Big Data, social physics, and spatial analysis: The early years

Trevor J Barnes[1] and Matthew W Wilson[2]

Abstract

This paper examines one of the historical antecedents of Big Data, the social physics movement. Its origins are in the scientific revolution of the 17th century in Western Europe. But it is not named as such until the middle of the 19th century, and not formally institutionalized until another hundred years later when it is associated with work by George Zipf and John Stewart. Social physics is marked by the belief that large-scale statistical measurement of social variables reveals underlying relational patterns that can be explained by theories and laws found in natural science, and physics in particular. This larger epistemological position is known as monism, the idea that there is only one set of principles that applies to the explanation of both natural and social worlds. Social physics entered geography through the work of the mid-20th-century geographer William Warntz, who developed his own spatial version called "macrogeography." It involved the computation of large data sets, made ever easier with the contemporaneous development of the computer, joined with the gravitational potential model. Our argument is that Warntz's concerns with numeracy, large data sets, machine-based computing power, relatively simple mathematical formulas drawn from natural science, and an isomorphism between natural and social worlds became grounds on which Big Data later staked its claim to knowledge; it is a past that has not yet passed.

Keywords

Social physics, Big Data, spatial analysis, computerization, William Warntz, macrogeography

Introduction

> Physicists not only know everything; they know everything better. (Stauffer, 2004: 1)

It is easy to think that Big Data has no history, that the clouds parted and Big Data appeared. That is the sense in Anderson's (2008) well-known account in *Wired*. For Anderson, the past is severed from the present. Nothing that occurred in the past constrains what will occur in the future. That is why history is irrelevant. The past is ignored because nothing should hamper or limit what is to come. It is a classic modernist move. In this case, only the bright and uncluttered future of Big Data matters (Barnes, 2013; Wilson, 2014).

Rather than focus on the future, our paper is about the past. William Faulkner (1994: 73) famously said, "the past is never dead. It's not even past." For Faulkner past and future are inextricably joined. We cannot escape history because the past never fully passes. Instead, we always carry the burden of the past into the future. The purpose of our paper is to unpack some of Big Data's historical burden. Its past is not a unified whole, however, a single complete history. It is much looser, fragmented, multiple, set in relation to many different historical processes (Jacobs, 2009).

Consequently, Big Data's past is complicated and fraught. It is tangled up with, among other things, complex histories of computerization, superpower weaponry and defence strategies, military funding,

[1]University of British Columbia, Canada
[2]University of Kentucky, USA

Corresponding author:
Trevor J Barnes, Department of Geography, University of British Columbia, 1984 West Mall Vancouver, BC, Canada V6T1Z2.
Email: tbarnes@geog.ubc.ca

commercialization and advertising, corporatization, government regulation, epistemological fashion and debate, and academic disciplinary research agendas. Clearly, we can't follow all these variegated histories in this short paper. Instead, our aim is much more modest. It is to set out historically only one of the antecedents of Big Data: social physics. We suggest that social physics is connected to Big Data historically through spatial analysis: social physics, as we will argue, influenced spatial analysis, and spatial analysis was then incorporated into Big Data.

We recognize that not all spatial analysis derives from social physics, however. For example, within spatial analysis there is the long tradition of German location theory that includes von Thünen, Weber, Christaller, and Lösch that doesn't subscribe to social physics at all (Barnes, 2003). The underlying rationales for the individual works comprising the German locational tradition vary with the particular author, and are also bound up with the historical and geographical context in which the particular theory was devised (Barnes, 2003). In none of these cases is social physics the motivation, though. But there is a strand of spatial analysis that comes from social physics, and which in turn has shaped contemporary practices of GIScience and Big Data (see Anselin, 2012; Goodchild, 2004, 2009). In particular, we contend that practices of exploratory data analysis, pattern recognition, and data mining within GIScience and Big Data, which have produced such analytical outcomes as prediction, modelling, intelligence, and design, stem in part from an earlier spatial analysis and social physics. This is not to preclude other influences on Big Data and other forms of spatial analysis, but it is to make the connections between social physics and spatial analysis explicit.

Social physics is a long-standing interdisciplinary mode of inquiry. From the beginning it was driven by a belief that laws, theories, and models of physics applied as much to social as to natural worlds. This view is known as monism, the idea that theories of physics explain both human and non-human worlds (and also represented by the chutzpah found in our paper's epigraph). The origins of monism are in Ancient Greece, but social physics as a specific form of enquiry did not emerge until the 17th century in Western Europe. Following several hundred years of fits and starts, social physics was finally institutionalized in the 1940s, the result of the energetic efforts of a motley of American academics of which the two most important were Harvard's George K. Zipf (1902–1950) and Princeton's John Q. Stewart (1894–1972).

Social physics entered geography in the mid-1950s through the American geographer William Warntz (1922–1988), who later became Director (1968–1971) of the Harvard Laboratory for Computer Graphics and Spatial Analysis. Warntz, who collaborated with Stewart early on, emphasized the importance of recognizing general geographical patterns through the spatial analysis of large sets of geo-coded data. He thought those patterns were best represented and explained by models and theories of physics, particularly the Newtonian (Lagrangean) potential model (Rich, 1980).[1] That is, his spatial analysis was based on social physics. To that end, Warntz made literally billions of calculations during the course of his career. At first, a small proportion were by hand using mechanical calculators and taking months of his life. Later, and by far the bulk, were undertaken by a computer first in only a few hours, which diminished to minutes, then to seconds. Those calculations were the bases of countless contour surfaces of population potential that Warntz made for the United States. On one occasion, Warntz even made a 3-D version on his dining room table, constructed from plywood, plaster of Paris and many, many nails, each one standing at exactly in the right spot and at the right population potential height.[2]

In telling historical stories of social physics, and particularly of Warntz's work within that tradition, we are not claiming that social physics is the equivalent to what now passes as Big Data, or even that it is its embryonic version. It is more complicated. The work that Warntz did, as well as other social physicists on whom he drew, contained some of the elements that were later incorporated into Big Data via advancements in spatial analysis, providing one of its proving grounds, helping to shape its formulations and practices. Our larger point is that Big Data and its use in geography didn't start *de novo*, from scratch, after the rubble of the past was cleared away. It was erected *on* that past, which consequently must be examined.

Historical foundations of social physics

In an early study, Sorokin (1928: 3) argued that the intellectual foundation of social physics was the Ancient Greek philosophical idea of monistic materialism in which "psychical and social phenomena [were] mere variations of natural phenomena." As a result, regularities and laws found to hold for natural phenomena applied equally as forcefully to social phenomena. Natural and social worlds were isomorphic: one world, one set of principles.

Monistic materialism, thought Sorokin (1928), became especially important from the 17th century with the beginning of the scientific revolution in Western Europe. That revolution was defined in part by the use of mathematics. Galileo famously said in the early 17th century that the Book of Nature was

"written in the language of mathematics," and "without [it] one wander[ed] about in a dark labyrinth" (quoted in Drake, 1957: 238). If society was isomorphic to nature, it therefore followed that the social should be written in mathematics too. That was what social physics did. It strove to convert all forms of society into numbers, using mathematics to shed light on its own dim labyrinthine corners. In some cases, leading scientists also became leading social physicists. For example, the English astronomer Edmond Halley (1656–1742) turned his astrophysical skills in the early 18th century to the life insurance business, using his mathematical expertise to invent actuarial science by setting out statistical regularities from a large-scale analysis of mortality data. In other cases, concern was with only the social. John Graunt (1620–1674) sifted through statistical data on disease, laying down an intellectual basis for epidemiology. And Graunt's contemporary, Sir William Petty (1623–1687), was the first to compile national statistics of income and wealth, providing a foundation for the "dismal science." According to Stewart (1950: 243), each of these three works of large-scale numerical data compilation and analysis represented early examples of "social physics."

Not that "social physics" yet existed as a name. When it finally arrived, it was in French: "*physique sociale*." The Belgian astronomer, statistician, and social tabulator Adolphe Quetelet coined the term in 1835 in his book *Sur l'Homme*.[3] Specifically, Quetelet assembled a large data set of Parisian crime figures, subjecting it to a demanding social statistical analysis, making scientific predictions of "terrifying exactitude" (quoted in Hacking, 1990: 105). He was able to do so because, following the monistic line, he believed social life was governed by statistical laws revealed by applying the methods of social physics.

Quetelet in his analysis drew especially on "the error curve" (the Gaussian normal distribution) to understand social regularities. Curiously, though, he eschewed his own home discipline of astronomy to provide other explanatory theories. But they came from that subject in 1858, although not from a trained astronomer, but oddly from a retired autodidact and Philadelphia publisher, Henry Carey. Carey unabashedly believed in monism, writing that "the laws which govern matter in all its forms, whether that of coal, clay, iron, pebble stones, trees, oxen, horses, or man are the same" (quoted in Sorokin, 1928: 13). But he went further. His three-volume *Principles of Social Science* applied Newtonian gravity formulations to terrestrial human movement. Carey wrote that under "the great law of molecular gravitation . . . man tends of necessity to gravitate toward his fellow man [sic] . . . . Gravitation is here, as everywhere else in the material world, in the direct ratio of the mass

and in the inverse one of the distance" (quoted in Tocalis, 1978: 68).

But all of this was just a protracted warm-up for what was to come. Beginning from the early 1940s and continuing to the present, social physics burgeoned, drawing in research money, establishing institutional bases, fostering specialized workshops and conferences, and launching its own journals as well as thousands of academic papers. The current form of the field, also called sociophysics or statistical physics, is dominated by physicists (see, for example, Battencourt et al., 2007; Gabaix and Ioannides, 2003; Stauffer, 2004, 2013). They call what they do "interdisciplinary": applying principles and methods of physics to subject matter traditionally investigated by other disciplines such as economics, psychology, sociology, political science, and geography (Stauffer, 2004, 2013). But the earlier mid-20th-century incarnation of social physics, and this paper's primary concern, was interdisciplinary in a different sense. It was constituted by researchers from many different disciplines and which, from the early 1950s, included geographers, producing an explicit spatial analysis.[4]

Mid-century social physics

Only one of the two founders of the mid-20th-century version of social physics was a physicist, John Stewart, Professor of Astrophysics at Princeton. The other was a linguist, George Zipf, who taught in Harvard's German Department. From 1939, he was given a one-off position of University Lecturer that allowed him free rein to teach in the College any subject he wanted: he chose social physics.

George K. Zipf

Zipf was fascinated, maybe obsessed, with numerical patterns. He was originally appointed at Harvard's Department of German. Strangely, though, his research was not about that language but Chinese. It involved taking 20 extracts of Chinese writing, each with at least a thousand syllables, and generating in total 20,000 separate individual syllables (Rousseau, 2002: 12). Carrying out a large-scale numerical analysis, making use of student "human computers" to undertake the calculations, Zipf found that the frequency distribution of syllables conformed to the power law[5]

$$P_n \sim 1/n^a \qquad (1)$$

where P_n is the frequency of any given Chinese syllable ranked n (all syllables are ranked in order of their frequency of occurrence, with the most frequent ranked 1,

the second most frequent 2, and so on down to rank *n*); and exponent *a* is almost equal to 1.

The power law finding perfectly catered to Zipf's obsession. It implied a set numerical pattern in which the second most frequent Chinese syllable occurred approximately one-half as often as the first, the third syllable one-third as often, and so on down the line. Moreover, Zipf found that the same power law distribution held in other languages too, and not only for syllables but also for words and phrases (Zipf, 1932, 1968). In each case, the frequency of use of an element of language was inversely proportional to the rank of its frequency. In English, for example, Zipf found by examining texts as different as the King James Bible and James Joyce's *Ulysses* that "the" is the most frequently used word in the language, which is twice as common as the second most frequently used word, "of," and three times as common as the third most frequently used word, "and."

After Zipf was appointed as University Lecturer, he took his original finding and in true social physics fashion applied it to other social phenomena. In this he was guided by his faith in monism. As he put it in a 1942 paper, "the same natural law governs the structure and behavior of our planet, of life on our planet, and indeed, of even the smallest minutiae of living activity including the most subtle emotional and intellectual elaborations of the human mind" (Zipf, 1942: 48).

His immediate work was on city-size distributions, and published as *National Unity and Disunity* (Zipf, 1941). Resting on an enormous amount of data collection and analysis, making use of even more human computers than for his studies of language, Zipf found his power law formulation applied to city-size distributions not only across different countries but also over time as well. Further, when plotted on log paper the power law distribution of city-size distributions took on a striking linear form, making the relationship even more beguiling (Figure 1). It established that within a given urban system the second largest city was half the size of the largest city, the third largest city
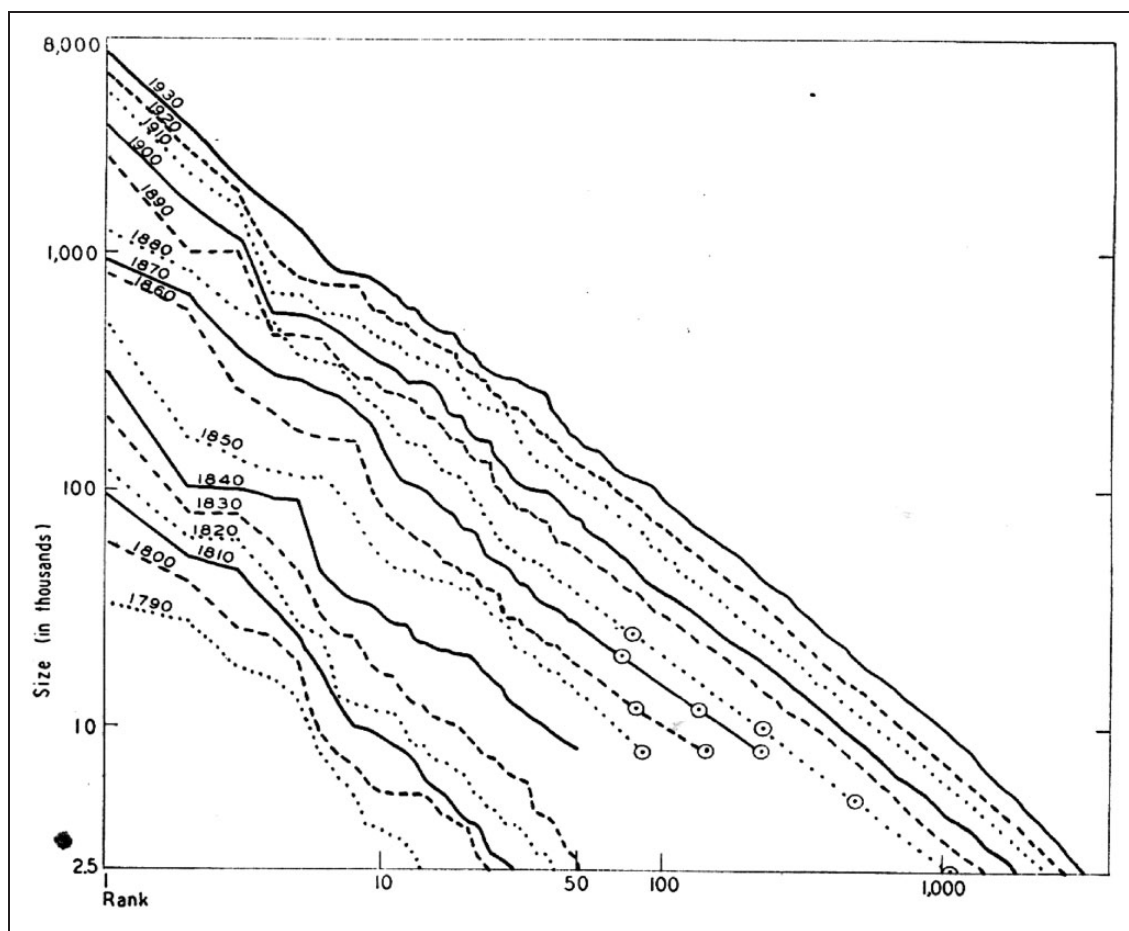

Figure 1. USA 1790–1930. Communities of 2500 or more inhabitants ranked in decreasing order of population size. Reproduced with permission from Zipf, 1949: 420.

a third of the size of the largest city and so on. Under such conditions the total population of the entire urban system was given by

$$\sum_{n=1}^{N} P_n = P_1 \cdot \sum_{n=1}^{N} 1/n \qquad (2)$$

where P_n is the population of the nth ranked city ($n = 1 \ldots N$); and P_1 is the population of the largest city.

Zipf called the right-hand side of equation (2) the rank-size rule, which was later eponymously dubbed Zipf's Law. For any given ranked city n, its population was equal to

$$P_n = P_1 \cdot 1/n \qquad (3)$$

Before his premature death from cancer in 1950 at age 48, Zipf, through prodigious feats of data collection and calculation, tried to show in his magnum opus, *Human Behaviour and the Principle of Least Effort* (1949), that the power law he had found for word frequency and city-size distributions applied to all phenomena.[6] He seemingly had no problem in showing that it applied to countable things like populations of cities, words used in English language publications from Beowulf to T.S. Eliot, and the "length of intervals between repetition" in Mozart's Bassoon Concerto in B^b Major (respectively Zipf, 1949: 420, 123, 337). But he also suggested that it applied even to more tricky phenomena that one might think would resist enumeration, such as "the death wish," multiple personalities, and "erotic substitute action" (it is best not to ask) (respectively Zipf, 1949: 240–241, 264–266, 269–270). In each case, the power law holds because, as he argued in his 1949 volume, that distribution is the manifestation of the ultimate monist principle, *lex parsionae*, the principle of least effort (Zipf, 1949: ch. 1). Zipf (1949: 1) contended that all parts of nature, and consequently all parts of society, beat the most efficient path to any given end, "minimiz[ing] the *total work* that [is] . . . expended in solving *both* . . . immediate problems, and probable future problems [original emphases]." That resolution always takes the form of the power law distribution; it is the statistical manifestation of the Principle of Least Effort.

Zipf's ambition was overweening. He thought social physics "could be applied to everything . . . down to and including man's innermost dreams. The soul. . . offers a perfectly legitimate problem to science" (Zipf, 1942: 62). Those who resisted that end he termed "verbalists," who he thought should be excoriated, if not suffer something even worse.[7] Moreover, he believed he was on a modern-day crusade: "My dear John," he wrote to Stewart on 22 May 1949, "the intellectual

battle of the second half of the twentieth century is on; you and I are both in it up to our necks with no earthly possibility of retreat (and there is no desire). I propose to keep the canons barking."[8] He did until the day he died. But his death was a half-century too early. He would have been in computational heaven had he lived to experience a world of Big Data. A world in which vast amounts of social data were available to be sifted, arranged, and analyzed, made to show the power of the power law distribution and *lex parsimonae*, and with neither a verbalist nor a human computer in sight. Our suggestion is that Zipf's work in part prepared social science, and geography in particular, for that very world.

John Q. Stewart

Even more important for the social physics movement than Zipf, who was a loner, was John Quincy Stewart. His work directly linked to geography, and he also had on his side, at least for a time, some heavy-weight US East Coast institutions and intellectuals including the Institute of Advanced Study at Princeton directed by Robert Oppenheimer, and Warren Weaver at the Rockefeller Foundation that funded his social physics project for four years (1949–1953). It was Stewart from his Princeton base who energetically and systematically constructed, maintained, and promoted social physics as a new stand-alone discipline. Continually on the look-out for new resources and allies, never missing an opportunity for endorsement (for both his project and for himself), he strove to turn social physics into a new social science.

Stewart was a Princeton man through and through. All his degrees were from there, and two years after his PhD he was back as a Princeton professor. Initially interested in matters extraterrestrial by mid-career his research focus became increasingly earthly. It centered on weather and navigation and, important for the later story, led to the publication of a widely circulated primer used in the US Army Air Force, *Coasts, Waves, and Weather for Navigators* (Stewart, 1945).

Part of that same terrestrial turn was his interest in models of population potential; that is, applying ideas of gravitational potential to the geographical distribution of population. There are different ways to interpret population potential, but perhaps the most common is as a measure of the influence of a given population at one location (j) on another location (i) (Rich, 1980: 5)

$$v_i = P_j/D_{ij} \qquad (4)$$

where v_i is the population potential at location i, P_j is the size of the population at j, and D_{ij} is the distance between locations i and j.

Equation (4) says that the population influence of location j on location i is proportional to the size of the population at j doing the influencing, and inversely proportional to the distance between the two locations. True to his roots, Stewart first applied this potential equation to the geographical distribution of the "Princeton family," that is, the 20,438 members listed in the 1940 edition of Princeton University's Alumni Directory.[9] Devoting his 1939 "Christmas holiday to statistical study," and drawing on the "vigorous assistance of Philip Wilkie ... [for] some of the computations," Stewart first measured the population potential of every state in relation to New Jersey, Princeton's home state. From that he worked out the expected number of Princeton alumni for every state, and compared that number to the actual count from the Alumni Directory. Stewart found that expected and actual numbers of alumni "more or less" lined up. It showed him that potential models applied not only to objects in the heavens but also to those down on earth within the social field, thus fulfilling the promise of monism; that is, a social physics one-size-fits-all theory.

Stewart subsequently expanded his inquiries, calculating total population potential for the US as a whole. For any given location i, total population potential is equal to the sum of all the individual potentials created by populations found at every location j ($j = 1, 2, \ldots N$)

$$V_i = \sum_{j=1}^{N} P_j / D_{ij} \qquad (5)$$

where V_i is the total population potential for location i.

With total population potentials known for all locations i, an equipotential (isopleth) map can be constructed. Each line on the map connects all places with the same population potential. Figure 2 shows population potential isolines calculated by Stewart for the United States for 1940. Population potentials are especially high in the North East (NE), indicating the considerable influence of the population on any given location within that region, but it falls off markedly for locations in the West until it rises again on the Pacific Coast.

Drawing equipotential maps took an immense amount of computational effort. That's why Stewart had to give up his Christmas holidays in 1939, shackled to log tables, slide rule, and Mr. Philip Wilkie. But by the late 1940s help was on its way in the form of electrical computation and cartographic production.
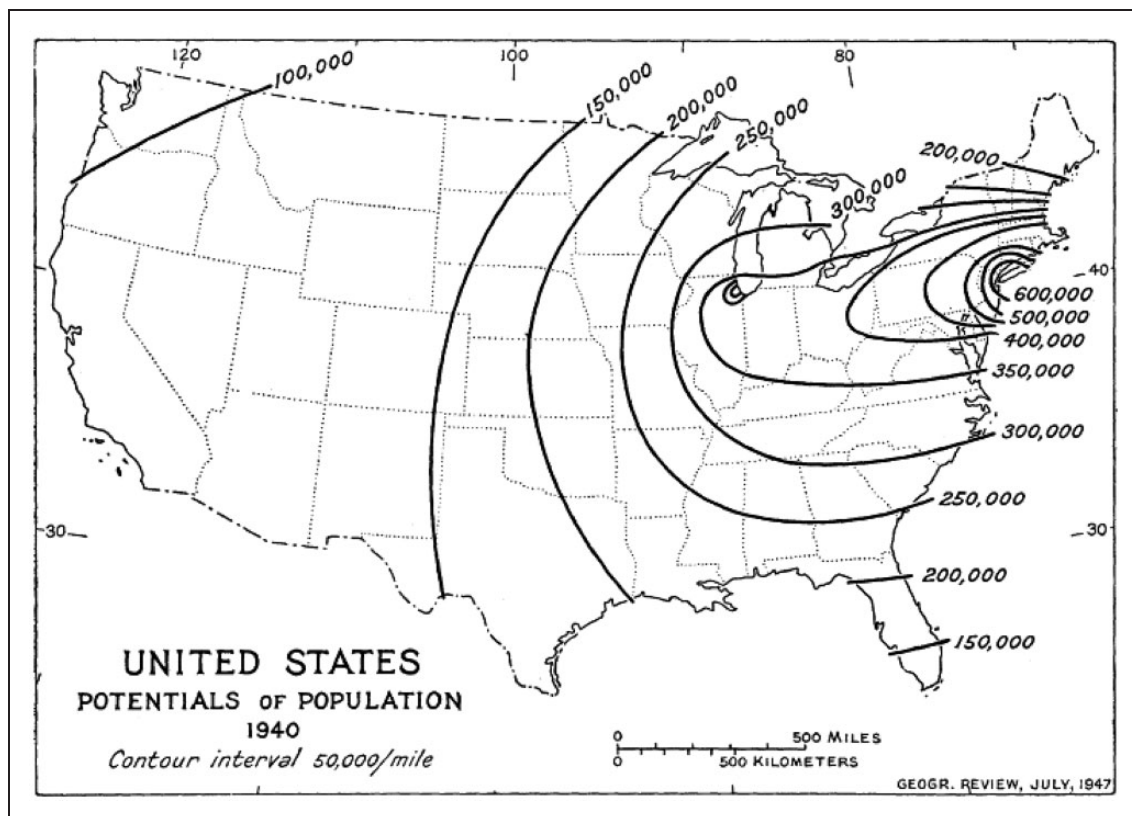

Figure 2. United States: potentials of population, 1940. Reproduced with permission from Stewart, 1947: 436.

Stewart enlisted into his project Thomas B. Bissett, a Princeton senior undergraduate electrical engineer, and inventor of a nationally award-winning "computing device." Taking second prize at the American Institute of Electrical Engineers' meeting at Pittsburgh in April, 1949, Bissett's machine "permit[ed] the rapid drawing of contours of equipotential." As Stewart described it, Bissett's was an analog machine that "applies Ohm's law . . . to the computation of potential of population (number of people divided by distance). Threads attached to the scriber controls a set of resistances as it moves."[10] Admittedly, Bissett's machine wasn't of the same caliber as other contemporary computers such as Jay Forrester's at MIT, the "WHIRLWIND" (Hughes, 1998), or John von Neumann's at the Institute of Advanced Studies, Princeton, the "JOHNNIAC" (Dyson, 2012), but it was certainly more reliable and quicker than Stewart's former standby of human computers.

By the summer of 1949, Stewart thought it was time for both he and the world to lose the quotation marks that usually accompanied the term "social physics." He approached a key post-war science manager and well-known scientist in his own right, Warren Weaver, head of the Division of Natural Sciences of the Rockefeller Foundation, for funding to develop social physics as a proper social science. Shannon and Weaver (1949) had published that same year *The Mathematical Theory of Communication*, which became a central text of the early post-war period, linking science, mathematics, computing machines and society and according to Edwards (1996: 203), enabling a "complete statistical analysis of human social activity." Weaver wrote back to Stewart, "I am completely sympathetic with your approach." And best of all, he offered him a grant of $15,000.[11]

Stewart used the money to mount seven conferences, as well as to help form in 1951 a Social Physics Committee of 40 strong, associated with the American Association for the Advancement of Science (AAAS), which produced its own bevy of special sessions. Stewart also increasingly adopted the language of "isomorphism" to describe the relationship between natural and social sciences. In a paper he gave in Boston at an AAAS meeting in December 1953 he suggested that "isomorphism means reliable item-to-item correspondence between two or more fields which superficially may seem altogether dissimilar. The fields are exhibited as more or less identical in their patterns of concepts."[12] Of course, this was another version of monism, and justification both for more computation, data compilation, and mathematization within the social sciences, and further exploration of points at which concepts from physics corresponded to items and relationships within the social world. In the December paper, Stewart was emboldened to assert isomorphisms between thermal, electromagnetic, and chemical energy and, respectively, human meaning, feeling, and authority, as well as an isomorphism between mechanical work (kinetic, elastic, and gravitational energy) and decision-making.[13]

But it was all a bit too much for Weaver, who wrote to him in 1953:

> To search for isomorphisms between social phenomena and physical phenomena is indeed an interesting idea. The real question, however, is whether or not it is a rewarding idea. It is interesting to suppose that there may be entities, social values, which play in social experience the same roles played by different forms of physical energy . . . But it is hard for me to sense how one can usefully assign quantitative measures to any significantly wide range of "values" in the social field. And when you link together such things as meaning, feeling, authority, and decision-making, this sounds to me like a very heterogeneous mixture.[14]

Stewart bristled at Weaver's use of "interesting." It "is more than merely 'interesting,'" Stewart wrote back. He was also "surprised" that it "sounds to you 'like a very heterogeneous mixture.'"[15] The argument here was around whether all social variables could, to use Weaver's terms, be assigned "quantitative measures." Weaver was sceptical, but Stewart bullish, in part because of his faith in monism. Clearly, this same issue continues into present discussions of Big Data, which seem to have sided with Stewart rather than Weaver.[16] But on that occasion, Weaver held sway because he controlled the purse strings. He told Stewart that his Rockefeller funding would not be renewed. Stewart needed more allies. And that's what he got later that same year, 1953, when William Warntz contacted him.

Macrogeography and William Warntz

The first communication between Stewart and Warntz was about the Bissett. Warntz wrote to Stewart on 1 October 1953, to ask whether the "ingenious electrical computer designed and constructed by Thomas B. Bissett . . . could be made available to him?"[17] Warntz was writing his PhD thesis at the University of Pennsylvania using Stewart's population potential model, trying to calculate equipotential lines. He needed that machine. But unfortunately it was not on hand. "No Bissett available here" replied Stewart, adding with a mathematician's zeal, "longhand computation is not so bad."[18] Possibly not, though. Warntz's obituary writer, Donald Janelle (1997: 724), reckoned that in completing his doctoral dissertation alone Warntz undertook "thousands of calculations to

derive population potential fields, all with a mechanical adding machine."

Although their first correspondence wasn't until 1953, Warntz had been a secret admirer of Stewart for some time. Warntz began his undergraduate degree at Penn in economics in 1940. But the war interrupted his studies. In 1943, he enlisted in the US Army Air Force, and trained as a navigator. In active service from 1944, awarded several military medals, Warntz served in a combat heavy bombardment group based in East Anglia. Following a crash landing as Warntz returned from a mission, he was sent to Cambridge to recuperate. It was there, presciently given his subsequent involvement with Stewart and social physics, that he read Isaac Newton's papers at the university library, the site of the largest single deposit of such papers in the world. After the war ended, Warntz remained in the Army Air Force and was posted to Gander at the Newfoundland Base Command for sea search and rescue missions. In the base's small library he found J.Q. Stewart's *Coasts, Waves and Weather for Navigators* (1945), which was to change his life. The book was primarily a how-to manual for navigators. But in an appendix, and quite out of character with the rest of the book (Warntz, 1984, 1, called it "an exotic chapter"; 1984: 1), Stewart "could not resist the temptation" to write about his new interest in population potential, likening equipotential lines to isobars within a cyclonic system. Warntz later said that was his "Ah-ha" moment (Dow, 1973: 56), when "social science and the things I learned about meteorology and navigation" came together.[19]

Once Warntz demobbed in 1948, he quickly completed his Bachelor's degree in economics at Penn, did also a Master's degree there, and in 1951 entered its doctoral economics program, specializing in economic geography. He was supervised by Lester Klimm, an economist with an unlikely interest in location theory. His dissertation, later published as *Toward a Geography of Price* (Warntz, 1959a), deployed the potential model to present a geographical analysis of price, empirically illustrated using commodity statistics taken from the US Department of Agriculture. The thesis involved the construction of both an aggregate demand curve derived from population potential estimates weighted by income, and an aggregate supply curve obtained from potential estimates of spatially varying production levels of different agricultural commodities. The intersection of the two curves determined the geography of price.

After the initial contact about the Bissett in October 1953, Stewart invited Warntz to come to the Princeton Observatory. They met on 3 December 1953, and the rest is history, with Warntz becoming Stewart's staunchest ally. The timing could not have been better given

Weaver's rejection of Stewart's project earlier that same year. It meant social physics got a second chance with a new home, in a discipline with a well-established provenance that came with a formidable institutional structure that included both the American Geographical Society (AGS) (where from 1956 Warntz was a Research Associate running his own "special program" in social physics) and the Association of American Geographers (AAG), which brought a large prospective audience, so far untouched by social physics but potentially ready to be converted.

But for this to happen in geography, things needed to be shaken up. In an address to the Geography session at the annual meeting of the National Council for Social Studies in Pittsburgh, in November 1957, Warntz said:

> In our researches in geography and our contributions to the journals of learned societies, I suggest the time has come for some unruliness, for discomfiting the feeble-minded, for treading upon the weak, and for general impatience. The time has come to raise a rumpus in geography.[20]

Warntz argued in that address that geography was getting no respect either from "the Ivies" – Ivy League universities like Harvard, which had shut down its geography program nine years before – or from geographers themselves who since the interwar period had pursued a parochial and prosaic version of their discipline, concerned only with "the immediate, the obvious, the microscopic, and the demonstrably utilitarian." For Warntz, the discipline could be so much more. But to do that, geographers needed to think big, to practice "macrogeography." He thought macrogeography would attract "the Ivies" and bring back to the discipline "originality," "erudition," and "brilliant scholarship."[21]

Predictably Warntz's conception of macrogeography was uncompromisingly monistic. Warntz was fond of repeating his own choice phrases, and there was no phrase he repeated more often than "social science... and physical science are but mutually related isomorphic examples of one generalized logic" (Warntz, 1957: 422, fn. 1). That was the strength of macrogeography; it embodied the "one generalized logic." That logic was of a single discipline: physics. It meant geography becoming like physics. And if a discipline failed to do so, as geography patently had, then it would suffer; it *had* suffered. That's why a "rumpus" was necessary.

Warntz thought geography had gone wrong when it turned to what he called microgeography, which was "preoccupied with the unique, the exceptional... and often the obvious" (Warntz, 1959b: 447). It relied on

Zipf's dreaded verbalism, "subjective description," such as "a feeling for the area," "the personality of the region," and "man-land relationship" (Warntz, 1957, 420). Under this regime of verbalism, Warntz said microgeography "has been tried, found wanting, and is justifiably being discarded."[22]

Now was the time to make amends, to do macrogeography – a geography unified with the "one generalized logic."[23] Although Warntz coined the specialized term macrogeography for his approach, it was nothing other than social physics. The name macrogeography served only a rhetorical function of making the project palatable to geographers, adding nothing substantively that wasn't already contained within Stewart's original formulation. But the name did resonate with, and in part was picked up by, other practitioners of spatial analysis who also emerged in the mid-to-late 1950s both inside and outside geography. Walter Isard's Regional Science Department was founded at Penn the same year Warntz finished his PhD therein 1955. Social physics was incorporated at least into Isard et al.'s (1960) tome on regional science methods, and Warntz maintained an adjunct position in regional science at Penn until 1966. Likewise there were various linkages between Warntz and the work of proponents of geography's "quantitative revolution" that occurred around the same time. Particularly important were the so-called "Garrison Raiders," a group of graduate students at the Department of Geography, University of Washington, Seattle, who worked with a young faculty member, Bill Garrison, who shared Warntz's dissatisfaction with "microgeography" and who sought an alternative in some form of spatial analysis.

Core to Warntz's spatial version of social physics was the computational task of confronting big data. In this he was fortunate, given that he and his macrogeography came of age during a period marked both by the rapidly improving power of machine-based computation as well as by increasing access to those machines. There would be no more cap-in-hand begging for the Bissett. In a paper, in the *Geographical Review*, Warntz (1964: 171–172) directly linked the development of macrogeography to computer development. The issue turned on the number of so-called "control points" used to calculate a given population potential map. A control point was a designated map location associated with a particular population level. Knowing those locations and population numbers was necessary before any potential calculations could be made and equipotential lines drawn. But the more control points, and clear from equation (5), the more gruelling the resulting calculations. In Stewart's first 1940 set of population potential calculations there were just 24 control points. Calculating for even that relatively small number by hand, even with the

assistance of the energetic Mr. Wilkie, still took Stewart all the Christmas holidays. Later Stewart used Thomas Bissett's machine, but it could only handle 10 control points at a time (Warntz, 1964: 177). The breakthrough was the arrival of the digital computer. In 1958, Warntz and Stewart used "an early IBM electronic computer...[to] compute...a map with 115 control points" (Wantz, 1964: 171). And by 1963, using Princeton's IBM 7090 computer, Warntz was calculating the US population potential using 3105 control points (one for every county in the conterminous United States). Warntz (1964: 177–178) reports that in this case:

> Inputs came from 3,105 punched cards...and [each card] contained for each county...the 1960 population, area, and geographical co-ordinates of the county's population center...[A]bout ten billion arithmetic operations were required to produce th[e] map, and three other, different maps computed at the same time. About twenty-three thousand words of storage (thirty-six bit words) were required, and the programming language used was Fortan...Running time for the computer was a few seconds short of two hours.

Warntz also used the same computer for his 200 hour-long, spare-time project of constructing a three-dimensional physical model, 2 by 3¼ feet, of the 1960 US population potential map. It was for an AGS exhibit, "New York: An International City," at the 1964–1965 New York World Fair (Janelle, 1997: 724). Warntz had hammered into the base map 3105 individual nails, the exact height of each calculated by the IBM 7090.

Warntz's work impressed Howard Fisher, who in 1965 had founded the Laboratory for Computer Graphics within the Harvard Graduate School of Design (GSD), with the intellectual support of Brian Berry of the University of Chicago and Waldo Tobler of the University of Michigan (both geographers and members of Bill Garrison's "Raiders"). Fisher had organized a series of luncheon presentations on the state-of-the-art in the spatial representation of quantitative information, to include Berry and Warntz.[24] In 1966, Warntz was made Professor of Theoretical Geography and Regional Planning (although without tenure) within the lab (the first professor of geography on Harvard's campus since Derwent Wittlesey's death 10 years prior in 1956). Fisher and Warntz's work within the lab was two-fold, according to Warntz: first, the "continued investigation into the uses of graphical analysis, and computer graphics in particular," of which Fisher no doubt led the effort, and, second, "pure research in the framework of general systems theory and spatial patterns,"[25] much of which was also

lumped by Warntz as theoretical geography and published within the *Harvard Papers in Theoretical Geography* (edited by Warntz, 1966–1971).

As Fisher was due to retire in 1970, at the exhaustion of the Ford Foundation grant that supported the lab, Warntz became Director in 1968 and the name of the lab became the Laboratory for Computer Graphics and Spatial Analysis. In addition to his duties within the Graduate School of Design, Warntz also led freshmen seminars within Harvard and was even asked to submit a proposal justifying the return of geography to the university (Janelle, 2000: 105). Seemingly, macrogeography really was what the Ivies wanted, with Warntz's work representing qualities of "originality," "erudition," and "brilliant scholarship." In the end it didn't work out, with geography never returning to Harvard after all, and Warntz resigning his position three years after he became Director (see Chrisman, 2006, for an excellent history of the Harvard lab and Warntz's role within it). What Warntz did achieve, however, was to prepare the ground of geography for Big Data, providing models of spatial analysis, stressing the importance of powerful computing machines, and championing the belief that sooner or later the entire social field could be turned into numbers.

Conclusion

> The strategic positions of graphics and geometry become even more clearly defined, while concomitantly the distinction between cartography and spatial mathematics decreases. (Warntz, 1977 [1967]: 5)

> The engine that drives social physics is big data: the newly ubiquitous digital data that is becoming available about all aspects of human life. By using these data to build a predictive, computational theory of human behavior, we can hope to engineer better social systems. (MIT Media Lab, 2014: n.p.)

Warntz's belief that numeracy held the key to spatial problem-solving was yoked to his curiosity about spatial representation as a means both to test hypotheses and to realize solutions graphically. That "we now look upon maps not only as stores for spatially ordered information" but also as part of a three-segmented theoretical geography of description, classification, and prediction, complemented the efforts of the lab under Fisher, an architect interested in assessing how computer graphics would alter the conventions of thematic cartography. [26] And this predilection can be felt as the impulse of Big Data, as can be seen in the materials for the new social physics at the MIT Media Lab, "to build a predictive, computational theory of human behavior." While at

Harvard, Warntz worked alongside many of the leading figures of what would become Geographic Information Science, including Jack Dangermond (founder of Esri, the company responsible for ArcGIS), and when he resigned his position at the GSD in 1971 he was appointed at the University of Western Ontario, where Bill Bunge (another Garrison Raider and author of *Theoretical Geography*) was serving as visiting professor and where Mike Goodchild was beginning his academic career in spatial analysis (and would later occupy the Dangermond Chair of Geography at UC-Santa Barbara).[27]

Big Data continues to vex GIScience, as can be seen in the precipitous rise of volunteered geographic information (VGI).[28] A term coined by Goodchild (2007: 220), "the most important value of VGI may lie in what it can tell about local activities in various geographic locations that go unnoticed by the world's media, and about life at a local level." The spatial analysis of these data inherits the affect, if not many of the techniques, associated with social physics. As Anselin and Rey (2012: 2211) write, the techniques associated with current spatial analysis and their subfield of spatial econometrics "has roots in regional economics and quantitative geography." Interestingly, this form of macrogeography – to understand the spatial patterning of Big Data – does not preclude attention to the micro (in fact, it presupposes recording the rich details of private life). Likely, Warntz would be pleased. While, perhaps, there are other epistemologies that enroll Big Data beyond the monism of data-driven methods (for instance, Crampton et al., 2013), our curiosity is increasingly piqued by the continuities of macrogeography with some contemporary GIScience work, made more conspicuous, we suggest, by the rise of Big Data.

Our suggestion is that we've been here before. That the call to quantify and to represent social-spatial activity enacts a specific visibility that potentially renders invisible other phenomena. Not everything can be quantified, or only partially quantified. What's not drops out of view. That was precisely the issue when critics in human geography took aim at Warntz's social physics version of the discipline during the 1970s. So-called humanistic geographers argued that what became invisible in Warntz's social physics was any real, sentient human being, who could never be described by numbers alone (Ley and Samuels, 1978). Radical geographers such as Harvey (1972) thought what dropped out were social relationships, power inequalities, and class conflict – for them, vital to understanding any kind of geography, large or small.

While these criticisms led to new forms of geography emerging, the social physics kind continued, indeed flourished, finding new applications and academic homes, taking on new manifestations, one of which

we've argued was Big Data. That's the reason we need to trace the connective tissues between Big Data and the mid-20th century. Not to do so would be a mistake. By showing that Big Data is historical, we show the assumptions that were built into it, as well as the contestations around them. Big Data becomes no longer a black box, self-contained, sealed and impregnable, but is opened up, available for verbalist discussion and contestation.

In 1967, Warntz provided a summary of the lab's position with respect to the mathematization of geography:

> Description of surfaces and paths including geodesics, spatial probabilities, central-place hierarchies in the broadest sense are the types that seem destined to replace the current arrangement characterized by specialties classified in terms of non-spatial properties and including such things as population geography, linguistic geography, cultural geography, economic geography, and so on. We agree.[29]

What Warntz is trying to do here is monism: to say that there is only one version of geography, the social physics kind, and all other kinds of geography are mistakes that must be abandoned. That same monistic urge, we've suggested, is also found in Big Data. We disagree. But as verbalists, we would.

Acknowledgements

Declaration of conflicting interests

Funding

Notes

1. Isaac Newton (1642–1727) did not recognize gravitational potential, but the Italian mathematician and astronomer Joseph-Louis Lagrange (1736–1813) did some 40 years after Newton's death. He found that when several planets were clustered gravitational attraction could be represented by a new coefficient, potential.
2. A photo of Warntz and his 3D map is found in Don Janelle's 2012 power point presentation, "William Warntz and the legacy of spatial thinking at Harvard University." Available at: https://cga-download.hmdc.harvard.edu/publish_web/Video/2012_09_05_Don_Janelle_slides.pdf (accessed 8 January 2014).
3. 'Physique sociale" was also chosen by August Comte (1796–1857) in 1838 to describe his positivist project. But the "statistical study of people... [was] anathema to Comte" (Hacking, 1990: 39), and to ensure there was no confusion with Quetelet, who he felt was a rival, Comte plumped instead for what he thought was a new term, "sociologie." Ironically, it too had already been coined in 1790 by Emmanuel-Joseph Sieyès.
4. In a 1954 paper presentation discussing the achievements of social physics, Stewart lists the full interdisciplinary glory of the field: "physics, chemistry, astronomy, mathematics, mathematical logic, geology, biological science, history of science, engineering and invention, industrial engineering, geography, sociology, psychology, economics, history, law and jurisprudence, public relations, theology, philosophy, and education." "Social physics and engineering," by John Q. Stewart, 30 January 1954, pp. 2–3. Box 58, Miscellaneous writing, John Q. Stewart Papers, Rare Books and Special Collections, Princeton University.
5. Richard Wallace says Zipf had a "room full of computers to count words.... Each 'computer' would arrive at work and begin tallying the words and phrases directed by Zipf." The computers, of course, were humans, Harvard undergraduates dragooned into service (available at: alicebot.org/articles/wallace/zipf.html; accessed 8 January 2014). In addition, Zipf's son Robert remembers also being collared into helping his father complete his work, "which required considerable hand calculations in that pre-computer time" (Prün and Zipf 2002: 3).
6. Zipf's colleague at Harvard, Clyde Kluckhohn (1950: 269–270) in his review of Human Behavior and the Principle of Least Effort, commented on Zipf's "staggering range of data.... To name only a few: number of different news items in various American papers; number of obituaries in the New York Times; charge accounts of Jordan March Co., Boston; bus, rail, and air passenger travel; telegrams exchanged between cities; density of traffic in Milwaukee; claims against insurance companies; marriage licenses related to residences of applicants."
7. Zipf wrote to Stewart in November 1949: "I gather that you too would like to hear lead pipes crunch on verbalistic skills. We have indeed suffered much from the verbalists. But being a peace-loving non-aggressive soul, my only advice is, 'Don't shoot the verbalist with blank cartridges'." Zipf to Stewart, 12 November 1949, Box 38, Zipf, George K., John Q. Stewart Papers, Rare Books and Special Collections, Princeton University.
8. Zipf to Stewart, 22 May 1949, Box 38, Zipf, George K., John Q. Stewart Papers, Rare Books and Special Collections, Princeton University.
9. "The gravity of the Princeton family," by John Q. Stewart, Princeton Alumni Weekly, 9 February 1940, pp. 409–410. Box 3, folder 9, John Q. Stewart Papers, Rare Books and Special Collections, Princeton University.
10. Memorandum to Research Corporation: Concerning the project, "The transfer of the social field of the methods

and some of the principles of physical science," *Progress in Social Physics*, September 1948 to May 1949. Box 2, Progress Report, quotations from p. 5 to p. 8. John Q. Stewart Papers, Rare Books and Special Collections, Princeton University.

11. Weaver to Stewart, 24 August 1949, Box 36, Weaver, W., John Q. Stewart Papers, Rare Books and Special Collections, Princeton University. The grant is reported in the Rockefeller Foundation Annual Report, 1950, pp. 175–176. Available at: http://www.rockefellerfoundation. org/about-us/annual-reports/1950–1959 (accessed 8 January 2014).

12. John Q. Stewart, "Remarks on the current state of social physics." Paper presented in Boston, 30 December 1953, American Association for the Advancement of Science conference, p. 2. Box 58, Miscellaneous writing, John Q. Stewart Papers, Rare Books and Special Collections, Princeton University.

13. Stewart (1950). "Remarks on the current state of social physics," p. 8.

14. Weaver to Stewart, 22 December 1953, Box 36, Weaver, W., John Q. Stewart Papers, Rare Books and Special Collections, Princeton University.

15. Stewart to Weaver, 3 January 1954, Box 36, Weaver, W., John Q. Stewart Papers, Rare Books and Special Collections, Princeton University.

16. The issue here is the relation of monism to Big Data. The most immediate invocation of monism by Big Data is its assumption that the social world can be mathematized in the same way as the natural world. Just as Galileo thought that the Book of Nature is written in the language of mathematics, there is a parallel belief within Big Data about the "Book of Society." Without the supposition that the social world can be fully made over as numbers, Big Data would have no purchase. We also suggest that monism is invoked by Big Data in a second form, at least implicitly. When Big Data deploys models of spatial analysis monism is presumed, because those models partly rest on a social physics that makes monism foundational.

17. Warntz to Stewart, 1 October 1953, Box 36, Warntz, W., John Q. Stewart Papers, Rare Books and Special Collections, Princeton University.

18. Stewart to Warntz, 14 November 1953, Box 36, Warntz, W., John Q. Stewart Papers, Rare Books and Special Collections, Princeton University.

19. William Warntz, "Social physics," *Princeton Alumni Weekly*, 18 September 1959: 8–13; both quotes p. 9, Box 58, Folder (B-St49-jq), John Q. Stewart Papers, Rare Books and Special Collections, Princeton University.

20. William Warntz, "The unity of knowledge, social science, and the role of geography," p. 1. An address presented at the geography session, National Council for the Social Studies 37th Annual Meeting, Friday, 29 November, Pittsburgh, Pennsylvania, Box 5, Folder 5, John Q. Stewart Papers, Rare Books and Special Collections, Princeton University.

21. Warntz (1950). "The unity of knowledge," respectively p. 3, p. 9, and p. 6.

22. Warntz (1950d). "The unity of knowledge," p. 6.

23. Mr. O.M. Miller, who worked with Warntz at the AGS, sought to differentiate macrogeography and microgeography by doggerel, and printed at the end of a joint paper by Stewart and Warntz (1958: 184) in the *Geographical Review*, "Macrogeography and social science":

> Geography,
> In small degree,
> Collects
> The facts that be
> On land and sea;
> Expects
> But little thought,
> Nor is it sought.
>
> Geography,
> In large degree,
> Is thus
> The gift to see
> Not thee and me
> But us,
> Summed up with grace
> In man and space.

24. "Announcing Five Seminars relating to Computer Graphics and Geographic Analysis," 24 October 1966, Box 13, Luncheon Reference, Howard T. Fisher Papers, Harvard University Archives.

25. William Warntz, "Introduction," March 1969, Box 12, Red Book Laboratory for Computer Graphics and Spatial Analysis, Folder 1, Howard T. Fisher Papers, Harvard University Archives.

26. Warntz (1977 [1967]: 3), full quote: "We now look upon maps not only as stores for spatially ordered information but also as a means for the graphical solution of certain spatial problems for which the mathematics provides to be intractable, and to produce necessary spatial transformations for hypothesis testing."

27. See Janelle (1997).

28. See recent special issues of *Environment & Planning A* by Wilson and Graham (2013) and *GeoJournal* by Leszczynski and Wilson (2013).

29. See Warntz (1977 [1967]: 5).

References

Anderson C (2008) The end of theory: The data deluge makes the scientific method obsolete. *Wired*, 23 June, 2008. Available at: http://www.wired.com/science/discoveries/magazine/16-07/pb_theory/ (accessed 8 January 2014).

Anselin L (2012) From SpaceStat to CyberGIS: Twenty years of spatial data analysis software. *International Regional Science Review* 35(2): 131–157.

Anselin L and Rey SJ (2012) Spatial econometrics in an age of CyberGIScience. *International Journal of Geographical Information Science* 26(12): 2211–2226.

Barnes TJ (2003) The place of locational analysis: A selective and interpretive history. *Progress in Human Geography* 27: 69–95.

Barnes TJ (2013) Big Data, little history. *Dialogues in Human Geography* 3: 297–302.

Battencourt LMA, Lobo J, Helbing D, et al. (2007) Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of the Sciences* 104(17): 7301–7306.

Chrisman N (2006) *Charting the Unknown: How Computer Mapping at Harvard Became GIS*. Redlands, CA: ESRI Press.

Crampton JW, Graham M, Poorthuis A, et al. (2013) Beyond the geotag: Situating 'big data' and leveraging the potential of the geoweb. *Cartography and Geographic Information Science* 40(2): 130–139.

Dow MW (1973) Interview with William Warntz, 15 April, Atlanta, Georgia. Available at Geographers on Film, Selected online GOF transcriptions, 56–60.

Drake S (1957) *Discoveries and Opinions of Galileo*. New York: Anchor Books, Doubleday.

Dyson G (2012) *Turing's Cathedral: The Origins of the Digital Universe*. New York: Pantheon.

Edwards PN (1996) *The Closed World: Computers and the Politics of Discourse in Cold War America, Inside Technology*. Cambridge, MA: MIT Press.

Faulkner W (1994 [1951]) *Requiem for a Nun*. New York: Vintage.

Gabaix X and Ioannides Y (2003) The evolution of city size distributions. Working paper, Department of Economics, Tufts University. Available at: http://ase.tufts.edu/econ/papers/200310.pdf (accessed 8 January 2014).

Goodchild MF (2004) The validity and usefulness of laws in geographic information science and geography. *Annals of the Association of American Geographers* 94(2): 300–303.

Goodchild MF (2007) Citizens as sensors: The world of volunteered geography. *GeoJournal* 69: 211–221.

Goodchild MF (2009) What problem? Spatial autocorrelation and geographic information science. *Geographical Analysis* 41: 411–417.

Hacking I (1990) *The Taming of Chance*. Cambridge: Cambridge University Press.

Harvey D (1972) Revolutionary and counter-revolutionary theory in geography and the problem of ghetto formation. *Antipode* 4(2): 1–13.

Hughes TP (1998) *Rescuing Prometheus*. New York: Pantheon.

Isard W, Bramhall D, Carrothers GP, et al. (1960) *Methods of Regional Analysis: An Introduction to Regional Science*. Cambridge, MA: MIT Press.

Jacobs A (2009) The pathologies of Big Data. *Communications of the ACM* 52(8): 36–44.

Janelle DG (1997) In memoriam: William Warntz, 1922–88. *Annals of the Association of American Geographers* 87(4): 723–731.

Janelle D (2000) William Warntz, 1922–1988. In: Armstrong PH and Martin GJ (eds) *Geographers: Biobibliographical Studies*. Vol. 19. London and New York: Mansell, pp. 102–118.

Kluckhohn C (1950) Review of human behavior and the principle of least effort. *American Anthropologist* 52: 268–270.

Leszczynski A and Wilson MW (2013) Theorizing the geoweb. *GeoJournal* 78(6): 915–919.

Ley D and Samuels M (eds) (1978) *Humanistic Geography*. Chicago, IL: Maroufa Press.

MIT Media Lab (2014) About Social Physics. Available at: http://socialphysics.media.mit.edu (accessed 13 January 2014).

Prün C and Zipf R (2002) Biographical notes on G. K. Zipf. *Glottometrics* 3: 1–11.

Rich DC (1980) *Potential models in human geography*. CATMOG 26, Study Group in Quantitative Methods, Institute of British Geographers. University of East Anglia, Norwich: Geo Abstracts.

Rousseau R (2002) George Kingsley Zipf: Life, ideas, his law and infometrics. *Glottometrics* 3: 11–18.

Shannon C and Weaver W (1949) *The Mathematical Theory of Communication*. Urbana and Chicago: University of Illinois Press.

Sorokin PA (1928) *Contemporary Sociological Theories Through the First Quarter of the Twentieth Century*. New York: Harper & Row.

Stauffer D (2004) Introduction to statistical physics outside physics. *Physica A* 336(1–2): 1–5.

Stauffer D (2013) A biased review of sociophysics. *Journal of Statistical Physics* 151: 9–20.

Stewart JQ (1945) *Coasts, Waves, and Weather for Navigators*. Boston: Ginn and Co.

Stewart JQ (1950) The development of social physics. *American Journal of Physics* 18: 239–253.

Stewart JQ and Warntz W (1958) Macrogeography and social science. *Geographical Review* 48: 167–184.

Tocalis T (1978) Changing theoretical foundations of the gravity concept of human interaction. In: Berry BJL (ed.) *Perspectives in Geography 3: The Nature of Change in Geographical Ideas*. DeKalb, IL: Northern Illinois University Press, pp. 67–124.

Warntz W (1957) Contributions toward a macroeconomic geography: A review. *Geographical Review* 47: 420–424.

Warntz W (1959a) *The Geography of Price*. Philadelphia: University of Pennsylvania Press.

Warntz W (1959b) Geography in the mid-twentieth century. *World Politics* 11: 442–454.

Warntz W (1964) A new map of the surface of population potentials for the United States, 1960. *Geographical Review* 54: 170–184.

Warntz W (1977 [1967]) Spatial Order – Concepts and Applications, Part 1. *Harvard Papers in Theoretical Geography: Geography and the Properties of Surfaces Series*: 1–83.

Warntz W (1984) Trajectories and co-ordinates. In: Billinge M, Gregory D and Martin RL (eds) *Recollections of a Revolution: Geography as Spatial Science*. London: Macmillan, pp. 134–150.

Wilson MW (2014) Morgan Freeman is dead and other big data stories. *cultural geographies*. Published OnlineFirst, doi: 10.1177/1474474014525055.

Wilson MW and Graham M (2013) Situating neogeography. *Environment and Planning A* 45(1): 3–9.

Zipf GK (1932) *Selected Studies of the Principle of Relative Frequency in Language*. Cambridge, MA: Harvard University Press.

Zipf GK (1941) *National Unity and Disunity*. Bloomington, IN: Principia.

Zipf GK (1942) The unity of nature, least-action, and natural social science. *Sociometry* 5: 48–62.

Zipf GK (1949) *Human Behavior and the Principle of Least Effort*. Cambridge, MA: Harvard University Press.

Zipf GK (1968) *The Psycho-Biology of Language: An Introduction to Dynamic Philology*. Cambridge, MA: MIT Press.