jewish communities tended to locate near each other for historical reasons. In order to satisfy the exclusion restriction, they construct their instrument for city *j* as a variable measuring whether any city further than 100 kilometers away recently expelled its Jewish community (a "push") with a measure of how close that distant city is to *j* (the "pull").

One final way that GIS has played a major role in analysis is by exploiting failures in Tobler's First Law. Regression discontinuity (RD) approaches have played a major role in what has become known as the "credibility revolution" in economics. Only relatively recently have spatial RD's been exploited. This is true especially in historical contexts due to the fact that political boundaries have often arisen for random reasons – e.g. the front line of a battle may have been cemented by a peace treaty negotiated far away, colonial cartographers liked to draw in straight lines, or perhaps a relatively innocuous landmark was used to anchor a border. An early example of spatial RD is Dell's (2010) study of the effect of the Peruvian mining Mita on present-day consumption and stunting.

Getting Started with GIS

Given the benefits of using GIS, it should be in the toolkit of all economic historians. Unfortunately, the barriers to entry into the GIS world can be significant. Among these barriers is the fact that the most common program used for GIS is the Arc Suite developed and licensed for a substantial fee by the ESRI corporation. While most research institutions maintain licenses for on-sight use of the program, many faculty and graduate students will still have trouble gaining access. Alternatives include the open source program QGIS as well as the statistical package R, though the learning curve for the latter is steeper. If one has access to the Arc Suite, then a practical and free guide to the basics of using the program is Dell (2009).

By facilitating the creation of new data, allowing for the exploitation of existing data, and generating new empirical approaches, GIS has opened new approaches in economic history. Many questions asked by economic historians, even if not primarily spatial in nature, can benefit from taking spatial issues into account. As more maps are digitized and databases geocoded, the library of resources at the command of the researcher will certainly make this claim even more true.

References

Alesina, Alberto, Paola Giuliano, and Nathan Nunn. "On the origins of gender roles: Women and the plough." *The Quarterly Journal of Economics* (Oxford University Press) 128 (2013): 469-530.

Anderson, Robert Warren, Noel D. Johnson, and Mark Koyama. "Jewish persecutions and weather shocks: 1100--1800." *The Economic Journal* (Wiley Online Library) 127 (2017): 924-958.

Becker, Sascha O., and Ludger Woessmann. "Was Weber wrong? A human capital theory of Protestant economic history." *The Quarterly Journal of Economics* (Oxford University Press) 124 (2009): 531-596.

Biraben, J. N. *Les hommes et la peste en France et dans les pays européens et méditerranéens: La peste dans l'histoire.* Mouton, 1975.

Bosker, Maarten, Eltjo Buringh, and Jan Luiten Zanden. "From baghdad to london: Unraveling urban development in europe, the middle east, and north africa, 800--1800." *Review of Economics and Statistics* (MIT Press) 95 (2013): 1418-1437.

Dell, Melissa. "GIS analysis for applied economists." *Unpublished manuscript, MIT Department of Economics*, 2009.

Dell, Melissa. "The persistent effects of Peru's mining mita." *Econometrica* (Wiley Online Library) 78 (2010): 1863-1903.

Dell, Melissa. "Trafficking networks and the Mexican drug war." *American Economic Review* 105 (2015): 1738-79.

Dincecco, Mark, and Massimiliano Gaetano Onorato. "Military conflict and the rise of urban Europe." *Journal of Economic Growth* (Springer) 21 (2016): 259-282.

Dittmar, Jeremiah E. "Information technology and economic change: the impact of the printing press." *The Quarterly Journal of Economics* (MIT Press) 126 (2011): 1133-1172.

Dittmar, Jeremiah, and Ralf Meisenzahl. "DP12037 State Capacity and Public Goods: Institutional Change, Human Capital, and Growth in Historic Germany." 2017.

Donaldson, Dave. "Railroads of the Raj: Estimating the impact of transportation infrastructure." *American Economic Review* 108 (2018): 899-934.

Donaldson, Dave, and Richard Hornbeck. "Railroads and American economic growth: A ``market access" approach." *The Quarterly Journal of Economics* (MIT Press) 131 (2016): 799-858.

Fenske, James, and Namrata Kala. "Climate and the slave trade." *Journal of Development Economics* (Elsevier) 112 (2015): 19-32.

Galor, Oded, and Omer Ozak. "The agricultural origins of time preference." *American Economic Review* 106 (2016): 3064-3103.

Grosfeld, Irena, Seyhun Orcan Sakalli, and Ekaterina Zhuravskaya. "Middleman Minorities and Ethnic Violence: Anti-Jewish Pogroms in the Russian Empire." 2017.

Jebwab, Remi, Noel D. Johnson, Mark Koyama, and others. "Negative Shocks and Mass Persecutions: Evidence from the Black Death." Tech. rep., University Library of Munich, Germany, 2017.

Jedwab, Remi, Edward Kerby, and Alexander Moradi. "History, path dependence and development: Evidence from colonial railways, settlers and cities in Kenya." *The Economic Journal* (Wiley Online Library) 127 (2017): 1467-1494.

Jha, Saumitra. "Trade, institutions, and ethnic tolerance: Evidence from South Asia." *American political Science review* (Cambridge University Press) 107 (2013): 806-832.

Johnson, Noel D., and Mark Koyama. "Jewish communities and city growth in preindustrial Europe." *Journal of Development Economics* (Elsevier) 127 (2017): 339-354.

Johnson, Noel and Mark Koyama. *Persecution and Toleration: The Long Road to Religious Freedom.* Cambridge University Press, 2018 (forthcoming).

Leeson, Peter T., and Jacob W. Russ. "Witch Trials." *The Economic Journal* (Wiley Online Library), 2015.

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

Nunn, Nathan, and Diego Puga. "Ruggedness: The blessing of bad geography in Africa." *Review of Economics and Statistics* (MIT Press) 94 (2012): 20-36.

Nunn, Nathan, and Nancy Qian. "The potato's contribution to population and urbanization: evidence from a historical experiment." *The Quarterly Journal of Economics* (Oxford University Press) 126 (2011): 593-650.

Tobler, Waldo R. "A computer movie simulating urban growth in the Detroit region." *Economic geography* (Taylor & Francis) 46 (1970): 234-240.