4-22-2014

Historical Health Conditions in Major US Cities: The HUE Data Set

Carlos Villareal
The University of Illinois at Chicago

Brian Bettenhausen
Booth School of Business

Eric Hanss
Booth School of Business

Jonathan Hersh
Chapman University, hersh@chapman.edu

Recommended Citation
Bettenhausen, B., Hanss, E., Hersh, J., & Villarreal, C. (2014). Historical health conditions in major US cities: The HUE data set. *Historical Methods* , 47(2), 67-80. doi: 10.1080/01615440.2013.874005

Historical Health Conditions in Major US Cities: The HUE Data Set

Comments

This is an Accepted Manuscript of an article published in *Historical Methods*, volume 47, issue 2, in 2014, available online at DOI: 10.1080/01615440.2013.874005. It may differ slightly from the final version of record.

Copyright

Cities in Time: The Historical Urban Ecological Dataset[1]

Abstract

The Historical Urban Ecological (HUE) data set is a new resource detailing health and environmental conditions within seven major US cities from the earliest available records through 1930. We collected and digitized 3365 tables of ward-level data from annual reports of municipal departments that detail the epidemiological, economic and demographic conditions within each city. We then drafted new Geographic Information System (GIS) data to link the tabular records to ward geographies. These data provide a new foundation to investigate the policies and investments that influenced the urban mortality transition and the growth of US cities.

Carlos Villarreal
Department of Economics
University of Illinois at Chicago
and
Center for Population Economics
Booth School of Business
5807 South Woodlawn Ave.
Chicago, IL 60637
carlos@cpe.uchicago.edu

Brian Bettenhausen
Center for Population Economics
Booth School of Business
5807 South Woodlawn Ave.
Chicago, IL 60637
bbett@cpe.uchicago.edu

Eric Hanss
Center for Population Economics
Booth School of Business
5807 South Woodlawn Ave.

[1] We cannot list here all of the outstanding assistance we have received in the completion of this project, but we would like to thank Lou Cain, Nathaniel Grotte, Hoyt Bleakley, and Dora Costa for their extensive comments. The inimitable Noelle Yetter, Tom Blaser, Heather DeSomer, Lewis Meineke, Alexander Orsini, Anna Wielgosz and Andrea Zemp provided outstanding research assistance. These data are compiled as a part of the *Early Indicators of Later Work Levels, Disease and Death* project (NIH program project grant P01 AG10120).

Chicago, IL 60637
ehanss@cpe.uchicago.edu

Jonathan Hersh
Department of Economics
Boston University
270 Bay State Rd.
Boston, MA 02215
jhersh@bu.edu

> "In all the problems we may devise for the sanitary or the social welfare of this great metropolis, we must accept and duly estimate the fact that its vast population is already more densely crowded in its domiciles than that of almost any other city; and that the evils attendant upon overcrowding and the aggregation of vast numbers will be continually augmented as the population increases, unless the resources of Sanitary Science and the beneficent operations of wisely-administered sanitary regulations are interposed."
> 1865 Report of the Council of Hygiene and Public Health of the Citizens Association of New York upon the Sanitary Condition of the City, p. xxxix

Introduction

The crude mortality rate in the seven largest US cities dropped from 22.31 deaths per 1000 persons in 1890 to 12.44 in 1930[2]. Early studies on this transition focused on reductions in mortality for cities on the whole (Meeker 1972; Conrad and Crimmins-Gardner 1978); water filtration, chlorination (Cutler and Miller 2004; Troesken 2004), the construction of sewer systems (Costa and Kahn 2003), sanitation capital (Cain and Rotella 2001) and gross population (Cain and Hong 2009) drove much of the city-level reductions in waterborne disease mortality through this period. Mortality rates, however, varied more within cities than between them,[3] and little is known about health patterns within cities. Existing intra-city studies are limited either to a specific city (Conrad and Cheney 1982), year (Higgs and Booth 1979) or disease (Craddock 2000; Hinman 2002; Hinman, Blackburn and Curtis 2006) and several fundamental questions have not been adequately addressed: what was the structure of neighborhood-level health changes during the urban mortality transition? What role did particularly good or bad neighborhoods play in driving this phenomenon? Were intra-city transition patterns comparable between cities? Improvements in the urban health environment improved unevenly within cities,

[2] Census Office, *Report on Vital And Social Statistics in the United States at the Eleventh Census: 1890. Part II – Vital Statistics. Cities of 100,000 Population and Upward*, 1896; Bureau of the Census, *Mortality Statistics*, 1931
[3] For example, in 1890 mean infant mortality (IMR) in the city of Chicago was 212 deaths per 1000, but it varied between individual wards from 95 to 415 deaths per 1000.

and until now researchers have lacked the neighborhood-level data and analytical framework to investigate intra-urban variation in health environment.

We built the Historical Urban Ecological (HUE) data set to investigate the changing health and environmental conditions within seven of the largest US cities from the earliest available ward-level records through 1930. The data include over 3,000 tables detailing crime, disease and health, vital and demographic statistics, property and land values, and tax information recorded at the ward level by municipal departments in Baltimore, Boston, Brooklyn, Chicago, Cincinnati, New York and Philadelphia. While Federal Censuses have recorded intra-city data at the ward level, the HUE data set is the first to digitize the large amount of data collected by municipalities, providing inter-censal statistics and greatly expanding the scope of feasible inquiries. These ward-level data would be of limited use without information on the spatial extent of the wards. Urban redevelopment programs in the mid- and late twentieth century razed many historically prominent neighborhoods, making contemporary maps inadequate for the reconstruction of the historical urban environment. Identifying the historical ward boundaries thus required the comprehensive reconstruction of historical cities in a geographic information system (GIS).

We used GIS tools to draft a new set of historical street centerlines in the seven study cities and from those data reconstructed ward boundary histories from 1830 through 1930. These GIS data provide the critical information to link and compare tabular ward-level data spatially and longitudinally. This project contributes to a growing body of historical GIS (HGIS) work. Over the last decade, the National Historical GIS Project[4] (Fitch and Ruggles 2003), the Atlas of Historical County Boundaries,[5] and a number of international GIS's (Gregory et al. 2002; De Moor and Wiedemann 2001) exposed historical census, parish, and other aggregated data to modern spatial methodologies. Extant work at the urban and intra-city level has heretofore been limited to either a specific year (Logan et al. 2011) or city (DeBats and Lethbridge 2005; Gilliland and Olson 2003; Gilliland, Olson and Gauvreau 2011; Lutz et al.). Historical changes to administrative boundaries have long impeded intercensal and intracensal quantitative spatial analysis. The NHGIS created tract-level boundaries from 1910 onwards, but earlier intra-city census geographies have not been consistently available. Wards were well-defined political and administrative areas within cities that, in addition to being censal aggregation areas from 1790 through 1940, were the primary scale at which municipal departments reported aggregate statistics. The HUE GIS captures all of the ward boundary changes, expansions, and subdivisions for the seven study cities, extending the usability of intra-city census data back to 1830.

The HUE GIS's historical street centerlines also create a framework for the development and integration of increasingly detailed historical data: sewer or school districts, transit and water

[4] https://www.nhgis.org
[5] http://publications.newberry.org/ahcbp/index.html

systems, and address-level points of interest or disease cases are a few possibilities. Statistics can be calculated at several levels of spatial granularity for the multilevel analysis of area effects. The core HUE GIS and collected tabular data will be available for visualization and download at the Center for Population Economics website[6] in January 2013 and further data releases of city infrastructure and health microdata will be published through the year.[7]

The HUE data are of interest not only to researchers interested in the health environment within cities, but also to economists, historians, geographers, and demographers interested in cities more broadly. Allison Shertzer (2012) used a preliminary version of HUE to examine the political integration of new immigrants within cities. Carlos Villarreal (dissertation) is using the HUE maps to examine the persistence of poor neighborhoods in Manhattan from 1830 to the present as measured by rents and incomes. He finds that areas with poor natural drainage, a disamenity before the installation of sewer pipes, remained undesirable despite the later introduction of water and sewer pipes which should have eliminated the initial disadvantage of marsh locations. Using the HUE data, he is able to investigate several explanations for the persistence of this initial disadvantage, including the durability of housing, the location of polluting manufacturers, and immigrant enclaves.

The remainder of this paper provides an in-depth description of HGIS methodology and collected tabular data. We conclude with early applications and research extensions.

2. Reconstructing Historical Cities Using GIS

A GIS of historical street centerlines forms the core of the HUE data set. We developed networks of street centerlines for each of the sample cities as they were in 1930 and subsequently added administrative boundaries, sewer and water pipelines and point-level address data. Each supplementary data set is spatially and topologically coherent and allow for spatial linking of disparate historical data sets. Wards were the most common geographies municipal departments used for statistical aggregation, and the HUE GIS data provide ward boundary histories for each sample city that span the century of vigorous urban development from 1830 to 1930. Street centerlines, wards histories with collected ward-level tabular data and additional smaller scaled spatial data, organized under the HUE data framework, will allow researchers to perform analysis over longer intervals than previously available at an intra-urban scale. The following sections document the creation of these core HUE GIS components and present potential applications.

2.1 Constructing Historical Street Network Maps

[6] Hue data will be accessible at the CPE's website http://www.cpe.uchicago.edu and through the NBER data portal http://www.nber.com/data.

[7] Please direct all inquiries to Research Manager Brian Bettenhausen bbett@cpe.uchicago.edu or Research Program Coordinator Eric Hanss ehanss@cpe.uchicago.edu.

The HUE GIS street centerline networks are based on a single source street map of each sample city produced in or around 1930.[8] By this time, Manhattan, Brooklyn, Boston, Chicago, Cincinnati, Philadelphia, and Baltimore had, for the most part, reached their modern spatial extents. Between 1830 and 1930 the street networks managed by each city grew by the annexation of neighboring areas and the development of new infill and sparsely populated areas. Street removal was difficult and unnecessary, so the growth of the street networks before 1930 was largely additive. As a result, earlier street networks in each city can be considered a subset of the 1930 street networks that compose the HUE GIS street centerline data, and they can be used for accurate spatialization of data throughout the period.[9]

While the 1930 HUE street networks serve as consistent references for earlier periods, they diverge significantly from those of the present day. In the mid- to late-twentieth century American city street networks experienced significant programs of demolition, redirection, and expansion resulting from urban renewal programs, redevelopment, and the construction of interstate highways and cross-urban expressways. For this reason, our survey of publicly available modern GIS data revealed that official census geographies and street centerlines available from the United States Census Bureau (TIGER/Line shapefiles) and private and municipal GIS resources would not be suitable for a historical GIS reconstruction.[10] The sections of street networks that have remained unchanged since 1930, however, provide the starting point for our historical reconstruction.

Those modern street centerlines that had not changed since 1930 were identified by comparing the 1930 street maps with modern satellite imagery. Once identified, unmodified centerlines were drawn referencing modern orthoimagery downloaded from the National Map, a contemporary spatial database provided by the U.S. Geological Survey. Unmodified streets in the historical core of each sample city were reproduced in this way, as was the entirety of Chicago and Manhattan Island. Where cities annexed substantial districts outside the historic urban core modern TIGER centerlines provided unmodified streets geographies.[11] This method

[8] Due to the sporadic availability of appropriate historical maps, we were unable to find comprehensive, detailed examples for every city at the appropriate extent and scale in 1930. The date of publication of maps used ranged from 1922 to 1932. Private and municipal cartographers created these maps for use as road maps, street directories, and for administrative and planning purposes.

[9] Many streets were renamed throughout the study period, largely as a result of name redundancy after annexation of neighboring areas. By employing period-specific maps as reference, historical addresses and ward boundaries can still be located. Several cities, such as Philadelphia, also maintain queryable street name change databases (http://www.phillyhistory.org/historicstreets/). Alternate or duplicate street names are not included in HUE GIS data.

[10] The US Census Bureau's TIGER/Line geographies were used extensively only as the MAF/TIGER Accuracy Improvement Project (MTAIP) approached completion. Highly accurate street networks are crucial, as errors in the initial street reference maps propagate into subsequent layers of spatial data.

[11] These areas include suburban Philadelphia and Cincinnati, the additional area annexed by Baltimore in 1919, the neighborhoods adjacent to Boston proper, and southern Brooklyn.

provides a level of spatial accuracy that would be difficult to achieve with georeferenced historical maps alone.

We then used these networks of unmodified streets to interpolate those streets and blocks that had been significantly modified or disappeared altogether during the intensive urban reconfiguration since 1930. Figure 1 shows a sizeable area west of downtown Chicago that was razed to accommodate a highway interchange in the late 1950's. Our interpolation of 1930 street centerlines is shown both in the left panel, which shows the modern orthoimagery that we used to create the centerlines, and in the right panel, showing the demolished neighborhood and the streets in the aerial photograph circa 1930. Despite significant alterations, many of the streets that surrounded the razed district persist through to the present day and allow us to reconstruct the missing streets by extending the extant centerlines through demolished areas in a manner consistent with the intersections and termini marked on the historical street maps.

[Figure 1 Here]

Some street centerlines, however, could not be satisfactorily extended through demolished areas and their realignment required additional effort. Streets that curved within razed districts proved to be especially complicated. For example, in Boston's West End the construction of a complex of high-rise apartment towers and the Massachusetts General Hospital altered the street network to the extent that our standard interpolation approach was inadequate. In these cases, centerlines were drafted over georeferenced Sanborn Fire Insurance Map plates and USGS topological surveys dating to the period around 1930.[12] Figure 2 shows the HUE 1930 centerlines, modern orthoimagery and the Sanborn map used in our reconstruction.

[Figure 2 Here]

The accuracy of the HUE street centerlines maps has two main caveats. First, while historical aerial surveys of Chicago and New York discovered after the reconstruction project was completed, such as that presented in Figure 1, provided reassurance that these reconstruction techniques provided a high degree of spatial accuracy, they are not the result of surveying, and cannot substitute for modern GIS data. Secondly, for data integrity purposes, HUE GIS centerlines reflect a single source street map. These source maps often reflected future plans that may or may not have been implemented. They do not provide perfect snapshots of the cities' contemporary street networks. Beyond confirming the spatial accuracy of HUE centerlines, the early aerial surveys also showed streets in Chicago and southeast Brooklyn that had not yet been completed by the 1930's. We have no solid evidence that HUE street centerlines in southwest

[12] Sanborn Fire Insurance Maps were produced irregularly in atlas form. Consequently, the publication dates of maps used differ and are frequently from years before or after 1930. Sanborn maps are among the most detailed sources of urban spatial data available before the second half of the twentieth century. They provide an immense amount of information on the built environment at fine spatial resolution (1:600 scale or 50ft to the inch).

Philadelphia were ever more than contemplated. Despite these limitations, the HUE street centerline networks function best as historically appropriate and spatially accurate reference layers on which further data can be constructed.

2.2 Constructing Ward Boundary Histories

The HUE data set provides a full set of the historical ward systems used in each sample city from 1830 to 1930. Descriptions of ward boundaries and the dates each system was adopted and discarded were determined from a survey of archival and reference documents, municipal reports and ordinances, collected ward maps and contemporary studies.[13]

Ward boundaries were delimited by street centerlines, surveyed city limits, and other geographic features such as shorelines, waterways and railroads. HUE ward boundary polygons were drawn along the 1930 street centerline networks[14] and additional bounding features, such as shorelines or abstract boundary lines, were derived from georeferenced maps and HGIS resources.[15]

The cities' ward boundary histories are highly idiosyncratic. In Manhattan ward boundaries were not modified after 1854, while Philadelphia changed its ward boundaries twenty times between 1830 and 1930. The quality of changes to ward systems also varied between cities. Figure 3 compares Philadelphia and Boston ward boundary changes between early and later points in the study period. Philadelphia wards in the central business district remained stable and comparable as peripheral wards divided through the study period. Boston ward boundaries, on the other hand, often changed quite significantly, and were further altered by extensive landfill.[16] The HUE ward histories provide scholars the opportunity to observe changing ward boundaries throughout the study period and to accurately link tabular ward-level data on population and public health compiled for the HUE data set to period-specific geographies.

[Figure 3 Here]

3. Building Tabular Ward-Level Data

---

[13] The ward boundary histories of Baltimore and Philadlephia were taken from secondary compilations, Manhattan from archival research aids provided by the New York City Hall Library. The remaining ward histories were compiled by CPE researchers. Specific documentation for each city's ward history is available for download on the CPE website.

[14] There were no cases in which earlier ward lines ran along streets that had been demolished by 1930. Street widening or straightening, however, was not taken into account in the construction of historical ward boundaries.

[15] In cases where historical city limits were analogous to historical county limits, as in the case of Philadelphia after 1854, the 1930 county boundary was drawn from the Minnesota Population Center's National Historical GIS. Where historical city limits remained constant to the modern day, contemporary city boundaries were used as a guide.

[16] See Gregory (2002) for a more thorough treatment of GIS techniques for working with changing aggregation geographies through time.

During the study period, municipalities primarily reported intra-city statistics at the ward-level. State and federal censuses often reported ward-level statistics as well. The sheer quantity of data available makes analyzing cities at the ward level an attractive option for intra-city analysis. Since collection and publication of the municipal ward-level data was at the discretion of the cities themselves, statistics collected varied across time and between cities. The statistics published by each city generally expanded over the study period. Early reports tend to be terse, reporting only key vital statistics or statistics idiosyncratic to the city. The scope of published statistics expanded during the latter half of the nineteenth century, responding in part to the cities' desire of to track and prevent disease outbreaks. As medical diagnosis methods improved, cities published more data on cause-specific case and mortality rates. City health departments were in close contact with each other, which resulted in the rapid dissemination of advances in medicine and the equally rapid introduction of new statistics into annual health reports.[17]

In our collection efforts, we attempted an exhaustive search of municipal reports' recordings of ward-level statistics during the study period. We collected 3,365 tables of ward-level statistics from 312 unique sources. We digitized each table by hand, attached source and metadata, and compiled the data into a relational database. The collection process yielded a total of 89,324 observations of 9,425 unique, ward-level variables from 1830 through 1930.

The database's size created several problems of taxonomy. Variable categorization was not merely required *post facto* for extraction purposes, but to aid inputters in the assignment and aggregation of variables as they were found across time and source material. Strict rules were maintained in variable naming conventions and categorization. Nonetheless, as the number of variables being input increased, their categorization had to adapt and expand to contain them. In the end, the tabular data are divided into five broad categories: crime, disease, municipal, property, and vital statistics and 19 sub-categories. Table 1 gives an overview of each main category and presents the earliest year each type of data are available.

[Table 1 Here]

Crime refers to statistics related to criminality, including homicides and crime reports. *Disease* refers to specific disease recording, and contains the cases, deaths and vaccination subcategories. *Municipal* refers to data on tax receipts, sewer and water outlays, schools, animal populations and elections statistics. *Property* refers to data on value and amount of personal property, chiefly real estate stock; and finally *vital statistics* refers to data on births, deaths and population. We see from the table that there is a great deal of variation for when each city began reporting different statistics. Boston began reporting vital statistics in 1849, followed by Philadelphia in 1860, New

[17] For example, the 1874 New York Report of the Board of Health includes a discussion (p.26-27) on whether a recent cholera infection in New Orleans was Asiatic cholera or its more benign cousin cholera-morbus. A doctor was dispatched to determine the nature of this strain, and in his subsequent report he describes conferring with physicians in Cincinnati, Nashville, and Murfreesboro, TN on its pathology and how it should be appropriately classified (p.415-418).

York City (only encompassing Manhattan Island at this time) in 1865, and Chicago later still beginning in 1866. Disease reporting expanded in the 1870s.

Table 2 highlights a sample of more specific variables available by time period.[18] The nine statistics presented illuminate the disparities between each city's reporting and collection efforts. Vital statistics are the most consistently reported across cities. Population and deaths are present by ward for the majority of the sample period, births are somewhat less represented. In the disease variables, we have shown two sample statistics: cases and deaths from scarlet fever and smallpox. Early reporting on these diseases was sporadic and generally linked to disease outbreaks in the cities. As time went on, reporting became more regular and consistent. Smallpox data appears to have been collected frequently in early years, but not collected with the same frequency in the decades of 1900s and 1910s, with the exception of Philadelphia and New York. As a sample of other municipal records available four statistics are presented: tax receipts, number of registered voters, acres of land, and value of real estate. Here we see the divergence of statistical reporting diligence; Boston published these statistics most consistently, showing relatively few gaps in the post-Civil War period. New York published acres of land and value of real estate consistently after the Civil War, while Chicago and Philadelphia only published these statistics intermittently.

[Table 2 here]

4. Four Sample Use Cases

This section shows four use cases for the HUE GIS. The first is a straightforward creation of chloropleth maps using the tabular ward-level data with the HUE ward boundary files. The next two show how the HUE GIS data, used as geospatial references, give a framework for creating new historical data and the fourth describes how HUE GIS data can revitalize existing datasets and provide new dimensions for linking and analyzing disparate data sets.

The creation of chloropleth maps, maps colored to represent values in data, are so fundamental to GIS visualization that the functionality has been built directly into the HUE data portal. For more rigorous analysis and customized visualization, the website provides tabular data extracts, GIS data, and pertinent source and metadata information for download. These files can be used with any standard desktop GIS software to create chloropleth maps to the users specification. Figure 4 is a simple example of such a map. We selected Baltimore infant mortality in 1906.[19] The city and year of the statistic indicate the relevant ward boundary map is Baltimore 1902 - 1918.

[18] For reasons of brevity, Cincinnati and Baltimore have been excluded in the table. Brooklyn and Manhattan have been aggregated into New York.

[19] Being between quinquennial census years, these data are not available from any federal census extract. They were collected and reported by the city in tables 3 and 5 of the *Annual Report of the Board of Health to the Mayor and City Council of Baltimore* for the year 1906, and digitized as part of our tabular ward-level data collection efforts.

Combining the ward boundary map and the tabular statistics we can reveal patterns in the spatial distribution of infant mortality in that period. So with relative ease, the HUE dataset can display data with new temporal and spatial granularity.

[Figure 4 here]

The HUE historical GIS street centerline networks also provide a base reference for the creation of new spatial data. Address- or block-level data can be accurately placed and novel relations between datasets can be discovered by aggregating to uniform areal units or by multi-scale methodologies. New spatial data sets may be address-level disease cases or census information, street level information on sewer and water systems or other aggregation geographies such as school and sanitary districts. Using HUE GIS street centerlines to place these data, they can be used independently or spatially linked to HUE ward-level tabular data.

In the next use case, we show the collection of block-level sewer pipes for the City of Chicago from 1857 to 1910. In this period Chicago's Board of Sewerage Commissioners and subsequently the Department of Public Works recorded the construction of sewers on a block-by-block basis.[20] Figure 5 shows the digitized map of these reports. The data are accurate to the year and presented in 3 gross periods. At the block-level, these data can aggregated to ward geographies to interact directly with HUE tabular data, they could be displayed over tabular data of dysentery cases and rates, or they can be linked to address level data to show in what year an individual would have had access to the public sewer system. The CPE is collecting and digitizing similar sewer and water system construction data for each of the target cities, though there is variance in data availability.

[Figure 5 Here]

The nineteenth century also saw the expansion of public transportation networks that significantly reduced travel times within cities, affecting the travel times between different neighborhoods, and creating new paths for infectious disease transmission. Figure 6 shows the New York City passenger transportation system in 1878. The lack of descriptions of their construction, however, precluded the block-by-block collection method used for city water and sewer. Being of general interest, many maps were discovered which contain information on the location of these historical transit routes. From these maps and collected local transportation histories we were able to digitize the urban transportation networks block-by-block, and derive approximate dates of opening and abandonment of each individual line.[21] The data were created

[20] Ward-level water and sewer data were rarely reported in annual municipal reports; it was more common for municipal departments overseeing infrastructure installation to report specific locations and cost of new pipe segments.

[21] Maps either were not published or were not available for every year in every city. Consequently, we do not necessarily know the precise years in which each line was first opened and/or finally abandoned. Special attention must be paid to the publication years of the maps used in the reconstruction. A detailed list of the map sources is available on the CPE website.

using HUE street centerlines and so can be used in combination with other data to derive distances to transit, travel times and, in concert with ward-level data, to examine the effects of transit on the development and make-up of the city.

[Figure 6 Here]

The HUE GIS also provides a framework to give new life to other historical data sets. The Philadelphia Social History Project (PSHP) directed by Theodore Hershberg at the University of Pennsylvania between 1969 and 1981 created a machine-readable database including detailed records of over 500,000 individuals from population census schedules, 100,000 firms from city business directories and manufacturing census schedules, and a detailed reconstruction of the transportation network spanning the period between 1838 and 1880. [22] These data, which advanced the study of industrialization, social stratification, individual mobility, and the uses of urban space in nineteenth-century Philadelphia (see Hershberg 1981), have not yet been systematically incorporated into a GIS framework due to the lack of appropriate historical geographies, though the data has seen limited use in several recent case studies (Condran 2008; Hargis and Horan 2004). The HUE GIS provides the necessary references for locating the address and ward-level data, but further efforts will be required to rebuild the grid unit system used by the PSHP. Once assembled in a modern GIS, researchers can analyze the PSHP data using techniques that had not been developed in the early 1980s. Using the HUE GIS as a construction framework, one could integrate the PSHP data with HUE ward-level data, increasing the richness of the data base and allowing for new research.

5. Conclusion

The cases presented in the previous section represent a small sample of the possible applications of the HUE data set. Users from a wide range of disciplines will find the spatial framework a useful reference over which any historical urban data can be projected and analyzed in conjunction with or separate from the rich ward-level data supplied with the GIS.

Researchers at the Center for Population Economics have already begun to employ HUE in ongoing studies of the effects of the urban environment on aging and work force participation. Under the auspices of the *Early Indicators of Later Work Levels, Disease and Death* project (EI), we have geocoded 13,607 addresses for 3,839 Union Army and US Colored Troops living in the study cities. Integrating military, pension and census records into lifetime health and economic histories, the EI data set is one of the richest collections of historical microdata publicly available. By linking these individuals to their physical locations and the environmental conditions they encountered throughout their lives we can measure the impact of disease

[22] www.icpsr.umich.edu/icpsrweb/ICPSR/series/229

environment, neighborhood socioeconomic and ethnic composition, and public health reforms on later-life health outcomes and explore disparities in longevity and work-force participation levels between healthy and unhealthy wards.

The HUE data set provides new tools by which to better understand the mechanisms of health and mortality distribution, disparity, and change. Preliminary research has concentrated on economic divergences within cities, but the opportunities for epidemiological, historical and demographic research are as yet under-utilized. By placing historical data in a modern GIS framework, the HUE data set opens the historical situation of American cities to modern empirical methodologies from a variety of disciplines, its extensibility allows for the integration of newly created or revitalized data, and opens new avenues into the discussion of urban form and development.

References

Beveridge, A. A. 2008. A century of Harlem in New York City: Some notes on migration, consolidation, segregation, and recent developments. *City & Community* 7:358–65.

Cain, L., and S. C. Hong. 2009. Survival in 19th century cities: The larger the city, the smaller your chances. *Explorations in Economic History* 46.6:450-63.

Cain, L., and E. Rotella. 2001. Death and Spending: Did Urban Mortality Shocks Lead to Municipal Expenditure Increases? *Annales de démographie historique*, 2001-1:139-54.

Condran, G. A. 2008. The elusive role of scientific medicine in mortality decline: Diptheria in nineteenth- and early twentieth-century Philadelphia. *Journal of the History of Medicine and Allied Sciences* 63.4: 484-522.

Condran, G. A., and R. A. Cheney. 1982. Mortality trends in Philadelphia: Age- and cause-specific death rates, 1870-1930. *Demography* 19.1: 97-123.

Costa, D., and M. Kahn. 2003. Public health and mortality: What can we learn from the past?. Paper presented at the Berkeley Symposium on Poverty, the Distribution of Income and Public Policy, December.

Council of Hygiene and Public Health. 1865. *Report of the Council of Hygiene and Public Health of the Citizens Association of New York upon the sanitary condition of the city*. New York: D. Appleton and Company.

Craddock, S. 2000. *City of Plagues: Disease, poverty, and deviance in San Francisco.* Minneapolis: University of Minnesota Press.

DeBats, D. A., and M. Lethbridge. 2005. GIS and the City: Nineteenth-century residential patterns. *Historical Geography* 33:78-98.

De Moor, M., and T. Wiedemann. 2001. Reconstructing territorial units and hierarchies: A Belgian example. *History & Computing* 13:71- 97.

Fitch, C. A., and S. Ruggles. 2003. Building the National Historical Geographic Information System. *Historical Methods* 36.1:41-51.

Gilliland, J. A., S. H. Olson, and D. Gauvreau. 2011. Did segregation increase as the city expanded? The case of Montreal, 1881-1901. *Social Science History* 35.4:465-503.

Gilliland, J. A., and S. H. Olson. 2003. Montréal, l'avenir du passé. *GÉOinfo* January-February:5-7.

Glover, J. W. 1921. *United States Life Tables, 1890, 1901, 1910, and 1901-1910*. Washington, DC: Government Printing Office.

Gregory, I. N. 2002. The accuracy of areal interpolation techniques: Standardising nineteenth- and twentieth-century census data to allow long-term comparisons. *Computers, Environment and Urban Systems* 26.4:293-314.

Gregory, I. N., et. al. 2002. The Great Britain Historical GIS Project: From maps to changing human geography. *The Cartographic Journal* 39.1:37-49.

Haines, M. R. 2001. The urban mortality transition in the US, 1800-1940. Cambridge, MA: NBER, Historical Research Paper No. 134.

Hargis, P. G., and P. M. Horan. 2004. Bounded by culture or culture bound: Ethnicity, schooling, and the interplay of theory and evidence. *Historical Methods* 37.1:23-33.

Health Department of the City of New York. 1874. *Fourth annual report of the Board of Health. May 1, 1873 to April 30, 1874*. New York: D. Appleton and Company.

Hershberg, T. 1981. *Philadelphia: Work, space, family, and group experience in the nineteenth century, essays toward an interdisciplinary history of the city*. New York: Oxford University Press.

Hinman, S.E., J. K. Blackburn, and A. Curtis. 2006. Spatial and temporal structure of typhoid outbreaks in Washington, D.C., 1906-1909: Evaluating local clustering with the G_i* statistic. *International Journal of Health Geographics* 5.13.

Logan, J. R., et. al. 2011. Mapping America in 1880: The Urban Transition Historical GIS Project. *Historical Methods* 44.1:49-60.

Lutz, J., Dunae, P., Gilliland, J., Lafreniere, D., and Harvey, M. In Press. Turning space inside out: Spatial history and race in Victorian Victoria. In *Assessing the landscape: Historical GIS research in Canada*, edited by J. Bonnell and M. Fortin. Calgary: University of Calgary Press.

Perry, N. 2012. Southern suburb/northern city: Black entrepreneurship in segregated Arlington County, Virginia. *Urban Geography* 33.5:655-74.

Preston, S., and Haines, M. R. 1991. *Fatal years: Child mortality in late nineteenth-century America.* Princeton: Princeton University Press.

Shertzer, A. 2012. *Immigration and Cities in the Twentieth Century*. UCLA: Dissertation.

Villarreal, C. *The Influence of the Initial Environment on Housing Prices: Evidence from New York City 1830-1940.* University of Illinois at Chicago: Dissertation.

FIGURES

Figure 1: Reconstruction of the Historical Chicago Streets
Left panel: historical street reconstruction with contemporary satellite imagery. Right panel: historical street reconstruction with 1939 aerial photography.

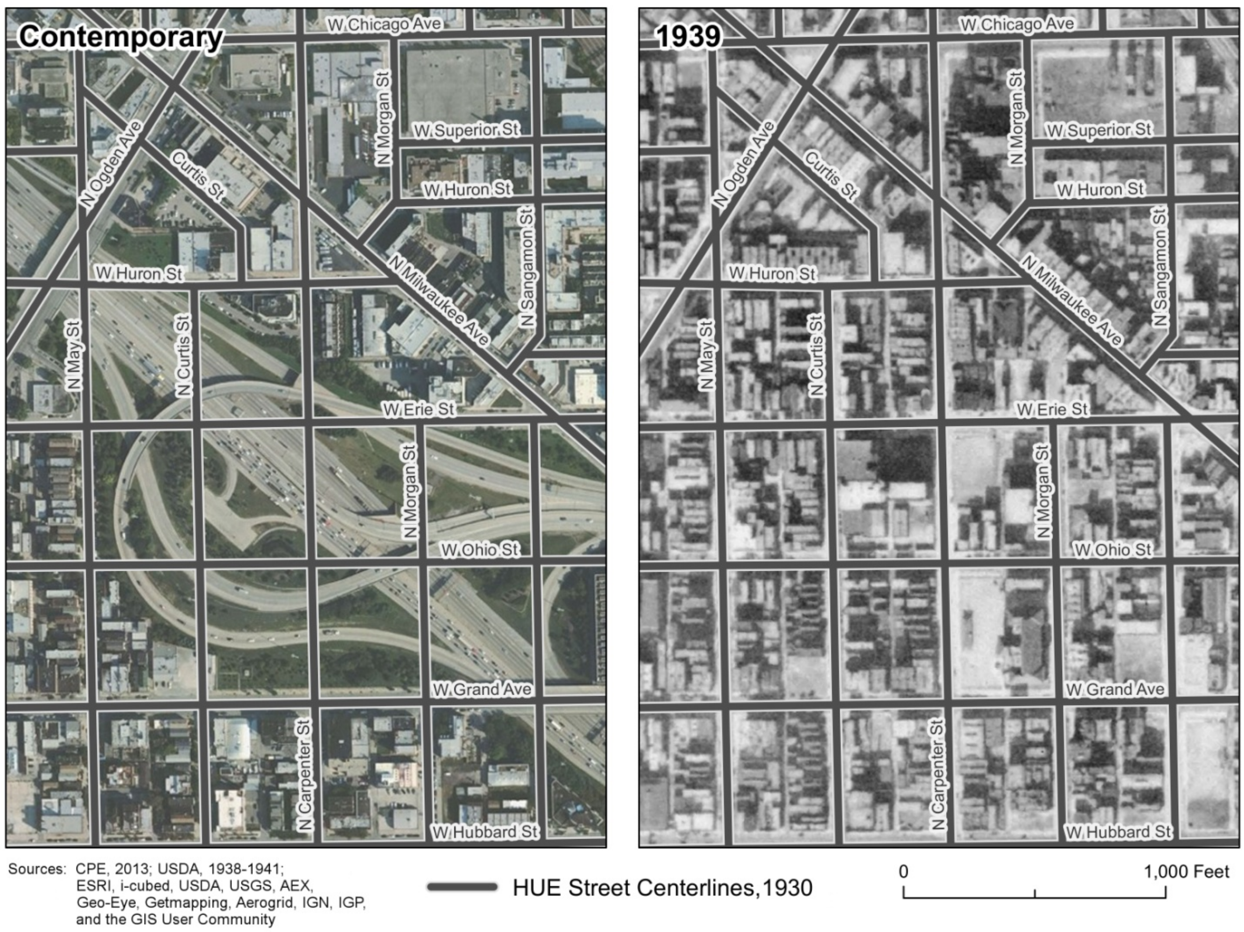


Figure 2: Reconstruction of Historical Boston Streets: West End Neighborhood
Left panel: historical street reconstruction with contemporary satellite imagery. Right panel: historical street reconstruction with 1885 Sanborn map plate.

Figure 3: Reconstruction of the Historical Philadelphia and Boston Ward Boundary Systems
Left panel: Philadelphia Wards 1855 and 1914. Right panel: Boston Wards 1830 and 1913.

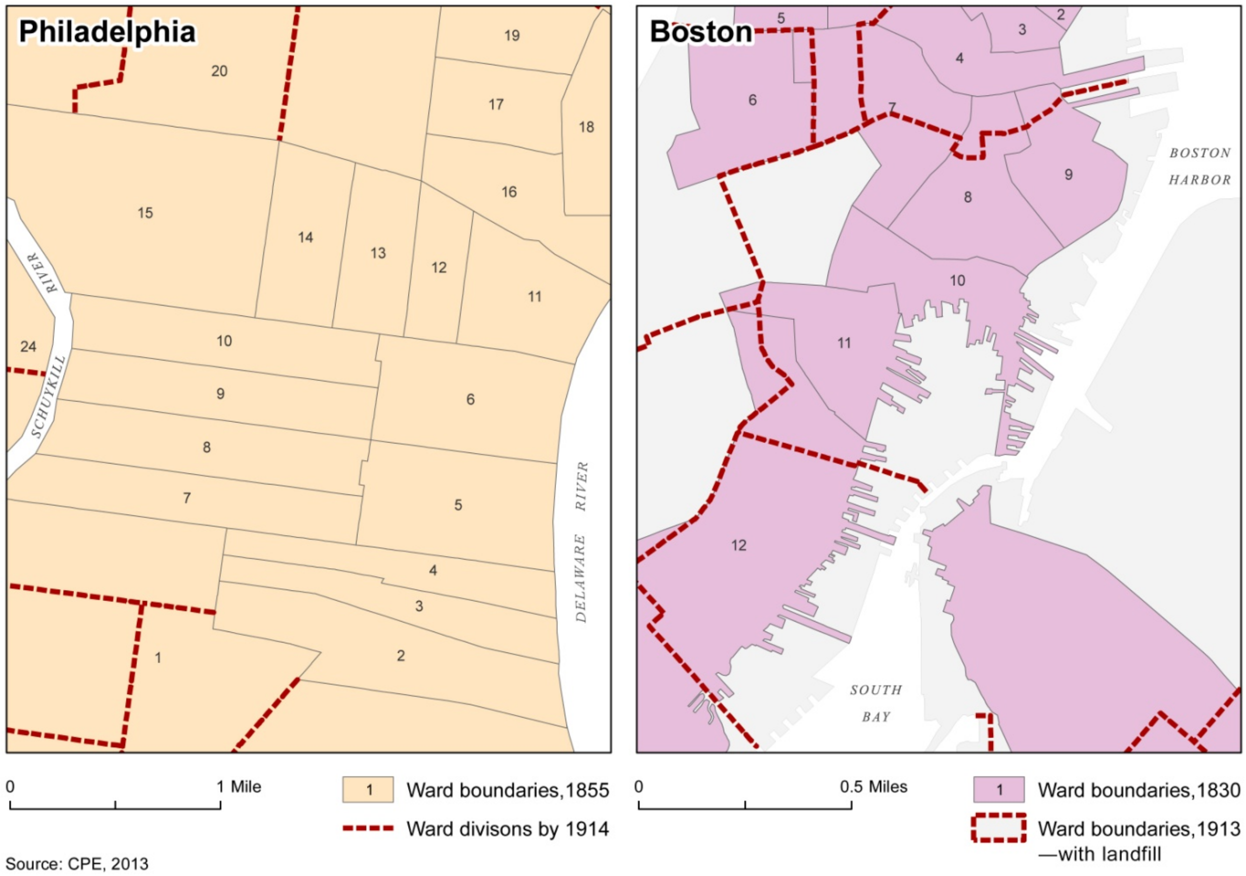


Figure 4: Ward-Level Infant Mortality Rate, Baltimore 1906

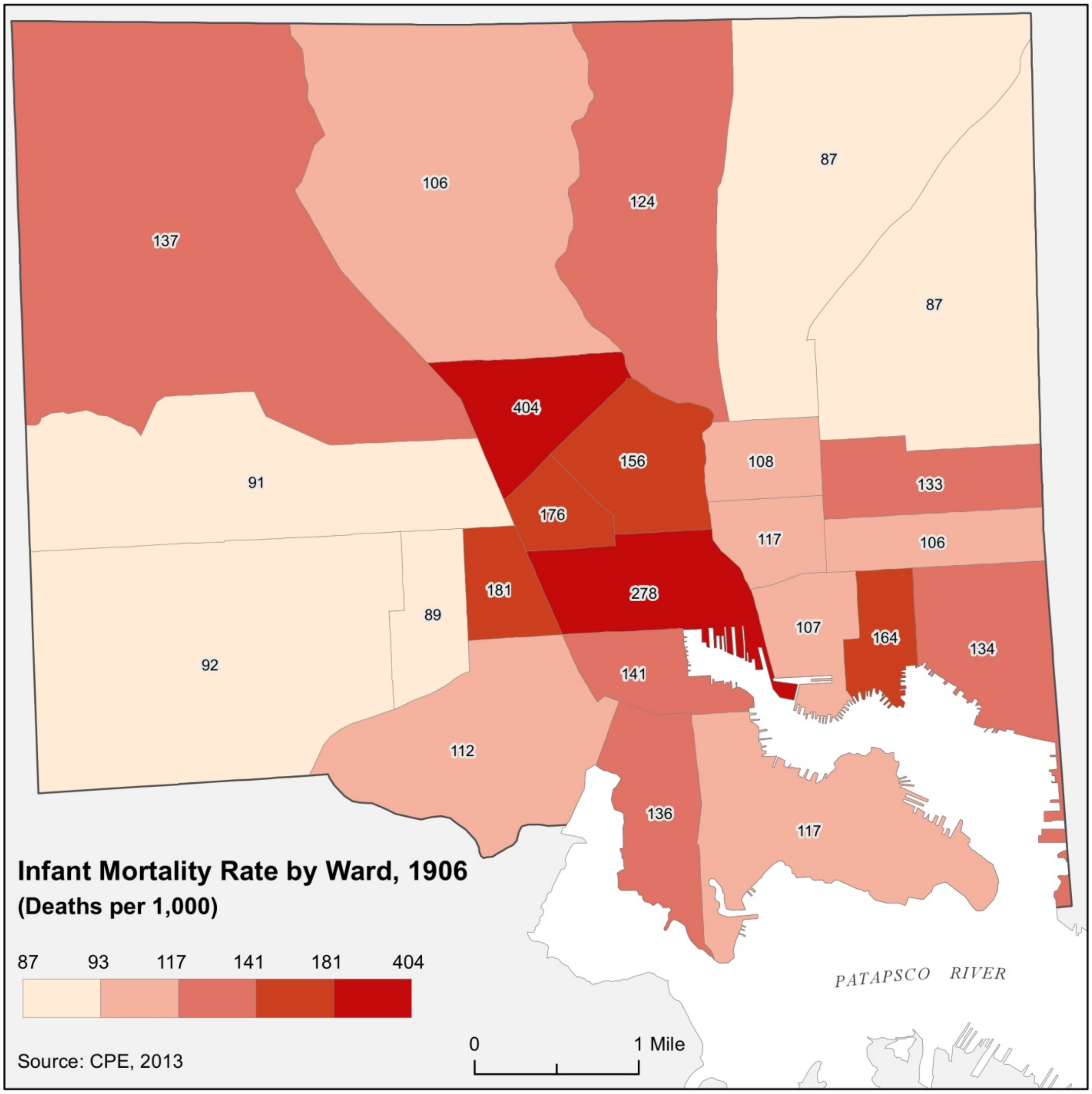


Figure 5: Block-by-block Installation of Chicago Sewer System from 1852 to 1910.

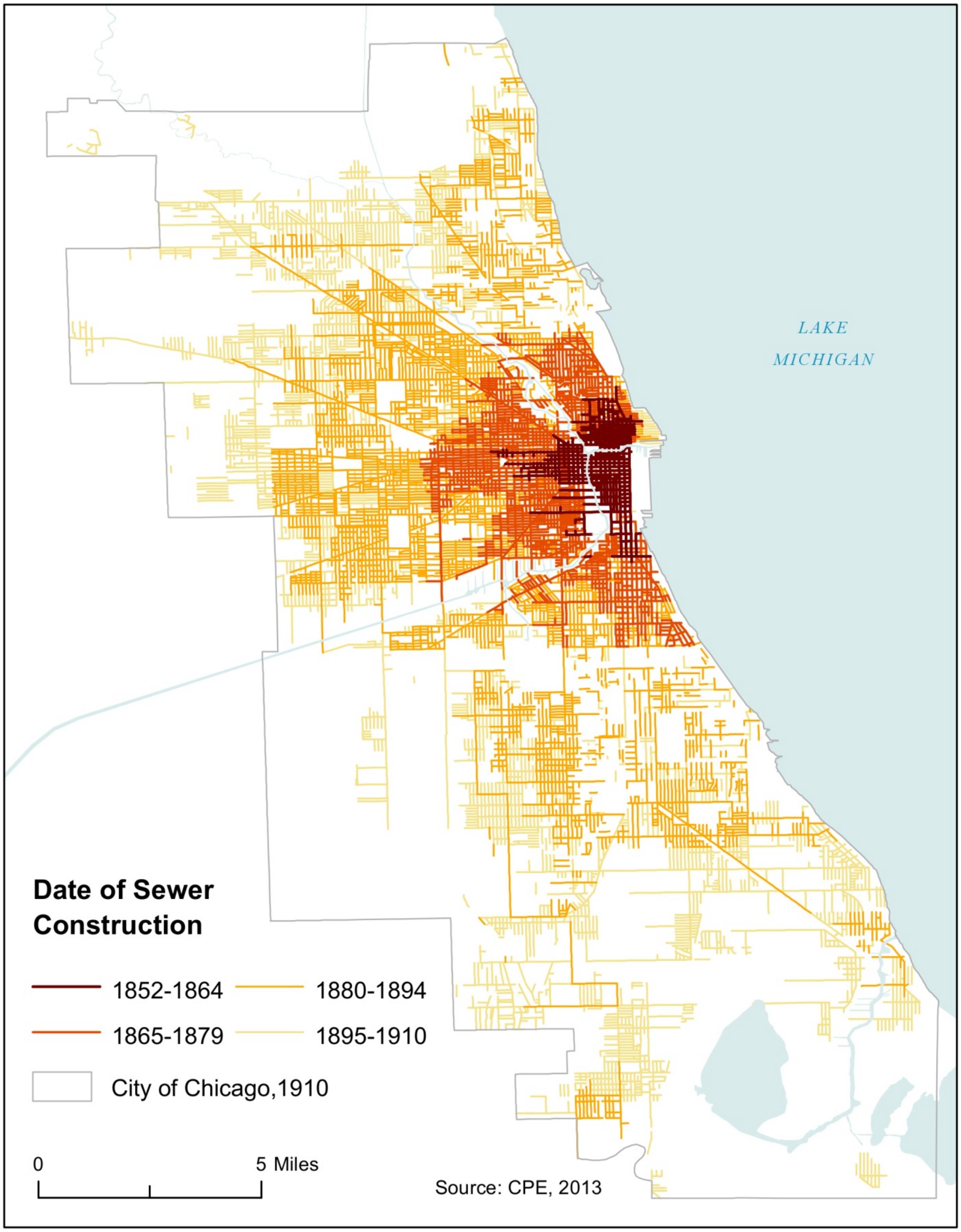
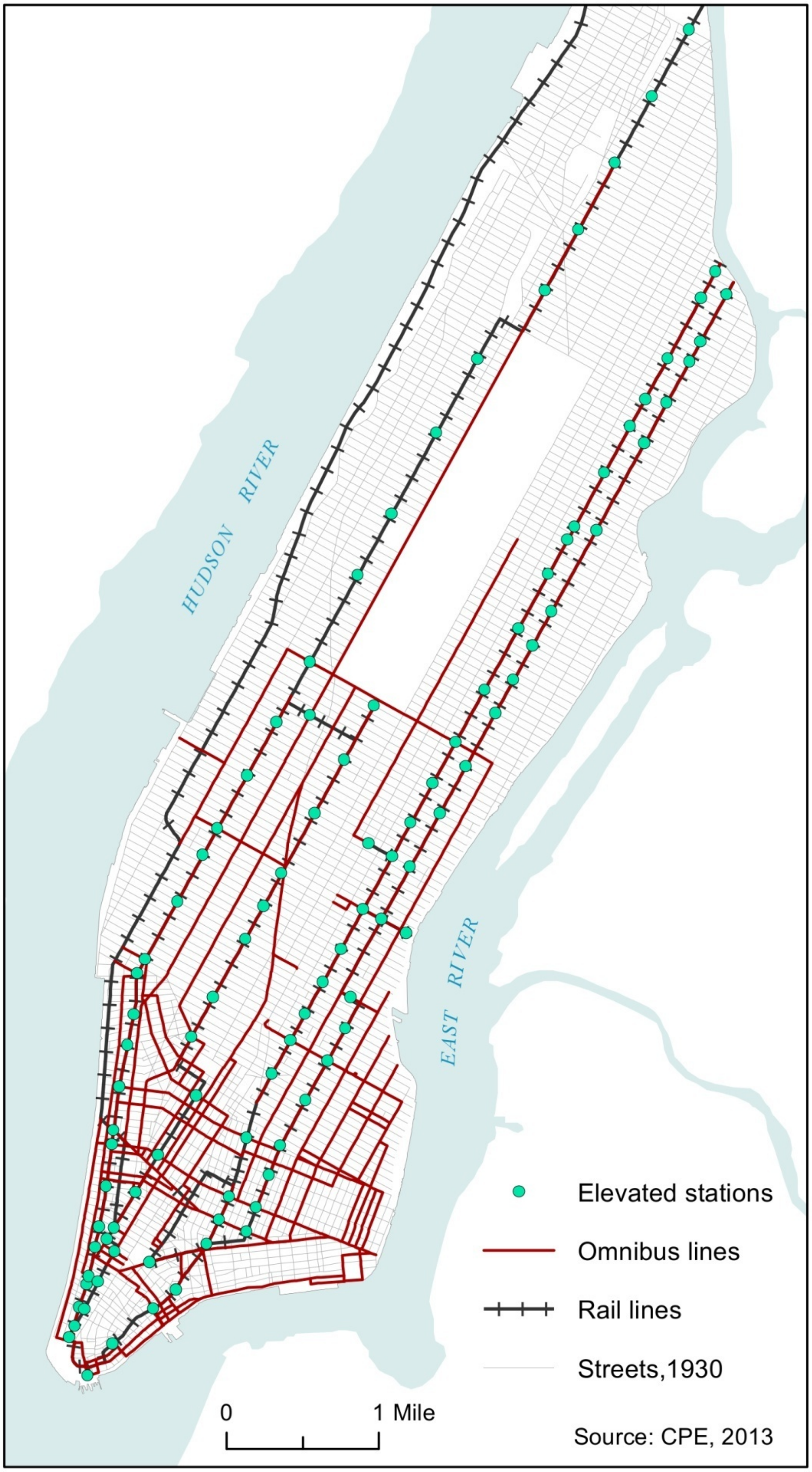


Figure 6: Manhattan Island Transportation Networks, 1878.



TABLES

Table 1: Earliest Available Ward-Level Data from Municipal Sources by Broad Category

City	Crime	Disease	Municipal	Property	Vital Statistics
Baltimore	-	1881	1888	1890	1879
Boston	1900	1877	1800	1810	1849
Brooklyn	1892	1867	1855	1838	1870
Chicago	1875	1866	1860	1866	1866
Cincinnati	1875	1874	-	1870	1867
Manhattan	-	1874	1854	1830	1865
Philadelphia	1863	1861	1877	1875	1860

Note: Crime refers to statistics related to criminality (e.g. homicides); diseases include cases of and deaths from specific diseases; municipal records include results from municipal elections and tax statistics; property refers to values and amount of personal property; vital statistics are primarily births and deaths.

Table 2: Summary of a Sample of Available Ward-Level Characteristics

City	Decade	Vital Statistics			Disease		Other Municipal Records			
		Population	Deaths	Births	Scarlet Fever	Smallpox	Tax Receipts	Number of Registered Voters	Acres of Land	Value of Real Estate
Boston	1830-39	x	-	-	-	-	x	x	-	x
	1840-49	x	-	x	-	-	x	x	-	x
	1850-59	x	x	x	x	x	-	-	-	-
	1860-69	x	x	x	-	-	-	x	x	x
	1870-79	x	x	x	x	x	-	x	x	x
	1880-89	x	x	x	x	x	x	x	x	x
	1890-99	x	-	-	-	-	x	x	x	x
	1900-09	x	x	x	x	x	x	x	x	x
	1910-19	x	x	x	x	-	x	x	x	x
	1920-29	x	x	x	x	-	x	x	x	x
Chicago	1860-69	x	x	x	x	x	-	-	x	-
	1870-79	-	x	x	-	x	-	-	-	-
	1880-89	x	x	-	x	x	-	-	x	-
	1890-99	x	x	-	x	x	-	x	-	-
	1900-09	x	x	-	x	-	-	-	x	-
	1910-19	-	x	-	x	-	-	-	-	-
	1920-29	-	x	-	x	x	-	-	-	-
New York*	1830-39	x	-	-	-	-	x	-	-	x
	1840-49	x	-	-	-	-	x	-	-	x
	1850-59	x	x	-	-	-	x	-	-	x
	1860-69	x	x	-	x	x	x	x	x	x
	1870-79	x	x	-	x	x	x	-	x	x
	1880-89	x	x	-	x	x	-	-	x	x
	1890-99	x	x	x	x	x	-	-	x	x
	1900-09	x	x	-	x	x	-	-	x	-
	1910-19	x	x	-	x	x	-	-	x	-
	1920-29	x	x	-	x	x	-	-	x	-
Philadelphia	1860-69	x	x	x	x	x	-	-	-	-
	1870-79	x	x	x	x	x	-	-	x	x
	1880-89	x	x	x	x	x	-	-	x	-
	1890-99	x	x	x	x	x	-	-	x	-
	1900-09	x	x	x	x	x	-	-	x	-
	1910-19	x	x	x	x	x	-	-	x	-
	1920-29	-	x	x	x	x	x	-	-	-

Note: An "x" indicates at least one record exists of the variable in that decade. For reasons of brevity Cincinnati and Baltimore have been excluded.

*New York includes Brooklyn and Manhattan.