Daniel Sui has sent me a written version of comments presented by five geographers at a panel on the first law of geography organized by him at the 2003 AAG meeting in New Orleans. The comments seem to fall into two camps: some reject the idea of “laws” in geography, and others feel that my notion has been of some merit. Interestingly, several other laws are cited in the comments; two by Isaac Newton (gravity, motion), two additional “first” laws (ecology, social science), four additional laws applying to people (utility maximization, primate city, human behavior, and Gresham’s), and three “second” laws (thermodynamics, spatial heterogeneity, things are never equal—the last two being only suggestions). Curiously, they do not mention other well-known laws, such as those of Zipf (1949), or Ravenstein’s (1885) 10 laws, or a second law of geography by proposed by Arbia, Benedetti, and Espa (1996) and one by myself (Tobler 1999, 87).¹

The comments then focus on three topics. One is a discussion of what constitutes a law, and whether “the first law of geography” fits into the appropriate definition. Much of the remaining material examines the concepts of “related” and “near.” Smith is quite correct in pointing out that this discussion would never have taken place if the specific word law had not been used. And Barnes is also correct when he puts it into context and observes that, by restricting myself to local effects, I used the notion to parse the possible complexities of simulating urban growth.

I am a great believer in simplicity, when this is possible. For example, the point in science is to achieve as many results as possible with the fewest hypotheses. So, in order to simplify the problem of depicting the growth of population in the Detroit region, I tried to eliminate complicating factors. This is when I invoked “the first law of geography: everything is related to everything else but near things are more related than distant things.” Doing this allowed me to concentrate on local effects—using the idea of a change in the “unit inhabitant,” and ignoring many other possible influences. As a result, Miller calls my simulation effort “crude,” whereas Barnes considers that it invokes complex ideas. Miller would undoubtedly consider economics, transportation, geology, and other factors to produce a much richer, but also more involved, model. Curiously there is no discussion of other concepts from the 1970 paper.

On Laws

Is the first law of geography simply an observation better referred to as an observed regularity or a principle and best not called a law? Semantics aside, how can one tell when something qualifies as a law? So what makes an observation a law? Two of the discussants on the panel placed great emphasis on this question. One, like many social scientists, believes that there are not, and cannot be, any such thing as laws where human behavior is involved. One argument is that once an empirical regularity—perhaps qualifying for the status of a law—is discovered, it will be modified by the people involved and thus rendered inoperable. Another strongly held view is that looking for laws of human behavior is a misguided effort involving “physics envy.” This approach argues that social science should be action oriented, and that the purpose is to change the world, or at least society. Many contemporary writers on human behavior represent this attitude, including Flyvbjerg and Giddens to whom I return later.

Smith points out that his stance is one of faith, not refutable by any evidence (“beyond validation or falsification”), and he uses philosophical citations to argue that “the first law” is not a law at all. I would counter this by observing that scientists have demonstrated that the pronouncements of philosophers are often suspect and incorrect (Reichenbach 1952; Lakatos and Feyerabend 1999, 19–112). A mathematician (Simmons 1992, 150) is even more severe, writing:

The growth of empiricism and the rise of science over the past three centuries have made it almost impossible to take seriously the extravagant pretensions of the a priori
philosopher, who sits in his study and spins a web of words, fanciful imaginings, and empty speculations out of the material of his own consciousness. Faith in reason alone is alien to us, and we believe that only careful observation and experiment can reveal anything of substance about the actual universe.

Kant is the classic example. His ideas concerning geometry were completely wrong as demonstrated by the invention of non-Euclidean geometries. Therefore a preferred—by me—definition of a law comes from a scientist, namely Richard Feynman (1967). In his book on *The Character of Physical Law*, he describes how to invent a new law. He points out that the first, and most difficult, step is to guess (Feynman 1967, 156). Then the criterion becomes to "compute the consequences of the guess to see what would be implied . . . . Then we compare the result . . . with experiment or experience . . . [or] observation to see if it works. If it disagrees with experiment it is wrong." So guess again. He also points out that laws can only be discovered by doing something radically different. The procedure used by previous discoverers of laws, therefore, cannot succeed. Feynman suggests that there is no fixed method that can lead to the discovery of laws.

Ravenstein's (1885) "Laws of Migration," although not given in a formal form, do allow prediction as pointed out by Dorigo and Tobler (1983)—in another laws' paper. Ravenstein, referring to migration between counties in Great Britain, wrote,

> even in the case of "counties of dispersion" which have a population to spare for other counties, there takes place an inflow of migrants across that border which lies furthest away from the great centres of absorption.

—(1885, 191)

Consider the case of Mexico, for which the center of absorption is the United States. Mexico, as predicted by Ravenstein, has an (illegal) immigrant problem, with people coming in from its neighbor to the south. This is clearly a correct prediction from the "migration law." This makes the case for the justification of the word law by Ravenstein, using the Feynman criterion.

As another example, in Tobler and Weinberg (1971), the spatial gravity model is inverted to make a prediction of the location of ancient Hittite villages in central Turkey. Is not the empirical efficacy of the widely used gravity model of human interaction, whether in its simplest form or as an entropy model, a sufficient demonstration of the validity of the first law of geography?

In Flyvbjerg (2001, 44–45) we find an assertion that the social sciences are incapable of producing laws. He advocates an action-oriented, invariably interventionist, social science and cites Giddens (1982) who later writes that it "is preferable not to use the term [law] in social science" (Giddens 1984, 347). In some ways Giddens' entire oeuvre emphases this theme. I do not find these pronouncements very creditable. Just as in previous centuries it was widely believed that humans could never travel to the moon, such assertions are not proofs.

Miller makes the further point that Newton's law of gravity does not explain the phenomena, it just describes the effect. His laws of motion do the same. Thus, reference to the need for an explanation, or a "cause," is not a useful argument, but more of a human psychological requirement. People seem to have a need to find "causes," and, therefore, laws about human behavior that only describe, but do not explain, are regarded as being inadequate.

In my reply to the discussants, I seem to have been forced to defend the first law of geography. Consequently, I looked around for several additional laws pertaining to human behavior. These include several other references to "laws" in the social sciences. For example, Zipf's several laws based on his principle of least effort—one on spatial interaction and another on word frequencies—are often cited, as is Pareto's law in economics relating to the distribution of incomes. The Auerbach-Pareto-Zipf law of city sizes is analyzed by Mandelbrot (1965). He also mentions (1965, 322) an Estoup-Zipf law of word frequencies. Lotka (1929) proposed a law concerning the frequency distribution of scientific productivity; Bradford's law (Garfield 1980) is similar. Also in economics there is Say's (1803) "law of markets: supply creates its own demand" (Patinkin 1948, cited in Weintraub 2002, 159). In sociology Merton (1973, 16) mentions Scheler's "law of three phases," and Thorndike's "law of effect" comes from psychology (cited in Lewin 1951, 27, 66). The Stanford nonmetric scaling pioneer psychologist Roger Shepard has a law attributed to him in relation to multidimensional scaling: "if a solution exists, probably it exists in two dimensions" (Coxon 1982, 87). The French engineer Lalanne in 1875 proposed "la loi des distances multiples," having to do with the distribution of towns and route connections, anticipating Christaller by half a century (Palsky 1996, 103). There is also a "first law of cognitive geography" (Fabricant et al. 2002). But my favorite discovery by far is David Lodge's "law of academic life: it is impossible to be excessive in flattery of one's own peers" (Lodge 1984, 152; emphasis as in the original—cited in Sokal and Bricmont 1998, 259).

I have pointed out that in social science some laws do exist, or at least that some regularities have been called laws. A longer discussion about laws in geography can be found in Golledge and Amadeo (1968). Also see Hempel...
Near Things

My use of the term near was equally ambiguous, as pointed out by several of the discussants. Geographers (e.g., Gatrell 1983) have often studied distance in its various forms. In addition to the metrical spherical (or ellipsoidal) geodesic (as the crow flies) distance there are distances in units of time, or travel cost, or intervening opportunities, city block distance (the so-called Manhattan metric), and various Riemannian or Finsler distances. Some speak of social distances, time-varying network distances, topological distances, genetic distances, ordinal distances (far, farther, farthest), and so on. Often these distances are not symmetric. We refer to the friction of distance. Thus, proximity and near can take on many meanings in different situations.

One interpretation, incorporating both relations and nearness, is the auto-covariance (or variogram) function, known for at least 40 years (Gandin 1963). This, along with Kriging and interpolation, seem to be the context in which the first law is most often cited, as evidenced by an examination of the phrase as referenced by Google on the Internet. In migration studies I have found strong spatial adjacency correlation of effects in movement patterns (Tobler 1995). The angle of repose in dry sand results in a phenomenon that is quite similar to spatial decay in social interaction as observed by Hägerstrand (1957), although the reason for the two is very different.

Multidimensional scaling, a frequently used technique in the social sciences, explicitly makes the assumption that similarity implies nearness (Goodchild and Janelle 1988). Seriation methods used in archaeology, with respect to closeness in time, and ordinations in political science and psychology invoke related ideas (Hubert 1974). Both the archaeologist and the historian have as one of their central dogmas the notion that the present is influenced by the past. There is a similar relation between events in space. Geographers are dogmatic that what happens at one location is influenced by events at other places. This relation between events and places allows one to make spatial predictions. For clarification this may be compared with temporal prediction.

The method of radiocarbon dating assumes that the degree of similarity of samples can be translated into a time difference to predict a date. Glottochronology—the study of language change over time—has a similar objective and uses a comparable strategy. It is occasionally used to estimate how long ago two culture groups separated. In these studies we can recognize the assumption that the long ago past has less influence than the recent past. Geographers assert that there occur comparable decays of similarity in space. Thus, given spatial locations (that is, latitudes and longitudes), one can to some extent predict (or estimate) the amount and types of interaction, or the degrees of similarity, between these locations. Is it more interesting when put the other way around: given the degree of similarity, or interaction, between places, can one predict their locations (that is,
the latitude and longitude coordinates) of the spatial origins? For example in 1962 W. Bunge suggested the use of geographical central place theory to predict the location of Mayan cities in Central America. Similarly, Susan Weinberg and I (Tobler and Weinberg 1971) predicted the location of many pre-Hittite towns in Cappadocia; this is an as yet unverified prediction. Philosophically, of course, prediction of this type is generally impossible, but, empirically, it is done. Gauss successfully predicted the location of Ceres, and Schliemann found Troy.

In one sense, prediction is actually too easy: tomorrow’s weather is likely to be like that of today. Both temporal and spatial forecasting must be discounted for the effects of persistence. Is this another instance of the first law of geography?

In economics, agglomeration tendencies result in geographic concentrations of businesses that are related to each other (Baldwin et al. 2003) and there are “spill over” situations (Palinck and Nijkamp 1975). The opening of a shopping center generally results in stores springing up on the periphery. These are all manifestations of the “stickiness” of temporal and spatial effects.

There is also obvious evidence that one must be carefully critical in applying the first law of geography. Anisotropic effects do occur, and so do discontinuities. The world is not always regular and predictable. For example, topographic interpolation in the badlands of South Dakota does not work well. The Mal Pais area in New Mexico is also a region of rough topography. Perhaps the name badlands (or Mal Pais) is in recognition of the local failure of the first law of geography.

Is not much social strife, including urban riots and racial discrimination, due to a perceived incompatibility between neighbors? Spatial autocorrelation need not always be positive, as evidenced by the NIMBY controversy. When one wanders into a new environment—a previously unexperienced part of a city, for example—it really does feel “strange.” Why is this? Can it explain a part of the attraction of tourism?

The Discussion

To turn to the individual contributions, Barnes seems to feel that the context of the discovery of a law somehow affects its validity. As anecdotal information, I have always found short historical commentary provides insight and puts the human nature of science into perspective but does not really change the result. Thus the validity would not be affected, even if the inventor’s (discoverer’s) name, or place in history, were not known. As mnemonic shorthand the names—the attribution of laws to individuals—are useful. Weintraub (2002, 205) notes that acceptance of a theory is a social process but also (p. 189) writes “Understanding the world in which Newton lived and made his contributions offers insight into the formation and acceptance of his contributions without denying the truth of his theories.” And Newton spent much of his time on mystical and alchemical subjects. Does this detract from his “laws”?

In the present instance the development of analytical geography and geographic information science has combined to produce a fertile ground for statements such as the first law of geography. My own background and mathematical interests also contributed, as Barnes rightly observes.

Miller expands on the notions of “related” and “near.” He does this in the context of spatial analysis. Here one notes immediately the distinction observable between the geographer as “natural scientist” and the individuals who are more interested in verstehen, a distinction nicely made by Catton (1965). The latter investigators hardly deal with autocorrelation or interpolation and similar techniques. Nor do they use analytical computational methods in studying “nearness.” He, Miller, goes on to point out how nearness can be extended to include temporal relations, adding “when you are” to “where you are.” This seems like a useful concept. Certainly, the long-ago past influences us less than the recent past. The statistician recognizes this in the concept of an “auto-regressive” model, also briefly explored in my paper “Cellular Geography” (Tobler 1979).

In “Doing Justice to the Law” Phillips is concerned, inter alia, with causal relations and takes as the point of departure his field of physical geography. He then rightly points out the many contingencies and conditions that must be attached to the statement of the first law of geography. In my paper I asserted that I need not take into account all of the activities throughout the world, thus ignoring butterflies in Brazil, as well as population growth in that country and immigration from Brazil. Nor do I know of any climate modelers who concern themselves with butterflies, even though the climatic system is multifaceted. Lewin’s “field theory” also excludes “events occurring at a remote distance” (1951, xii). On the whole I am in agreement with Phillips’s indication of the complexity and contingent nature of our field of study. Most model assertions require the ceteris paribus [everything else being equal] assumption.

I find that Smith fits quite well into Catton’s model of the verstehen class of scholar, also emphasized by Giddens and Flyvbjerg. He makes many good points, but I find it difficult to see even a remote connection between population growth in the Detroit region to shoes and
blisters. He deprecates the debasement (perversion? evolution?) of language, but I do not find this particularly important; rather, if true, it seems a natural human linguistic trait.

The panel discussant, Goodchild, takes a stance similar to that of Miller, focussing on geographical information science, and the attendant analytical procedures. He goes beyond this to speculate on further “laws,” making the important point that the first law, as constituted, deals with a bivariate relation, a second-order effect. He then considers other, perhaps more fundamental, possible “principles” of geographic science. Spaces of social interaction, mentioned by Goodchild, are a much-studied, dynamically growing endeavor, often using the multidimensional scaling principle, which equates similarity with distance, but even allows the influence of distance to vary from place to place, and varying by direction (Davies and Coxon 1982). His closing paragraphs remind me of a statement found in Lösch (1954, 363), namely that the purpose of theory is “to test reality,” not the opposite, as is often assumed.

Goodchild also introduces an interesting argument with respect to the physical environment, namely “what if [the first law] were not true?” He interprets this as meaning that there would be infinite local variation and, consequently, complete unpredictability. Every spatial environment would become excessively chaotic. Does this also apply to social environments?

He also raises the intriguing question of whether the law can be extended to other spaces. Could it be that the connectivity and interaction of neurons in the human brain might exhibit properties similar to those observed in society? Within an order of magnitude, both seem to have the same number of components, and both are characterized by relatively sparse connectivity matrices and exhibit local interaction. What I mean by this is to suppose that you could label all of the human neurons and then list them in a table, with the source neuron on the left margin and the destination neuron across the top. You would then have a table that is approximately 10 billion by 10 billion in size. Indicate by a one in the table all those neurons that send messages directly to another neuron. Each neuron seems to interact with about 1,000 others, so that the table is asymmetric but mostly empty. Permute the table so that the interacting neurons lie near the diagonal, and the sparseness of the table is then more easily seen. Are the neurons now ordered by their spatial separation?

Next do the same with all of the people in the world. This conceptual table is approximately 6 billion by 6 billion in size. But most people routinely interact directly with only a few others, perhaps also only a thousand or so. It is again a sparse matrix, and the permutation might arrange the people by geographic distance. In spite of this sparsity the “small-world” phenomena suggests that everybody is only about six steps away from everybody else (Milgram 1967). Is there a similar effect in the neurons in the brain? Perhaps this is a clue as to how the brain works, or how society works. What is left out here is the temporal variation of the connections.

One worm (caenorhabitis elegans) has been studied to the point where its nervous system has been completely mapped. It has 282 neurons and the topic of its connections—as many as 74 for one neuron—is actively being pursued (Watts and Strogatz 1998).

Similar to this, it is my understanding that when one receptor in the human eye is activated nearby (i.e., adjacent) receptors are inhibited. In both of these situations geometry plays an important part.

Conclusion

To summarize a long discussion, it can be seen that there are quite a number of references to laws relating to human behavior. Just as my reference to things being related is not terribly novel, the assertion about near things being similar also has lots of precedents. Perhaps what is unique is that I put these two things together and called the result a “law.” The fact that near things are more related than distant things seems a fundamental property of geography and rather easily explained.

Finally, I would like to convey my thanks to Professor Sui for arranging this discussion panel and for challenging and stimulating questions concerning my reply. The participants, Barnes, Miller, Phillips, Smith, and Goodchild, have all taken their obligation seriously. Their contributions have led me to contemplate many ideas and references with which I was previously unfamiliar. It also has provided a forum for a study of the diversity of viewpoints within the field of geography. I did not expect such a discussion when I wrote that paper in 1970. I was just having fun doing an animation in order to bring time into geography more explicitly. Another instance of the law of unexpected consequences?

Notes

1. Arbia’s second law of geography reads “Everything is related to everything else, but things observed at a coarse spatial resolution are more related than things observed at a finer resolution.” This suggests that aggregation has a smoothing effect, as is well known (see Tobler 1969, 1990). My second law of geography asserts that “the phenomenon external to [a geographic] area of interest affects what goes on in the inside; a sufficiently common occurrence as to warrant being
called the second law of geography." This is comparable to the need for "boundary conditions" in many physical problems. It also relates, perhaps inversely, to Foucault's (1979) emphasis on "confinement."

2. Giddens (1984) also comments extensively on the relation between geography and sociology. Concerning distance, he states (p. 363), "Distance in space is apparently easy to comprehend and to cope with conceptually" [sic]. He goes on to say, "human beings do make their own geography." He finds great value in the concept of regionalization (p. xxv), but eschews the idea of laws in the social sciences (pp. xxxii–xxxiv).


4. This phrase, "the law of unexpected consequences," is quite common—as can be seen by entering it into Google on the Internet—and apparently, quite well understood, though the origin is unclear. The distinguished historian H. Wayne Morgan, for example, uses it in his 1998 Brewster lecture in Greenville without his feeling any need for justification or elaboration.

5. An anonymous reviewer suggested looking at a book by R. A. Fisher (1935). I did this and found the following on page 66: "the widely verified fact that patches in close proximity are commonly more alike, as judged by the yield of crops, than those which are further apart."

References


