Spring 5-11-2018

Urban Scaling and the Impact of Socioeconomic Performance on Health and Environmental Outcomes

Sky Tallman

Candidate

Sky Tallman

Department

Community and Regional Planning

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

Bill Fleming, Chairperson

Caroline Scruggs

Richard Epstein

Urban Scaling and the Impact of Socioeconomic Performance on Health and Environmental Outcomes

by

Sky Tallman
Bachelor of Arts in History, minor in German, University of New Mexico, 2005
Master of Arts in Liberal Arts, St. John's College, 2016

THESIS

Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master's in Community and Regional Planning

The University of New Mexico
Albuquerque, New Mexico

July, 2018

Urban Scaling and the Impact of Socioeconomic Performance on Health and Environmental Outcomes

Sky Tallman
B.A. History, University of New Mexico
M.A. Liberal Arts, St. John's College
M.C.R.P. Community and Regional Planning, University of New Mexico

Abstract:

This thesis explores the question of whether correlations between socioeconomic indicators and health and environmental indicators can be observed at the metropolitan level of analysis. Indicators are considered both in terms of per capita values and scale-adjusted values. Scale-adjusted values are a concept based on research on urban scaling which account for the agglomeration effect of population on socioeconomic output and describe a city's performance in terms of how it compares with expected performance on an indicator for its population size based on the power law scaling of that indicator to population. These indicators provide an alternative baseline for comparing the relative economic performance of cities of different sizes. Considering these scale-adjusted indicators a potentially more meaningful measure of socioeconomic performance, they were expected to correlate with health and environmental outcomes better than per capita values, although this was only true for median household income.

One of the questions driving this inquiry was whether the theory of urban scaling might be useful in explaining the variation in indicators which do not scale with population, but which have been shown in other research to relate to socioeconomic factors. In this study, those factors are air quality, tree cover and twelve health outcomes and risk factors collected by the CDC. In all but a few cases, stronger correlations could be seen with per capita socioeconomic values than with scale-adjusted values.

Data for metropolitan and micropolitan statistical areas were gathered on a wide range of socioeconomic indicators and tested for correlations with health and environmental outcomes. Socioeconomic data that scaled with population were also calculated as a scale-adjusted metropolitan indicator (SAMI) value and tested for correlations with health and environmental outcomes. The expectation was that SAMI values would help explain health and environmental outcomes, which was found generally not to be true. From the data, however, some correlations between socioeconomic and health indicators could be observed, while tree canopy and impervious surface showed inconsistent correlations with health and economic indicators in different ecoregions.

Summary of key findings

- Median household income was the only indicator tested for which the scale-adjusted value correlated more strongly with health indicators than the unadjusted value (p. 84-5, 89, 91, 103).
- More binge drinking in a city correlated with better health, crime and socioeconomic outcomes. It correlated most strongly with walking to work. (Tables 3 and 4).
- Asthma is least influenced by socioeconomic indicators, with the GINI coefficient having the strongest negative correlation (meaning that as inequity rises, asthma rates fall) and walking to work shows the strongest positive correlation (meaning that asthma rates rise as more people walk to work), although the coefficient is too low to be significant (Tables 2, 3 and 4). The AQI was most strongly correlated with federal civilian spending; aggregate travel time per person and walking showed some correlation but were not significant (Table 2).
- Means of transportation were most strongly correlated with overall health and obesity, but were not correlated with air quality or the prevalence of asthma, and travel time to work correlated significantly only with sleep (Table 3).
- As commute times get longer, more people walk to work (Figure 9).
- A 1% change in driving alone to work has the same effect on COPD as a $2631 change in scale-adjusted median household income (about 0.5% of MHI) (Table 15, p. 78).
- Means of transportation, whether walking, bicycling or driving alone to work, correlated strongly with all health outcomes except mental health and asthma (500 Cities data), but the portion of people taking public transportation did not. BRFSS data on obesity correlated significantly and negatively with use of public transportation. Means of transportation to work did not, however, correlate with any measurement of air quality, which is surprising due to the substantial contribution of automobiles to air pollution and the negative correlation between density and commute times. (Tables 2 and 3)
- Tree cover is higher in wealthier neighborhoods, but not in wealthier cities.
- There were no correlations between socioeconomic or health indicators and tree cover that were similar in all regions. Obesity was most consistently and negatively correlated, followed by high blood pressure, sleeping less than seven hours a night and physical health. Tree cover in the transition ecoregion, between the woodlands of the east and the desert and plains of the Midwest, showed the greatest number of significant correlations with health and socioeconomic indicators. At a national level, mental health was significantly correlated with tree canopy, but was only significantly correlated in the mountain west when data were separated by ecoregion. (Table 10).
- Tree cover was more influential than crime on mental health and crime was more influential than tree cover for high blood pressure. Both crime and tree cover had a similar effect on sleeping less than seven hours a night. (Table 19).

- Driving time to work and choices of mode of transportation were not correlated with air quality indicators at a metropolitan level (Table 2).
- Poor neighborhoods have more exposure to air pollution, but measures of economic performance at a metropolitan level are not correlated with air quality (Tables 2 and 16).
- Some correlations have been found between neighborhood tree canopy and air quality. At the metropolitan level, this was found only in the transition and mountain west ecoregions (Table 10).
- AQI is not correlated with health outcomes at the metropolitan level (Table 2).
- No apparent loss of productivity or increase in crime correlates with binge drinking (Tables 4 and 5).
- The city can be considered a risk factor based on how it provides alternative transportation infrastructure, walkable neighborhoods and commercial districts, decent wages and other services to residents, such as increasing the length and quality of education achieved by and available to residents of a city, and to the extent that it creates policies that redistribute wealth and alleviate poverty, as all of these factors can be linked to health outcomes (p. 83, 104-5).
- The healthiest/least healthy cities performed similarly on the GINI index, poverty, and the lowest quintile's share of aggregate income (Table 7).
- In the healthiest cities, the percentage of people who bicycled to work was ten times higher, while the percentage who walked to work was twice as high than in the least healthy. There was a 14% difference in attainment of higher education and the healthiest cities generally had *less* tree canopy coverage than the least healthy (Table 7).
- The healthiest cities had higher MHI, SMHI, educational attainment, rent, walking bicycling and state and local spending and lower unemployment, driving alone and murder (Table 7).
- Sublinear scaling of infrastructure: There should be no assumption that infrastructure costs offset by scale-based translate into more cost-effective provision of other services. The health impact of state and local government spending per capita does not depend on population size (p. 88).
- Five of the seven cities with the highest cancer rates had the lowest rates of obesity and five had the lowest percentage of people sleeping less than seven hours per night. Five of the eight cities with the highest rates of diabetes had the lowest rates of asthma, and three of the five cities that had among the highest rates of both diabetes and obesity had the lowest rates of cancer (p. 52).
- The SAMI for median household income showed no significant correlation with any scale-adjusted crime category, although it was significantly correlated with the per capita values for property crime and burglary. The SAMI for Aggregate Household Income showed significant correlations with the SAMI values for most crime indicators. The SAMI value for walking and driving alone to work showed the strongest negative and positive correlations with the SAMI values for most crime indicators (Table 5).

Contents

List of Figures and Tables

Introduction

Urban scaling is a statistical phenomenon widely observable in urban systems suggesting that some common aspects of individual behavior produce similar city-scale results regardless of geographic, economic or cultural context. The most important of these common behaviors is the way in which social networks form and behave. With this as a baseline, scaling can, theoretically, be applied to any system of cities with similar results. Many socioeconomic and infrastructural indicators scale non-linearly with population in similar and predictable ways in diverse systems of cities. This happens because certain properties of cities, such as productivity, crime and infrastructure are emergent properties of complex systems that arise from common behaviors of individuals. How cities scale is an empirical description of nonlinear patterns explained in terms of network theory, but lacks a practical application beyond affirming that there are economies of agglomeration and savings on infrastructure as cities get larger. Indeed, scaling is a sort of simplified model of agglomeration, yet the theory offers little that helps to explain the socioeconomic over- or underperformance of particular cities. That is, over- or underperformance should be a reflection of the physical and social connectivity or dissipation of energy in a city (Bettencourt 2013A), yet high and low-performing cities are not clearly distinguishable in these terms.

One key difference between scaling theory and other models of economic agglomeration is that scaling uses population as the key figure against which to predict a wide variety of other urban outcomes. Population is not the best predictor of other indicators, including health outcomes, tree cover or air quality, but because. Scaling with population predicts a wide range of socioeconomic and infrastructural indicators in different urban systems, and the scaling of socioeconomic and infrastructural output is explained as an emergent product of nearly universal behavioral tendencies of individuals. The theory argues that social networks and mixing are a common explanation, but has not gone so far as to explain differences in economic performance or in or infrastructural economies of scale between particular cities. Given the broad applicability of urban scaling to different systems of cities, this thesis asks whether scale-adjusted socioeconomic values are better predictors of health and environmental values that do not scale with population.

Scaling theory considers population as an aggregate proxy for a wide range of other factors, and presumes a few common underlying “rules” driving human behavior that are scale invariant. Population essentially determines certain aspects of the large-scale results of individual behavior patterns because the size of social networks and nature of social mixing are largely a function of population size. Productivity and other social and economic metrics scale with population in regular ways because they are strongly influenced by local social networks and mixing. The structure of social networks in cities represents a sort of universal constant on which the scaling of socioeconomic and infrastructural values is based (West 2017).

Agglomeration models can predict specific socioeconomic outcomes for cities with a higher degree of accuracy than scaling, even without using population as a variable, but these models cannot be as readily translated from one system of cities to another, and are often geography or indicator-specific. For this reason, scaling theory may represent a sort of breakthrough that will lead to something like a universal theory of urbanization. However, it has yet to become a tool for policymakers and urban planners to be able to describe or explain the performance of individual cities in more detail than in how they compare to other cities. In theory, travel costs and spatial connectivity within a metropolitan area should correlate with social mixing, and therefore economic output, but this cannot be easily observed in cities with high or low scale-adjusted socioeconomic performance. The size and quality of social connections have been shown to be related to health outcomes including tuberculosis, psychological health and mortality (House et al. 1989). If social connectivity is a significant part of the explanation for scale-adjusted over- or underperformance, health indicators would be expected to show a consistent correlation with scale-adjusted economic performance. In this study, only one scale-adjusted economic indicator, median household income, showed stronger correlations with health outcomes than corresponding per capita indicators.

Scaling allows for the comparison of cities based on a non-linear, moving baseline of expected performance, accounting for how a city would be expected to perform on a particular indicator based on its size. By accounting for scale, cities can be ranked based on their scale-adjusted performance, which is useful in comparing the relative performance of cities of different sizes. By itself, this may be a useful way of describing performance on certain indicators, but it has limited utility in that it does not provide insight into how policymakers might improve performance in a low performing city. Planners could use these results to intentionally address issues of public health by targeting socioeconomic indicators associated with certain health outcomes, or to garner public support for economic initiatives by showing how they will impact public health. A policy, for example, to raise minimum wages, and thereby the median household income could be correlated with measurable expectations for how it would impact a range of health indicators and how it would correlate with a change in tree cover in certain ecoregions.

This paper explores the question of whether scale-adjusted measurements of socioeconomic performance correlate with or help explain cities' performance on health and environmental indicators that do not scale with population size. With the exception of median household income (MHI), indicators measured in per capita terms showed stronger correlations with health and environmental outcomes in cases where there was any correlation at all. The case of MHI allows some of the relationships between economic performance and health outcomes to be explained through the lens of scaling theory, while the correlations observed between other socioeconomic indicators and health and environmental outcomes reveal some interesting insight. For example, tree cover at the metropolitan level does not correlate with income, except in a few ecoregions, and tree cover does not correlate with the Air Quality Index, and only in two ecoregions is there a significant correlation between tree canopy and the percentage of days with unhealthy levels of ozone or particulate matter of 2.5 microns; metropolitan level air quality indicators do not correlate with health outcomes, and air quality is unconnected to economic performance at the metropolitan level.

At a neighborhood or individual level of analysis, environmental indicators such as tree cover and air quality have been shown to correlate with socioeconomic qualities such as income or race in certain cities that have been studied (Gerrish and Watkins 2018), though these correlations weaken considerably or disappear when multivariate analyses consider race along with other variables (Schwarz 2015). Tree cover is significantly correlated with income, with poorer neighborhoods tending to have less tree cover. Health indicators, too, at a neighborhood level of analysis, have been shown to correlate with the socioeconomic status of individuals and neighborhoods.

Asthma tends to be more severe (though perhaps not more prevalent) in poorer neighborhoods. Exposure to traffic-related air pollution contributes both to the severity of asthma symptoms and to its onset (Guarnieri and Balmes 2015, Rona 2000), while obesity has also been shown to correlate with poverty in developed economies (Levine 2011), and the prevalence of chronic obstructive pulmonary disease (COPD) has also been shown to have a strong connection with socioeconomic status.

If the correlations between socioeconomic performance and health outcomes or environmental qualities are persistent and regular, then they should be observable at different scales of analysis. There is strong evidence, for example, that urban structure at a neighborhood level and density at the level of an urban area have a significant impact on CO_2 emissions. Urban form at a neighborhood level has a significant impact on transportation choices (Holtzclaw 1994, Kühnert 2006, Ewing 2010), though these impacts are often insignificant at the level of analysis of metropolitan statistical areas. Thus, the CO_2 emissions of metropolitan areas scale sublinearly with population size, regardless of the density of the urban core (West 2017).

Considering whether the median or scale-adjusted performance of a city on certain socioeconomic metrics correlates with health outcomes could reshape conversations around policy in which the city could be considered a risk factor, rather than individual circumstances, choices or neighborhoods. Considering the socioeconomic performance of a metropolitan or micropolitan area as a risk factor for health outcomes could create more equitable outcomes than considering risk as uniquely affecting only certain populations. Choices in mode of transportation taken to work, for example, are strongly influenced by urban form (Ewing 2010) and have a moderate or strong correlation with health and environmental outcomes. While these health outcomes may disproportionately impact lower income communities or individuals, it may be more meaningful to think of mode of transportation as a collective choice or as the result of policy, infrastructure and safety than as a question of individual choices. Given the strong correlation between urban form and choice of mode of transportation, mode of transportation may also be thought of as a proxy variable in its effect on health, representing the relationship between urban form and health outcomes.

The degree to which a whole metropolitan area over- or underperforms relative to its scale-adjusted expected performance may be expected to show a similar correlation to health and environmental factors as has been observed in higher and lower performing neighborhoods within cities. Adjusting for scale should provide a more meaningful baseline with which to compare the economic performance of cities (West 2017). It is surprising to find that only median household income shows a stronger correlation with health outcomes, yet scale-adjusted values for other indicators, including aggregate income, GDP, real income (adjusted for regional price parities), rent, educational attainment, poverty rates, state and local spending, usual hours worked, crime rates for different types of crimes, and modes of transportation to work did not show stronger correlations when adjusted for scale. The SAMI for income deficit and federal spending showed stronger correlations with some indicators, though correlations with federal spending were generally very low.

What this has shown is that, although there is a relationship between economic performance and environmental and human health, scaling does not generally illuminate this relationship better than per capita data. Adjusting for scale only proved a stronger correlation with data in the case of MHI. Because scaling theory describes the superlinear scaling of a wide range of socioeconomic indicators as the outcome of the same underlying pressures, and only this indicator shows stronger correlations with health indicators than its per-capita corollary, scaling theory is probably not a strong explanation for this correlation. Yet, for some reason, scale-adjusted median household income is more strongly associated with most health outcomes than unadjusted median household income.

Understanding the relationship between metropolitan-level socioeconomic data and environmental and health outcomes may prove a useful tool in advocating for policies that more holistically address social, economic and environmental goals. Although the question of whether scaling would show a stronger relationship to health and environmental data than data using per capita or percent values, the analysis of the data unveiled some interesting patterns of correlation between socioeconomic performance and health outcomes. It also showed that correlations that have been observed at the neighborhood level between air quality, tree canopy and socioeconomic status are not clearly observable when comparing urban areas across metropolitan areas, even when accounting for ecological zones.

The audience for this work includes people who see scaling as a new potential tool for crafting urban policy as well as people who are interested in the relationship between health outcomes and socioeconomic performance. Although the results relating to tree cover and air quality are not conclusive, the positive trends shown in this analysis indicate that this is an area of great interest to urban foresters and policymakers, and these results may further other research projects.

There is a popular literature on urban scaling, and an audience that lacks training in urban economics, could easily be led to believe that scaling represents a new lens through which to understand urban economic development. This thesis is not an argument against the possibility that urban scaling may represent a new tool, rather it is an exploration into one possible way in which urban scaling may prove applicable.

The relationship between health outcomes and socioeconomic indicator levels at the metropolitan level may be of interest to many disciplines, including epidemiology and insurance.

Context for the Problem

The context for this problem is primarily theoretical and will focus on data for US metropolitan and micropolitan statistical areas. This Census definition has been chosen because Metropolitan and Micropolitan Statistical Areas (MMSAs), or Core Based Statistical Areas, are defined as a core urbanization of at least 10,000 people and include the county containing the urbanization and any counties from which a significant portion of the population regularly commute to the urban core. Thus, MMSAs approximate a web of socioeconomic and face-to-face interactions, which is the important aspect of a city through which scaling theory explains why cities scale with size. The scaling of socioeconomic indicators has been intensively studied, but because environmental and health indicators do not scale with population in the same way, the significance of socioeconomic scaling to a city's health and environmental performance, or vice versa, remains largely unstudied. If there proves to be a significant relationship, it could have implications for understanding the interrelationships between environmental, health and socioeconomic development policies.

Because scaling relationships simply compare the performance of cities within a system, they do not necessarily reveal a great deal of meaningful information about the conditions within a particular city or the causes of its over- or underperformance (Olpadwala 2017), thus investigating whether health or environmental indicators correlate with scale-adjusted socioeconomic performance could potentially add significance to such measures of relative performance.

This work does not attempt to make an argument that socioeconomic factors are causative of health outcomes, so although the correlations that were found were only weak or moderate, the finding of any correlation is of interest as it suggests that socioeconomic indicators have a significant influence on health indicators but lack a consistent influence on environmental outcomes. The correlations with health outcomes are in line with what has been observed in other studies, however the correlations, or lack of correlations, with environmental outcomes suggest that existing literature focusing on certain cities or areas may be more idiosyncratic than universal.

Analytic timeframe

The last Economic Census was done in 2012, and this is the source of some of the socioeconomic data that will be used in this analysis. Because other data are generally available for 2012, wherever possible, 2012 data will be used to make results and discussions most meaningfully comparable. Scaling relationships with indicators that have been studied show a great deal of stability over decades (Bettencourt 2010), therefore data from 2012 should scale statistically similarly to current data and should be roughly comparable to data from other nearby years, should some data prove to be unavailable from 2012. This study will aim to analyze urban scaling and its relation to health and environmental data in 2012 and will not focus on changes over time.

Literature Review

A Brief Overview of Scaling

Over the past decade or so, a vein of research on cities has explored the phenomenon of urban scaling, which describes how socioeconomic and infrastructural quantities in cities scale nonlinearly, as a power law function, with population size.

Similar patterns of urban scaling can be observed within systems of cities in different cultural and national contexts and even in ancient systems cities. Cities tend to display a remarkable degree of regularity and predictability in how socioeconomic factors and infrastructure lengths, such as the lengths of pipes, wires, roads, etc., scale with population size. The power law scaling of urban indicators means that any city twice the size of another city in the same system will not simply have twice as much of everything, rather, it is likely to demonstrate economies of scale on many aspects of infrastructure and increasing returns to scale for many socioeconomic indicators such as per capita earnings, rate of innovation, and productivity. Even quantities such as the number of restaurants, lawyers and doctors in a city increases superlinearly at about the same rate as socioeconomic quantities (see Bettencourt et al. 2007, 2010, 2013, 2014; Batty, 2013; Schlaepfer 2014; Ortman 2015, West 2017). That is, socioeconomic factors tend to scale superlinearly with city size, while infrastructure tends to scale sublinearly, all at a rate of about 15%.

Aggregate metropolitan socioeconomic quantities, such as gross metropolitan product and productivity, tend to rise around 115% with a doubling of population, while the total infrastructural quantities, such as lengths of roads, wires or pipes used in the city tend to rise by only 85% (ibid, 275). Some other quantities such as crime and disease also scale superlinearly at about the same rate as socioeconomic factors. Cities also scale with respect to land area, using on average use only about 2/3 as much land per capita with a doubling of population (Bettencourt – Origins of Scaling).[1] A sublinear scaling between land area per capita and population has also been reported by Louf (2014), and it makes intuitive sense that as cities get larger, they tend to become more dense.

West (2017) claims that population size alone in the United States, "can predict with 80-90% accuracy what the average wage is, how many patents [a city has] produced, how long all of its roads are, how many AIDS cases it's had, how much violent crime was committed, how many restaurants there are, how many lawyers and doctors it has, et cetera" (West 2017, 278). Beyond the US, population size accounts for 65-97% of the variance in observed data when looking at systems of cities. West goes further to claim that we can use scaling to know a great deal about indicators in other countries. By knowing how the scaling of US cities compares to that of Japanese cities, one could estimate the crime rate in a Japanese city of a given size based on crime rates in any US city (West 2017). West concedes that the wide variation among cities conditions

[1] The relationship between land area and population is based on observations of Metropolitan Statistical Areas, the definition of which is includes all of the counties containing or significantly connected to the urban core. With this definition, the sublinear scaling of land area per capita does not necessarily represent a greater efficiency of land use per capita.

the above claim, but emphasizes that scaling gives a baseline from which to describe an indicator (such as crime rates) as high or low (ibid).

Population should be understood not as the cause of scaling, but as a strong approximate variable representing a measurement for the diverse set of socioeconomic interactions and outcomes that characterize a city (Bettencourt 2010).

Scaling patterns describe economies of agglomeration and increased savings on infrastructural costs as cities grow. The scaling of productivity with population size emerges out of the relationship between physical and social networks; location, therefore, is a function of interaction (Batty 2013).

All systems of cities (that have been analyzed, at least) scale in the same way because of a “universal socioeconomic dynamic reflecting average organizational behavior of human interactions” (Bettencourt 2010) regardless of historical, cultural, economic or technological circumstance. West (2017) defines cities as an “emergent self-organizing phenomenon that has resulted from the interaction and communication between human beings exchanging energy, resources and information” (West 2017, 280).

Bettencourt (2013), explains the apparent regularities in how such a diverse range of indicators scale in cities as evidence of a "universal socioeconomic dynamic reflecting average organizational behavior of human interaction in cities"(2). Scaling is a function of population size and urban area as well as the average number of local social interactions a person is likely to have, the average value of those social interactions (which can be positive or negative), and the cost of travel. Essentially, it is a function of the ability of a city to facilitate the generation and maintenance of the social networks and a mixed population.

Scale-Adjusted Metropolitan Indicators

One of the key ways in which scaling may prove useful to planning is that it enables the prediction of expected values or outcomes in a city based on population size. Scaling can be used to account for expected differences between large and small cities by adjusting urban metrics based on how they scale across cities in the system. By adjusting for scale, cities of different sizes can be easily compared based on the deviation of their actual performance from the value predicted by the power law trend line that describes an indicator's scaling (See Appendix B. for national SAMI distribution with labels on New Mexico cities).

Researchers at the Santa Fe Institute have suggested that we might measure cities' performance based on a scale-adjusted residual. These are referred to as Scale-Adjusted Metropolitan Indicators (SAMIs), which are calculated by measuring the deviation of a particular city's performance on a particular indicator from the national power law trend line for that indicator (Bettencourt 2010). By using scale to compare a city's performance to its expected performance, the non-linear effects of population size are accounted for. This provides a more useful measurement of a city's performance than comparing its performance to the national average, which assumes a linear scaling of the indicator with population size (ibid). Ranking cities based on scale-adjusted performance may be useful in comparing like cities; It allows a way of grouping cities based on similarities in scale-adjusted performance and for making meaningful measurements of the impact of local policy and circumstance (ibid). For example, Juneau, AK, San Jose-Sunnyvale-Santa Clara, CA and Bridgeport-Stamford-Norwalk, CT all have similarly high SAMI values, even though their populations range from 31,275 to 1,836,911.

Four Principles of Urban Scaling

There are four basic principles that have been outlined by Luis Bettencourt et al. of the Santa Fe Institute, which help to explain why scaling relationships occur with such regularity, and which are implicit in the models of urban scaling. These are:

1. Cities develop for the purpose of fostering and enhancing social interactions and social mixing. As such, they "develop so that citizens can, in theory, meet anybody else in the city" (Bettencourt 2013 B. 2). Thus, cities can be imagined as a "web of social interactions embedded in space" (ibid, 2). For a city to maximize its potential, it is also assumed that the minimum resources available to each person match the cost of reaching any point in the city (Bettencourt 2013A).

2. "Urban mobility is essential for mixing, but it comes at a cost" (Bettencourt 2013 B, 2). This assumes that social interactions take place in physical space, and that social networks require the ability of their members to share space with each other to form and to function. Therefore, the cost of movement, necessary for interaction, is a critical determinant of any social network. Most of urban structure "can be explained by the very simple and universal desire for the best achievable balance between income and commuting cost" (Louf 2014, 8).

3. "City infrastructural networks grow incrementally, and this growth is decentralized because it arises locally from an adaptation to human social needs rather than from a central master plan" (Bettencourt 2013 B, 2). Based on empirical observations of US road networks and of the road networks of over 3600 cities around the world, incremental growth leads to sublinear scaling of infrastructure. The population times average distance between individuals is approximately equal to the average length of the infrastructure network per capita: $d = n^{-1/2} = (A/N)^{1/2}$. The total infrastructure length per capita can also be expressed as a function of the baseline area and population: $a^{1/2}N^{5/6}$ (Bettencourt 2013 A). Christopher Alexander found that an urban form exhibiting a cohesive sense of wholeness, emphasizing pedestrian accessibility and aesthetic coherence could be achieved through a process of incremental planning, through a process by which each phase of development followed a few simple rules, rather than a master plan (1987).

4. "Human effort is bounded" (ibid). The amount of energy people will expend in a day to move around and to build and maintain their social networks is limited. Thus, by crowding people together, cities make it possible for people to have a greater number and diversity of possible social interactions without any additional effort (ibid). To build this into a model, Bettencourt assumes that there is a sort of conservation law in which the density of roads times GDP divided by the population density = 0 (so, $dG/dN = 0$, where G = is equal to GDP times road volume per capita, and G is independent of N) (Bettencourt 2013A). Since GDP scales superlinearly and road volume scales sublinearly with population size to approximately the same degree, the ideal ratio G:N should remain constant, and a significant deviation would likely imply an inefficient transportation network or insufficient density to keep the population mixed. Bettencourt (2013A) presents a theory that there is an optimal value for G around which cities fluctuate. A city that deviates too far from the optimal value, Bettencourt predicts would cease to hold itself together as a city.

The impact of economic success on urban form in a city would be expected take the form of geographical expansion, such that the increase in transportation costs is proportional to the increase in the average transportation budget. The spreading out of the population incurs new costs, including increased commute times, distances and congestion, which are, theoretically, proportional to increases in wealth. Cookson (2017) estimates that 9% of travel time in OECD countries is spent standing in traffic). This is modeled by adopting equations used to describe the dissipation of energy in electrical networks. Increased distance between people, increased transportation costs or other barriers to mixing incur an energetic cost that dissipates the energy required to keep the population mixed. If dissipation exceeds travel budgets, the nature of the city or network should be expected to adapt or change. The spreading out of the population as a result of decreased transportation costs at one time might lead to the development of various centers if transportation costs were to rise at another time, or the city could be held together with new adaptations such as more efficient transportation or increased density (ibid). Transportation costs are approximated by assuming that the cost of transportation is approximately $(\varepsilon A^{1/2}N)/N$, where ε is the average energy/time of transportation, A is the network area of the city and N the population (Bettencourt 2013A).

Social Networks and Travel Costs

"The number of interactions in a city that the average person maintains scales inversely to the way the degree of infrastructure scales with size" West writes, connecting social networks and socioeconomic output to infrastructural efficiency; "the degree to which [the scaling exponent of] social interactions exceeds one, matches the degree to which infrastructure is less than one" (West, 321).

To model the relationship between socioeconomic output and the spatial structure of a city, Bettencourt (2013) uses the variable, a_0, as a measurement of the average distance traveled by people in a city, and as a variable in determining the cost of travel and the efficiency with which a city can keep its population mixed, and argues that the socioeconomic outputs of a city are a byproduct of social mixing in the population. Bettencourt is not alone in claiming that travel distances and travel times within a population are highly regular, but the claim that the way average travel patterns relate to the spatial dimensions of a city are the underlying cause of urban scaling implies that high and low performing cities should also exhibit different spatial patterns.

One thing this thesis will examine is the role of spatial patterns that should impact travel costs and the mixing of an urban population in high and low performing cities.

The average number of social interactions per person can be derived as the product of the volume of public space spanned by the movement of people, goods and information in the city (a_0l) times population density ($n = N/A$) (ibid). And the total social output of a city can be expressed as the product of the total number of social interactions and their outcome ($\bar{g}$), which can lead to the formulation of the average cost per person to mix throughout the city (G), with $G = \bar{g}a_0l$, which measures social output per capita times network area per capita. G relates to the total social output of a city, Y, in terms of the formula, $Y=G*N^2/A$ (ibid). Because a_0, the network area, depends on how far the average person in the city travels, so the physical size of a city is dependent on transportation costs and is independent of G (see principle four above). While there is an optimal "efficiency" for cities, there is no optimal size (Bettencourt 2013A).

A city is held together by the ability of its citizens to, theoretically, be able to mix with people in any other part of the city (principle two, above). This assumption requires that the cost of transportation to mix throughout the city must be covered by the average individual's budget (ibid). So the larger the city, the larger the amount of energy necessary to hold the city together as a mixed space, and the greater the rate of dissipation of that energy. The relationship between energy used in transportation, the area of the city and the population, shows that as transportation technology makes transportation faster and cheaper, the size of a city grows as a function of transportation costs. This can be observed with patterns of urban sprawl across the world that have accompanied rising levels of wealth and decreasing transportation costs (Bettencourt 2013A). Bettencourt argues that "Low G cities such as Brownsville, TX or Riverside, CA would benefit from increased mobility or density. High G cities, Bridgeport, CT are economically and infrastructurally developed, but would benefit from more compact living or increase in transportation energy efficiency" (1441).

Urban population densities in the US have fallen by one-third over the last half-century (Clark 2016). Falling densities have been "accelerated by growth in real per capita income and declining unit (e.g., per mile) transportation costs as households seek to consume more housing and locate farther away from the city center" (TRB 2009), a finding consistent with principle two. West and Bettencourt relate the role of energy used in a city to the city's "metabolism", but the analogy is criticized by Shalizi (2011), Strano (2016) and others.

Statistically, the number of *potential* social connections grows with population as the square of the population (Batty 2013), which is much faster than the rate of growth observed in social network sizes as a function of population size (Schlaepfer 2014). A study by Schlaepfer (2014) attempted to analyze what effect city size has on social networks. The study looked at metadata from 1.6 million cell phones in Portugal and 24 million landlines in the UK, and showed that the total number of reciprocal contacts (a number was contacted by and contacted the same number) and activity per capita do grow superlinearly with city size with an exponent of around 1.12 in Portugal (1.05 in the UK) (Schlaepfer 2014). This meant that the average person in Lisbon (population over 500,000) had about twice as many reciprocated contacts as a person in Lixa (population 4000) (ibid).

Although this study showed that people in larger cities were more socially connected, the "probability that an individual's contacts are also connected with each other remains largely constant, which indicates that individuals tend to form tight-knit communities in both small towns and large cities "(ibid). This study both adds nuance to the common perception of increased social alienation in larger cities. A larger social network should not be interpreted as more friends – this tends to remain relatively constant regardless of where a person is living (with 3-5 close friends, 12-15 in a 'sympathy group', ~50 significant people and a social world of around 150 people we know personally, whom we can trust, and for whom we feel some emotional affinity (Dunbar 2010)). Dunbar's assertion that the size of social groups remains constant can be observed in the size of more tightly knit groups, but the number of social contacts does not show the same stability and are influenced by city size. The increasing number of contacts, then, implies that in larger cities, people interact with a greater number and diversity of other people, which makes productive social interactions more likely.

If walking were the only mode of transport, the radius within which a social network would likely form would be within the space each individual can "afford" to reach regularly, that is, what a person could dedicate to walking in a day. Modern transportation technology affords a greater mobility of people and services and leads to a larger geographical representation of the social network. One expression of this was found by Marchetti (1995) in the historical distribution of villages in Greece, which reflects a regional network held together because the distance between villages is within the average person's daily transportation "budget," that is, villages are located at regular intervals about an hour's walk from each other, each containing a surrounding radius of farmland of about 2.5km. Among walled cities built before 1800, none had a walled radius of larger than 2.5km, implying that any citizen could theoretically afford to meet any other citizen at a cost of about an hour's walk per day. Likewise, all ancient empires had maximum radii equivalent to a 15-day journey. (West 2017). Implying that beyond a month between messages, the social networks required to maintain political power fall apart. This, Speck (2012) argues is what was at the root of the division of the Roman Empire — when Rome lost control of the seas, the eastern half of the empire was beyond a 30-day round-trip and the empire split apart.

Scaling theory connects travel costs and speeds with socioeconomic outputs and spatial patterns. It stipulates that if the size of the city exceeds that at which all of its citizens can afford to reach every other part, the city should begin to split apart and develop new centers. The theory also speculates that the average radius of a city should reflect the predominant speed of travel, and today, modern large cities have an average radius of about 40 miles, representing an analogous travel cost to walking across cities before the 18^{th} century (about an hour a day) (Speck 2012). Data from the Department of Transportation suggest that total daily travel times for drivers in the US is closer to 85 minutes (BTS 2017). This is a worrying number, because Americans are already emitting about ten times more carbon into the air than the UN benchmark for achieving the 2-degree Celsius climate change goal set in the Paris agreement (Prakash 2017).

The connection between mobility and spatial patterns was further investigated by Holtzclaw, in a 1994 study for the California Department of Transportation on the impact of transit and the built environment on vehicle miles traveled (VMT), or the number of miles traveled by personal motorized vehicles, car ownership and transportation mode choice. He found that density had a significant impact on VMT, followed by the accessibility to transit. Holtzclaw found that VMT related to density and a transit accessibility index with exponents of -0.25 and -0.076, respectively (Holtzclaw 1994). Another author, Ewing (2010), examining the impact of various urban design factors, found that Intersection density and street connectivity have an even greater effect on VMT than employment density. "One startling observation from this analysis" Holtzclaw writes "is how poorly household income predicts auto ownership or VMT" (op. cit. 37). By excluding the extremes of the income scale, no correlation was found, which means that large-scale spatial patterns that result from changes in average transportation costs are not being driven by certain classes.

Negative congestion externalities and positive externalities associated with the agglomeration of employment tend to be of similar magnitude, though when transportation costs rise, employment tends to become more spread out, which weakens the agglomeration advantage of clustering, which can attenuate over very short distances (Brinkman 2016). When transportation costs rise, it is similar in economic terms to a decline in technology, and all aggregate measures of economic activity tend to decline in response, including production, employment, rents and wages (ibid).

In a study of 98 cities, there was a strong trend of new job growth taking place outside of the urban core (TRB 2009), indicating that travel or congestion costs are high enough to be impacting the urban form of many cities, putting increasing pressure on dispersed development patterns. This will likely equate to less urban open space, and in forested areas, less urban tree canopy.

Longer commute times have been associated with higher risk of stress-related illness, lower levels of happiness and greater levels of pollution. While better health is generally associated with higher incomes, higher paying jobs tend to be located in areas of greater congestion (TRB 2009). Therefore, cities with high income and low commute times may be expected to show a lower rate of stress-related illness, while low incomes and high commute times may be expected to show higher levels of stress-related illness.

Density and mixed uses have been recognized by the United Nations' New Urban Agenda as important contributing factors to urban productivity, driving economic growth through the efficient use of resources, and improving the sustainability of small and medium-sized businesses (World Economic Forum 2016). Further emphasizing the impact of physical structure on social connectivity, the World Economic Forum (WEF) states that "dense urban structures promote an easy exchange of ideas, goods and services [and] are considered good for business, innovation, arts and culture and are environmentally friendly" (ibid, 33). Conversely, Speck (2012) argues that a lack of housing density can contribute to the struggling of businesses in other economic sectors. Using density to improve services while saving on resources is considered a key aspect of resilience by 100 Resilient Cities (2017).

Other Perspectives on Social Networks

The importance of social networks as drivers of socioeconomic output is supported by urban economic theory. Storper (2013), explains that networks are the “basis of the functioning of the formal institutions of metropolitan government and governance, and they themselves amount to informal institutions that carry out the mobilization and transportation of skills and capacities across different domains of the economy as well as time periods” (103). Mills and Zhang (2013) have demonstrated that the strength of social networks is very important in low-income people’s ability to avoid hardship. The strength of social networks was measured on people’s assessment of how likely they felt that family, friends, or others would be able to support them in case of hardship.

Social networks are not easy to measure directly, but social institutions may serve as a meaningful proxy. Roman (2004) shows that the presence, capacity, and proximity of social institutions are closely linked with “increased levels of public safety, reduced violence, improved supervision of children, reductions in physical decay, disorder and fear of crime, as well as increased participation in community organizations as well as higher levels of physical health” (I).

Storper (2013), instead of focusing on the ability of each citizen to fully mix to explain why cities scale, looks at how the specialization of industrial or business sectors within a city facilitate the ability of social networks within specialized fields ('insider groups') to form more productive relationships. For Storper, the linkages between firms, workers and skilled people explain the wealth of cities, not population size. Storper’s work supports the assumption used in scaling theory that social networks are a fundamental driver of economic activity. There is a "minimum threshold of colocation of firms to generate critical movement of people that in turn raises the benefits of face-to-face contact" and generates a positive feedback for further economic growth and specialization (ibid, 181-182).

Transportation Costs and Sustainability

Several authors have been tempted to conclude that bigger cities are greener, which relates to principle four of scaling, arguing that greater efficiency in infrastructure costs creates decreased dependence on driving and increased energetic efficiency. However, this is a matter of some contention. Geoffrey West (2017) explains that the sublinear scaling of infrastructure leads to greater efficiency with size and that “on average the bigger the city, the greener it is and the smaller its per capita carbon footprint” (see also: Batty 2013). Doug Saunders (2011), writing on the emergence of urban patterns resulting from rural migration to large cities, argues that cities “[reduce] ecological damage and carbon emissions by decreasing distances and increasing shared technologies: Cities, in the words of one major study, ‘provide an opportunity to mitigate or even reverse the impact of global climate change as they provide the economies of scale that reduce per capita costs and demand for resources’” (23).

The argument that bigger cities are greener has been challenged based on how emissions are measured; if the per capita carbon footprint is measured only in the context of what is emitted directly by individuals within the city, urban areas, account for just over half the world’s population, and only 30-40% of global CO_2 emissions. Fragkias (2013) demonstrates that when measured using consumption-based accounting, a measurement that accounts for CO_2 embodied in the products that are imported and consumed in urban areas, the urban share of global CO_2 production rises above 60%, with the majority being produced by a few wealthy cities. The same study found that consumption-based CO_2 emissions among US cities scale approximately linearly with city size, implying that population size itself has little to do with how “green” a city is. “The ratio of outputs is a function of the proportion of population sizes, but not of [population size]” (ibid e64727).

Other factors, such as wealth and density were also analyzed to see if they could account for the variation in CO_2 outputs. Density showed, with a strong R^2, an inverse scaling of CO_2 with population of about 17%, while emissions showed a positive correlation with wealth (with a weak R^2) such that a 1% increase in wealth corresponded with a 0.36% increase in CO_2 emissions. When measured in terms of CO_2 emitted per dollar of economic activity, increasing density decreased CO_2 per dollar of economic activity (TRB 2009). Because density decreases reliance on personal vehicles, and personal vehicles account for about 20% of CO_2 emissions (TRB 2009), promoting density would not only improve energy efficiency, but would likely have economic benefits as well by helping cities avoid “diminishing economies of agglomeration” attributed to sprawl (World Economic Forum 2016, 3). Another way to analyze the relationship between density and CO_2 emissions is to look at the degree to which growth has taken place in the dense urban cores and in the less dense peripheries to test if there is a correlation with their changes in

CO_2 outputs. UN Habitat (2014) has shown that there is a strong correlation between the degree to which cities have contained growth to their dense cores and their CO_2 footprint. The positive correlation of CO_2 with wealth was only significant for a portion of the years studied after 2005 (Fragkias 2013). By weighting all three factors, population, density, and wealth with their respective scaling with the formula,

$$\ln(CO_2) = 1.685 + 1.028\ln(\text{population}) - 0.172\ln(\text{density}) + 0.364\ln(\text{per capita income}),$$

the result is a scaling exponent for CO_2 with population between (0.971, 1.084), indistinguishable from one (ibid). Fragkias speculates that the apparently linear scaling of CO_2 output, “could be that the compact spatial form of cities is associated with gains in energy efficiencies but that these gains are offset by the increased consumption facilitated by higher productivity levels induced by larger urban agglomerations” (ibid, 3).

Other studies have found that CO_2 and GHG emissions scale superlinearly with city size when a city is defined as connected urban space (rather than a density-based definition), and the relationship between city size and its CO_2 and GHG emissions seem to present different scaling regimes in developed and developing countries, with developed countries showing a linear or slightly sublinear relationship and developing countries showing a superlinear relationship (Oliviera 2014).

Energy consumption and pollution, it appears, are not related to how socioeconomic or infrastructural indicators in cities scale with population size. These differences have more to do with the ways in which manufacturing, international trade, farming and forestry practices, energy production and other factors define the disparity between the carbon footprints of large and small cities and in developing and developed economies.

Challenges to Scaling Theory

An important challenge to scaling theory comes from studies that show that factors other than population are both more useful and more accurate in describing socioeconomic output such as productivity and GMP. Though scaling theory supposes population to be a proxy for other socioeconomic interactions, demonstrations that population is less determinant than other factors both challenge the theory and offer different kinds of insight into scaling that may be useful to planners.

Cervero (2000) found that city population was not a “reasonably significant predictor” of economic output (1656). This study found that employment density, location quotients for the finance, insurance, and real estate sectors, and commuting speeds and distances, accounted for 50% of the variation in worker productivity. The only two statistically significant variables in predicting worker productivity at the metropolitan level were employment density and demographics (1666), not population. Cervero (2000) also found a correlation between congestion and productivity: congestion at the intra-metropolitan level corresponded with increased worker productivity, while at the inter-metropolitan level, the opposite appeared to be true.

Prud’homme and Chang-Woon (1999) found that productivity scales with travel speeds within a city such that a 10% increase in speed corresponds with a 2.9% increase in productivity. Labor productivity, they found, could be largely explained by how many jobs could be reached in a location given a commuting time. Though the variation in travel speeds does not generally have an impact on smaller towns where all locations tend to be accessible within a reasonable commute time, the impact was significant in larger cities. They found that productivity scaled as a function of labor market size (with slightly different scaling exponents depending on the travel time). In three Korean cities with a 60-minute radius, a 10% increase in labor market size led to a 2.4% increase in productivity; in France, the same size increase led to a 1.8% increase within a 20-minute commute and 1.3% within a 30-minute commute. Prud’homme and Chang-Woon conclude that “containing sprawl and improving transportat [sic] speed in a city both increase the productivity and therefore the output of a city” (1857).

Although productivity appears to scale positively with increased urban travel speeds, commute times to work scale negatively with population size according to Levinson (2012). Levinson finds that with a 10% increase in population, commute times to work increase by 1.14%. Levinson attributes this to increased traffic and less road capacity per capita: roadways per capita decreasing with an exponent of 0.21, while total roadways and street density increase with population with exponents of 0.67 and 0.099 respectively (ibid).

Shalizi (2011), criticizes the theory of scaling by arguing that the results of studies by Bettencourt et al. are not statistically valid on a number of accounts. One of his strongest criticisms is that despite the remarkably strong fit of a power law least squares regression showing the scaling of GMP to population size, with its small exponent of 1.2 and an R^2 of 0.96, it is insignificantly better than a linear fit to the same data with an R^2 of 0.94 (Shalizi 2011).

Furthermore, Shalizi argues that controlling for the size of four high-value-added industries in cities generates a better predictive framework than population, effectively arguing that GMP scales better as a function of key industrial sectors than of population (ibid, see also Storper 2013). Accounting for urban productivity through industrial sectors falls more in line with previous thinking in urban economics (Cervero 2000).

A study on scaling comparing cities between 2005 and 2010 in western Europe to cities that were part of the Soviet Union finds that wealthier cities (defined as Large Urban Areas, a definition most similar to MSAs in the US, in order to make the study comparable to studies done in the US) in western Europe scale linearly with size (exponent 0.05), while poorer, post-Soviet cities scale superlinearly with size (exponent 1.25 in 2005 to 1.42 in 2010) (Strano 2016). Strano suggests that superlinear scaling may represent a phenomenon of economies in transition rather than a physical quality of stable and mature economic systems. Superlinear scaling, he suggests, "represents an unbalanced situation of rapid growth of large cities and economic segregation of smaller ones, which makes redistribution of income increasingly difficult" (ibid, 6). Superlinear scaling may represent not only a sort of transitional economy, but also uneven growth, with rapid population growth going disproportionately to the largest cities. Strano prefers the theory that scaling is a reflection of macroeconomic, regional processes to the criticism of other authors, such as Cottineau, who argue that city form and structure determine the nature of scaling, because their work "implicitly accept[s] that scaling does occur within a metropolitan boundary" (ibid, 2).

Olpadwala (2017) argues that the studies that have been done should not be relied upon as evidence of a universal theory of urbanism. The majority of the studies have been conducted in culturally and economically similar countries, and many use sample sizes that should call into question the universality of the claims made. Ultimately, he claims, trying to derive universal laws that describe patterns of collective human behavior is limited by the inherent unpredictability of the free will and behavior of individuals.

At one point, he argues that the context of capitalism clouds the validity of the universality of the claim that empirical observations of cities offer insight into the nature of cities that transcends the larger economic framework, writing that "to investigate urbanization in capitalist contexts without explicit attention to its rules of engagement is to restrict a priori the power of the inquiry" (45). Although writers on scaling claim only to describe the relationship between size and these aggregate measures within a system of cities implies cities within a particular economic system. Claims that scaling may hold some potential as a starting point for a scientific theory of cities are based on the appearance of similar scaling regimes in different countries, and although, except, perhaps, for China, the countries studied are capitalist, their diversity still lends their findings the aura of transcending economic arrangements.

Health Indicators and Socioeconomic Performance

The literature below was found as a result of searches for articles linking health outcomes in cities to socioeconomic performance or status. Even when strong correlations between socioeconomic status and health or environmental outcomes is the topic of investigation, economic factors will also be considered as a proxy term to describe behaviors, attitudes and access to other resources or services. One way that health outcomes may relate to scaling theory is through findings that the rate of social participation is considered one of the primary mechanisms explaining the relationship between health outcomes and economic status (Cheng 2012).

That there is a connection between health and the economic status of individuals and neighborhoods has been well established. Robinette (2017) shows that for every $10,000 increase in the average income of a neighborhood, residents are significantly more likely to experience better health. An increase in the median household income of a county of one standard deviation corresponded with a 13% decrease in premature mortality, and an 8.4% increase in people with a college degree was associated with an 18% decrease in mortality in low income counties and a 12% decrease in high income counties (Cheng 2012). Income and educational attainment have been cited as significant determinants of Chronic Obstructive Pulmonary Disease (COPD), such that people in the lowest socioeconomic strata are about twice as likely to be diagnosed (Gershon 2012, Kanervisto 2011). Even when rates of COPD were corrected for smoking habits, low SES was still found to significantly increase the risk factor (Bakke 1995, Kanervisto 2011)). The association between COPD and low socioeconomic status cannot be entirely explained through class-based differences in occupational exposure (Prescott 1999) although exposure to biomass smoke was considered a significant contributor to risk for COPD in one small study (Business Insights 2017). Though other types of illness are influenced by socioeconomic status, COPD is more strongly correlated than other diseases (Prescott 1999, Gershon 2012). In a study of non-smoking males, there was a 400ml difference in lung capacity between the highest and lowest social classes, and lung-related illness showed a particular sensitivity to levels of education (Prescott 1999). Having only basic educational attainment has been shown to increase the risk factor for COPD with an odds ratio of 1.8, and the association between COPD and both poverty and shorter education remains significant even when adjusted for age, gender, BMI and smoking history (Kanervisto 2011).

Cancer has been shown to be influenced by level of education, and with economic status in men but not in women (Dalton 2008). This was only found to be true for certain types of cancer, including mouth, pharynx, larynx, oesophagus, stomach, lung, kidney and cervix (ibid). The data analyzed in this study groups all types of cancer together, so will be less sensitive to this nuance. Dalton also found an urban-rural gradient in cancer prevalence, with higher rates in capital cities and lower rates in peripheral, rural areas.

Considering the relationship between asthma and air pollution, much research supports the connection between the severity of symptoms and levels of air pollution, but there is some disagreement on whether pollutants are causative of the onset of asthma. Nitrogen dioxide is most strongly associated with the proximity to traffic, and local levels of NO_2 show a small but significant association with an increased prevalence of asthma (Graziella 2014). Several other studies have shown that the severity and prevalence of asthma is linked to socioeconomic status, and that poverty likely explains what has, in the past, been described in terms of racial or ethnic differences in susceptibility to asthma (Rona 2000). While people of African American or Puerto Rican descent are more likely to be diagnosed with asthma, (White non-Hispanic males and females under 18 years old have an asthma rate of 14.1% and 10.1%, while the rate for Black non-Hispanic boys and girls was 22.7% and 14.8%) the effect of poverty and living in poor neighborhoods is strong and historical and current patterns of segregation and unequal economic opportunities explain different outcomes for different racial groups (DePriest 2017). According to a 2018 study, as a result of discriminatory lending practices, white and black homeownership rates are now more unequal than during the Jim Crow era (Glantz), making it likely that outcomes associated with poverty and poverty-related stress would show great disparities in how they affect people of different races. Diets may also be impacted by demographics, Block (2004) found that predominantly black neighborhoods in New Orleans have a significantly higher concentration of fast-food restaurants (2.4/square mile) than predominantly white neighborhoods (1.5/square mile).

In 2011, Galea calculated that 176,000 premature deaths could be attributed to racial segregation (2011).

International studies have shown that ethnicities that are more economically disadvantaged have higher rates of asthma in several different countries, suggesting that genetic predisposition is less influential of a factor than social discrimination (Rona 2000). Another study has shown that people of color are more likely to be exposed to unhealthy levels of NO_2 than whites in the United States, showing that race was more significant than income in predicting exposure. In 2000 and 2010, black and Hispanic people experienced a 40% and 37% higher rate of NO_2 exposure than whites, even while annual concentrations of NO_2 have dropped over the same time frame (Milman 2017). Like the continuation of redlining in lending practices, the worst polluters still tend to be located near minority neighborhoods (ibid).

Historically, before the inflammatory nature of asthma was well understood, it was considered a psychogenic disease related to anxiety (Wright et al. 2018). Today, stress is still understood as having a significant influence on asthma and the severity of symptoms (ibid). Given the debate on whether there is a significant relationship between air quality and prevalence of asthma, it is interesting to note that the level of inequality as measured by the GINI coefficient had the strongest correlation with asthma rates (as measured by the BRFSS, but not the 500 Cities data) of any socioeconomic factor considered in this study. The association between asthma and the level of inequality in a city would support evidence that stress plays a significant role in the prevalence of asthma, and is exacerbated by the increased exposure of poor neighborhoods to higher levels of pollution. The argument that psychological stress influences asthma is based on the fact that many of the same mechanisms produced by stress also are involved in the activation of asthma, including the neuropeptides, adrenaline and noradrenaline and catecholamine released as a result of stress are also known to contribute to the inflammation of the airways (Wright et al. 2018).

One study examined the correlation between the perceived threat of violence and asthma rates; even for people living in the same neighborhood, those who perceived the threat of violent crime to be higher were more likely to be diagnosed with asthma (Camacho-Rivera et al. 2014). Thus, it is not only the crime rate, but the perceived threat which influences the onset of asthma. Health and mortality rates have also been shown to be sensitive to negative social interactions such as discrimination, potentially exacerbated by decreased access to resources (Galea 2011).

“Social perceptions of inequality or stress resulting from the acceptance of the social stigma of inferiority may have pathogenic consequences that adversely affect health” (Cheng 2012), which likely exacerbates observed differences in outcomes between social classes and racial groups. An international study looking into the question of why ethnicity appears to be an important factor in asthma prevalence found that in each country studied, ethnic groups that were most socially disadvantaged and discriminated against had higher rates of asthma in each country (Rona 2000). At a state level of analysis, after adjusting for other social variables, it was found that state median household income had a significant effect on high blood pressure, with adults in lower income states and states with higher rates of poverty were associated with higher rates of high blood pressure (Fan 2015). In fact, independent of individual socioeconomic status, living in a state with low median household incomes and a high percent of poverty increased the odds of high blood pressure in residents (ibid), and people living in a county in the bottom quartile have a high blood pressure rate 9% higher than those in the most affluent counties (Shaw 2016).

House (1989) writes on the connection between health risk factors and social networks. His work shows that the level of social connectivity has a significant influence on health outcomes, though the most significant impact is noticeable between those who are very isolated and those moderately isolated. There is little difference between health outcomes of people moderately and highly socially integrated, and men are more strongly affected than women. An increased risk of coronary heart disease was also noted among socially isolated men by Eng (2002). Galea et al. (2011) explain the link between social connectivity and health in how it influences health behaviors and reduces stress. The result of social interaction is an impact on immune function, cardiovascular activity, and the progression of existing disease (ibid).

In this study, binge drinking generally correlated with socioeconomic data in the opposite direction from other health indicators, and high rates of binge drinking were significantly and negatively correlated performance on all other health indicators except cancer, which showed a negative, but insignificant correlation. Binge drinking is considered a risk factor by the CDC for its association with “alcohol poisoning, hypertension, acute myocardial infarction, sexually transmitted infections, unintended pregnancy, fetal alcohol syndrome, sudden infant death syndrome, suicide, interpersonal violence, and motor vehicle crashes” (Cremeens 2004). Binge drinking has been found to be more prevalent among wealthier individuals, but more frequent among poorer individuals, and men who had a negative perception of their neighborhood were eight times more likely to be binge drinkers (Jintarin 2015). Neighborhood poverty is associated with a greater risk of alcohol-related problems, but that risk is significantly greater when a neighborhood also shows signs of disorganization (graffiti, litter, perception of crime, lack of participation in neighborhood/social organizations) (Eckart Washington 2017).

Because of the variables considered in this study are unconnected with the reasons why binge drinking is considered a risk, the true impact of binge drinking cannot be taken into account. However, hypertension, or high blood pressure, which is linked to binge drinking, is considered the leading cause of chronic disease and premature death in the United States (Fan et al. 2015), and is measured by the 500 Cities project, but it does not show a positive correlation with binge drinking. Rates of interpersonal violence attributed to binge drinking, but violent crimes including murder and aggravated assault correlated negatively with binge drinking (though aggravated assault was a very weak correlation, -0.25 and domestic violence was not a category of crime in the FBI data used).

Linking the potential impacts on health from socioeconomic status to tree cover is the theory that one of the reasons for worse health outcomes in individuals and communities with lower socioeconomic status is less access to parks and recreation facilities, as well as commercial areas, schools, jobs and safety as well as being less likely to live in neighborhoods that support active transportation to work (Shaw 2016).

Environmental indicators and socioeconomic Performance

Urban Tree Cover

A number of ecological services are provided by urban tree cover and would be expected to have a positive impact on public health as a result of these services. Services include mitigating air pollution and particulate matter, reducing urban noise, improving infiltration of rainwater runoff and improving water quality in streams and rivers, as well as reducing urban temperatures and moderating the urban heat island effect through evapotranspiration. Although generally considered an amenity, trees also provide some disservices, including increasing pollen levels, promoting invasive species, hosting pathogens or pests, inhibiting human mobility and safety, increasing water usage in places where water is scarce and even increasing greenhouse gas emissions (particularly N_20 releases from denitrification) (Pataki 2011). Tree cover has been shown to have a positive effect on people's health, but it is likely not a direct effect of a reduction in levels of pollution or other ecosystem services (ibid).

Urban tree cover, not surprisingly, shows great variation between naturally forested regions (31%), grasslands (19%) and deserts (10%) (Nowak 1996). In all three environs, land use had a significant impact on canopy coverage, with residential and park land having the highest rate of tree canopy. Residential areas in forested cities having a tree canopy coverage of 53%, grassland cities with 43% and desert cities with 33%, and with a negative correlation (-0.64) between density and tree cover (ibid).

Within cities, urban tree cover has been shown to be sensitive to neighborhood economic conditions with poorer neighborhoods frequently having less tree canopy coverage (Gerrish and Watkins 2018, Schwarz 2015). Though the magnitude of the effect of income on tree cover varies widely between studies based on methodology and how various mitigating factors were treated, especially spatial autocorrelation, Gerrish and Watkins (2018) found in a review of literature that the level of inequity in tree cover is substantial. Schwarz et al. (2015) found that an increase in median household income of $1000 correlated with an increase in tree cover of 0.05-0.20% increase in tree cover in most cities studied. Even so, two cities in the study (of seven cities) showed negative correlations between tree cover and income (ibid). Race showed a significant correlation with tree canopy in a study of seven cities in a bivariate model, but in a multivariate model, race became insignificant or minimal, and varied between cities, and was not significant in spatially autocorrelation models (ibid).

Tree canopy is correlated with a range of other indicators, including, heat island effects, surface water runoff, particulate matter, carbon monoxide levels, Sulphur dioxides and nitrogen oxides (Hirabayashi 2016), all of which can have negative impacts on human health. However, many of these have not been adequately quantified in empirical studies of urbanized areas (Pataki 2011). The cooling effect is highly correlated to the specific location of trees in relation to infrastructure and varies widely by species of tree, while the effect on local air quality has few empirical studies, but has been estimated by Pataki to be extremely low (ibid).[2] Another study found that tree cover near industrial areas had the potential to reduce air pollutant emissions by as much as 63% in adjacent neighborhoods (Rao et al. 2004). Although the ability for vegetation to

[2] In a study which modeled the effect of doubling the tree planting density in the Salt Lake Valley, only an estimated 0.2% of total CO_2 would be offset over 50 years (Pataki 2011).

remove a substantial amount of particulate matter is likely limited, and claims that trees and vegetation significantly improve urban air quality are not well supported by empirical evidence, Pataki (2011) found less than 2% of PM10 were removed in urban areas with 26% tree cover. Urban trees may indirectly reduce O_3 production by reducing heat island effect, although this has not been well quantified (ibid). Despite the positive impact on air quality of urban tree cover, Pilat et al. (2012) found no relationship between overall urban vegetation and air pollution.

Tree cover has also been shown to mitigate stress related illnesses such as high blood pressure and mental health. The effect of trees has been shown to be significantly stronger among youth, the elderly, people with lower socioeconomic status (Maas 2018) and men, and can vary over the course of a person's life (Astell-Burt 2014). One mechanism which helps to explain this is that exposure to green space reduces cortisol levels (Gaffield 2003). Tsai (2018) found a significant correlation between depression and access to urban green space, finding that when green space is distributed more evenly through a city, it has a greater impact reducing levels of depression. Trees impact on stress levels likely helps explain findings that greater urban tree cover is related to a decrease in the prevalence of asthma (Lovesai, G.S. et al 2008, Ulmer 2015). Tsai (2018), McCormick (2017), Ulmer (2015) and Holtan (2015) all show that green spaces may have a positive impact on social engagement, which allows green space to show a positive impact on mental distress even after adjusting for income levels, which show a strong correlation with improved mental health. This would support the social network aspect of scaling theory, and should lead to the finding that increased tree cover in cities should correspond with most of the socioeconomic indicators with strong superlinear scaling with population size.

Because of the relationship between tree cover and the amount of impervious surfaces in a city, tree cover has also been shown to correlate with water quality. Generally, the increase in impervious surfaces has had a negative impact on water quality, while tree cover has a positive impact (Nowak 2012).

In summer months, tree cover could also impact health and air quality in indirect ways: by reducing the heat island effect, less energy will be used for cooling, potentially impacting air quality not only through what the trees absorb, but also to what they make unnecessary to emit. Through mitigating the effects of urban heat islands, which have been linked to increased levels of ground level ozone, respiratory symptoms and increased heat stroke, increased tree cover should also correlate with certain health outcomes (TRB 2009). Urban trees can reduce ambient summer temperatures by 2-9°F through evapotranspiration. Tree-shaded areas can be 20-45°F cooler than unshaded areas (EPA 2018), and strategic planting of trees can lead to reduced energy demands for cooling by up to 40% (Roseland 2012). However, the benefits of evapotranspiration come at the cost of irrigating trees in drier climates, which can dampen the positive effect when the cost of water is high -- in California, for example, irrigation accounts for 30-70% of urban water use (Pataki 2011).

Urban trees have also been shown to have an impact on mental health and by providing spaces to enjoy being outside, trees have also been shown to improve social capital (ibid) although the presence or absence of parks in a neighborhood had no significant association with social capital (Holtan 2015). If trees increase connections between people, scaling theory tells us that one result should be increased economic productivity, yet there are no clear connections between tree canopy coverage and economic activity or productivity.

The correlation between urban tree cover and income is so strong that there may be some feedback loops the reinforce the correlation. For example, the impact of trees on property values lead high-income neighborhoods to promote and increase tree cover, while low-income neighborhoods may actively oppose tree plantings to prevent gentrification, or out of the inability or unwillingness to afford the cost of upkeep and watering (Schwarz 2015). It is also common for people to perceive an association between urban trees and a higher threat of crime, although an actual connection has not been shown, and in some cities, such as Baltimore and Chicago there is a negative correlation between crime and urban tree canopy (Schwarz 2015, Ulmer 2015). In areas with naturally high tree cover, more tree canopy may be a sign of disinvestment, and this partially explains apparent correlations with increased tree canopy in African American neighborhoods in Baltimore (Schwarz 2015)

Methodology

This study involved the gathering and analysis of nationwide data from the sources listed below. The number of cities for which data were available from each source is included. When comparing data across several categories, the number of cities analyzed was that for which data in all categories was available.

Data sources:

CDC: Risk Behavior health risk factors. 185 cities. Asthma, depression, obesity, diabetes, overall health

CDC: 500 Cities (2015 data) - cancer prevalence, chronic obstructive pulmonary disease, diabetes, asthma, high blood pressure, obesity, sleep less than 7 hrs. Medication for high blood pressure, stroke among adults, physical health not good >14days, mental health not good for more than 14 days, binge drinking.

FBI: Crime – violent crime, murder and nonnegligent manslaughter, rape, robbery, aggravated assault, property crime, burglary, larceny and theft, motor vehicle theft.

Census Bureau: 1990 Time Series median household income: cph-1-124 https://www.census.gov/data/tables/time-series/dec/cph-series/cph-l/cph-l-124.html

American community Survey: Population, poverty status, GINI coefficient, education, mobility, transportation, all other data

Economic Census: NAICS sector data, GDP, value added in the manufacturing sector, Productivity

Decennial Census 2010: Land area, Shapefile for MMSA boundaries

BEA: Gross Domestic Product by Metropolitan Area; State and Local Government Spending by Metropolitan Area; Federal Civilian Non-Military Spending by Metropolitan Area; Real Income by Metropolitan Area. Contains data on 386 cities

EPA: Air Quality Survey - 532 cities Air quality indicators: Air quality index, Sulphur dioxide, Ozone, Nitrogen dioxide, carbon monoxide, PM2.5 and PM10.

United States Department of Agriculture Forest Service *Northern Research Station Unit 08. (2011) Urban and Community Forests of the Continental United States.* Percent tree canopy within places. Calculated by MSA using ArcGIS.

EPA: National Walkability Index

Socioeconomic data above were collected and analyzed for correlations with health and environmental data. Socioeconomic data were analyzed both as a percent or per capita value and, where applicable as a scale-adjusted value. Scale adjusted metropolitan indicator (SAMI) values are the difference between the observed aggregate or per capita value and the expected value based on the trend line of the power law fit when that indicator (y-axis) is plotted against population (x-axis).

$SAMI = Y_{observed} - Y_0 N^{\beta}$

Where Y_0 is an parameterless intercept, N is the population and β is the scaling exponent.

The last Economic Census, which is a census of all known business entities conducted every five years, was done in 2012. Two measurements of productivity were calculated using Economic Census data, one by dividing GDP by aggregate hours worked, the other by dividing value added in the manufacturing sector by aggregate hours worked in that sector. Wherever possible, 2012 data is used to compare data at a common point in time. 1990 data for median household income are also used to test the impact of historical wealth on tree canopy and impervious surface. Scaling relationships with indicators that have been studied show a great deal of stability over decades (Bettencourt 2010), therefore data from 2012 should scale statistically similarly to current data and should be roughly comparable to data from other nearby years in cases in which 2012 data were not available.

Data are analyzed in two ways: the relationship between socioeconomic performance and health and environmental indicators is of primary interest. The relationship between health and environmental indicators as well as the relationship between tree canopy and air quality is of secondary interest and is also tested.

Center for Disease Control (CDC) data were collected by incorporated cities, not by metropolitan or micropolitan statistical area (MMSA). Therefore, health data should be understood to represent the urban core and not necessarily the whole MMSA. Some CDC data, that from the 500 Cities project, is from 2015, while CDC data from the Behavioral Risk Factors survey (BRFSS) is from 2012. Because scale-adjusted indicators demonstrate a great deal of stability over decades, data from nearby years is used in some cases, and is specified in each case. The error that results from this is minimal: the difference in coefficients of correlation between aligning 2015 CDC data to 2012 ACS data is 0.02 to 0.03 in most cases.

In this analysis, Metropolitan and Micropolitan Statistical Areas were used as the primary unit of analysis because it is the only functional definition of a city for which data are widely available. Scaling theory is based on measuring the influence of population on the aggregate product of the whole web of social and economic interactions that comprise a city, and makes no claim to be able to describe infrastructural or socioeconomic outputs based on political or other definitions of a city. For this reason, census-defined places will not work; cities that are a single socioeconomic unit are comprised of multiple places, and often multiple urban areas. Urban areas are also not suitable, given that they are defined based on density. Although an MMSA is not a precise definition of such a web of social and economic interactions with an urban core, it is as close an approximation as possible for which data are readily available.

Urban areas are a density-based definition, and so these necessarily exclude the urban-rural interface and populations living in less dense areas surrounding cities that are otherwise integrally connected to the social and economic life of the city. Places also create divisions between populations that are otherwise interdependent and that have high degrees of interaction and exchange. Therefore, Metropolitan and Micropolitan Statistical Areas, which are described as the counties containing an urban core and nearby counties that have sufficient interaction with the urban core as measured by traffic patterns, offer the best available geographical representation of a complete web of the social and economic interactions that define a city.

Tree canopy was analyzed using the USFS tree canopy database which contains percent tree canopy for all Census-defined places. Using ArcGIS, places were merged with metropolitan and micropolitan statistical areas and percent canopy from all of the places contained by each MMSA were averaged to define an average percent urban tree canopy for each MMSA. The USFS categorized each place in its dataset within one of sixty-six ecological zones. For the purpose of analyzing statistical patterns, these ecological zones were combined to reclassify the country into six ecological zones to correct for differences in tree canopy as a result of climate or ecological zone. The data were exported to an Excel spreadsheet and analyzed along with socioeconomic and health data.

Methods for each part of this investigation are described below:

Establishing scaling relationships between population size and socioeconomic indicators.

A first step in this analysis is to establish scaling relationships between population size and socioeconomic indicators using 2012 data. This was done using LibreOffice 5.0 spreadsheets and graphing population along the x-axis and socioeconomic data along the y-axis and establishing a power law trend line and R^2 for the fit of the line.

1. Creating scale-adjusted metropolitan indicator (SAMI) values for each indicator by subtracting the observed value of each indicator in each city from the expected value based on power law trend line for the scaling of that indicator.

   Using the fit line established for a given socioeconomic indicator, the deviation from the fit line, or the scale-adjusted metropolitan indicator or SAMI was established by subtracting the observed value from the value predicted using the fit line formula and population.

   $Y_{observed} - (Y_0 * N^{\beta})$

   Where Y_0 is an intercept or scaling constant, N is the population and β is the scaling exponent.

2. Coefficient of Correlation between socioeconomic indicators and health/environmental indicators.

   To test whether SAMIs for socioeconomic indicators or per capita data correlated more strongly with health and environmental indicators, a coefficient of correlation was determined for all SAMIs and per capita data available. As a general guideline, a coefficient of correlation between 0.3 and 0.5 is considered weak, between 0.5 and 0.8 moderate and above 0.8, strong. This was determined using the "=CORREL" spreadsheet function to compare the sets of data.

3. Graphing indicators with higher coefficients of correlation.

4. Identifying cities that rank high or low in multiple areas.

5. Extracting tree cover data using ArcGIS. Data on tree coverage in all Census-defined places were obtained from the USDA Forest Service website. In ArcGIS, places were clipped by MSA, place polygons were merged and intersected with MSA polygons. Tree

cover data were averaged over all Census-defined places within each MSA using the Summary Statistics tool and then exported to an Excel spreadsheet.

7. Median hours worked and travel time to work are two indicators that are expected to reflect particular aspects of scaling theory, and were given special attention to test for whether scaling theory could help to explain the results. Median hours worked relates to productivity and travel time to work is a proxy for the cost of travel within a city.

8. Multinomial Regressions

   In order to analyze the interactions between various indicators, multinomial regressions were taken using the =linest function in Google Sheets. In these models, the response variable is the health indicator being analyzed and the predictor variables are selected from the socioeconomic indicators. Where multiple indicators appear to have an influence on health outcomes, a multinomial regression model will help to decipher the magnitude of the effect of each predictor variable.

   The basic model to be implemented is:

   $$Y=B0 + B1*X1+B2*X2 \ldots Bn*Xn +E$$

   Where Y is a health outcome, B0 is an intercept, Bn is the coefficient for a particular socioeconomic indicator and Xn is the value of that indicator and E is the residual error (Grace-Martin 2000), although the residual error is not calculated in the results in this paper.

   Where coefficients of correlation were very low (below 0.3), predictor variables were not considered. The assumption is that where correlations are particularly low, there is no consistent or significant impact on the outcome. This means that the regression run for each health outcome and AQI only considered variables with significant correlation values (above 0.3), thus a value was not calculated for every predictor variable for each health indicator.

Results

Scaling of socioeconomic indicators to population

Each of the following socioeconomic indicators was evaluated at the metropolitan or micropolitan level to determine the degree to which it scales with population. The scale-adjusted metropolitan indicator (SAMI) for each indicator that showed a significant correlation were calculated by subtracting the observed value for each indicator from the predicted value based on the population and the formula for the power law trend line of the data.

Separated by their superlinearity or sublinearity, and organized by their exponent, it becomes clear that the majority of socioeconomic indicators scale superlinearly. Those that scale sublinearly are, for the most part, not surprising. A few interesting scaling relationships do stand out. Private goods producing industries increase sublinearly with population, while service producing industries increase superlinearly. The number of people who walk to work and who drove alone to work increases slightly sublinearly relative to the population over 16 years old, but slightly superlinearly relative to total population. Unemployment increases slightly sublinearly relative to the size of the labor force, but slightly superlinearly relative to total population. Burglary and rape scale sublinearly, while all other crime indicators scale superlinearly.

The exponents for indicators that were measured on a per-capita rather than an aggregate basis as well as value added per hour, (indicators shaded grey) should be read as β-1, so positive numbers represent a superlinear trajectory and negative numbers a sublinear trajectory. Because there is more variation in per capita indicators, the R^2 values are much lower.

Table 1:
Scaling Exponents: All socioeconomic indicators considered with scaling exponents based on power-law scaling with population

Indicator	Superlinear Scaling exponent	R^2	Indicator	Sublinear Scaling exponent	R^2
GINI coefficient	0.011	0.04	Land area per capita	-0.655	0.533
Rent as percent of household income	0.024	0.067	Establishments with more than 20 employees (mfg)	0.4698	0.20
Mean travel time to work	0.052	0.141	GDP-agriculture, forestry, fishing	0.686	0.18
Median household income	0.071	0.195	GDP-Natural Resources and mining	0.696	0.28
Productivity in Mfg (value added/aggregate hours worked)	0.134	0.10	Number of firms with or without paid employees	0.705	0.46
Larceny or theft	1.001	0.91	Population 25+ high school diploma	0.93	0.96
Property Crime	1.003	0.91	Burglary	0.939	0.84
Population 25+ some college	1.004	0.98	Population 25+ Less than high school	0.941	0.89
Aggregate usual hours worked	1.015	0.99	Aggregate income deficit	0.9597	0.90
Total population to Drove alone to work	1.019	0.997	Poverty - below 100% poverty level	0.96	0.93
GDP-Manufacturing	1.024	0.64	GDP Private Goods Producing Industries	0.962	0.77
Total population to unemployment	1.024	0.94	Population 16+ Walked to work	0.987	0.81
Poverty - at or above 150% poverty level	1.029	0.99	State and local spending	0.987	0.886
Real aggregate income	1.035	0.97	Rape	0.987	0.82
GDP-Retail	1.035	0.95	Population 16+ Drove alone to work	0.996	0.997
Aggravated assault	1.045	0.81	Aggregate number of rooms	0.997	0.99
Vacancy (# of rooms to # of vacant rooms)	1.056	0.91	Civilian labor force to unemployment	0.997	0.93
GDP-Food Service Industry	1.064	0.93			
GDP-Construction	1.064	0.93			
Aggregate household income	1.074	0.98			
Aggregate travel time to work	1.078	0.78			
GDP-Transportation and Utilities	1.09	0.83			
Value added per hour (mfg)	1.093	0.91			
Violent Crimes	1.093	0.87			
GDP	1.097	0.94			
Population attending college	1.100	0.89			
GDP-Trade	1.113	0.93			
Murder and nonnegligent manslaughter	1.119	0.76			

Table 1, Cont'd		
Population 25+ bachelor's	1.135	0.96
Aggregate gross rent	1.1599	0.96
Population 25+ graduate or professional degree	1.169	0.94
Federal non-military spending	1.175	0.678
GDP-Private Services Producing Industries	1.18	0.93
Motor vehicle theft	1.194	0.80
GDP-Finance, real estate, rental, leasing	1.263	0.88
GDP-Publishing Industries (except internet) including software	1.305	0.63
Robbery	1.325	0.86
GDP-Information Industry	1.329	0.82
Population 16+ Public Transportation to work (excluding taxi)	1.488	0.79
GDP-Performing Arts, spectator sports, museums and related	1.534	0.67

The strongest correlations between health and environmental indicators and socioeconomic indicators, (coefficients of correlation above 0.5) though only moderate, are associated with per capita measurements rather than scale-adjusted levels. Only in a few cases did scale-adjusted metropolitan indicators (SAMIs) provide a coefficient of correlation at or above 0.5. The SAMI for median household income (MHI) correlated with obesity and overall good health (BRFSS) (coefficient of correlation (r) of 0.61 and -0.50), though the coefficients of correlation for these factors were similar without adjusting for scale (r=0.55 and -0.52). A weak correlation (r between 0.3 and 0.5) can be seen between good or better health and rent, the GINI index of inequality, driving a car alone to work, walking to work and income deficit. Obesity is weakly correlated with walking to work, using public transit and the SAMI for GDP. The Air Quality Index is weakly correlated with federal non-military spending (r=-0.33), real aggregate personal income (0.38), and aggregate travel time per person to work (0.24). Although the correlation was only moderate, the SAMI for government spending showed much stronger correlations with depression and air quality than did spending per capita.

Table 2: Coefficients of Correlation: BRFSS Health Indicators and AQI to Socioeconomic Indicators					
Color Scale: significant negative correlations are green; significant positive correlations are red.	Good or Better Health	Obesity	Currently suffers from Depression	Air Quality Index	Currently suffers from Asthma
SAMI Median Household Income	0.61	-0.50	-0.20	-0.05	0.03
Median Household Income	0.55	-0.52	-0.35	0.19	-0.05
Real per capita income	0.36	-0.28	-0.22	0.02	-0.07
GDP per capita	-0.14	-0.23	-0.35	0.17	-0.23
SAMI GDP	0.16	-0.30	-0.22	0.04	-0.08
SAMI Gross Aggregate Rent	0.13	-0.27	-0.23	0.01	-0.07
Median Gross Rent	0.36	-0.57	-0.48	0.22	-0.10
Rent as Percent of HHI	-0.15	-0.21	-0.14	0.07	-0.05
GINI Coefficient	-0.32	-0.04	-0.15	-0.13	-0.26
Drove Alone	-0.40	0.54	0.26	0.06	0.04
% Bicycled to work	0.37	-0.48	0.02	-0.04	-0.06
Walked to work	0.40	-0.39	0.08	-0.20	0.18
SAMI Aggregate Travel Time	0.08	-0.17	-0.17	0.05	-0.03
Aggregate Travel Time per person 16+	-0.08	-0.13	-0.27	0.24	-0.01
Percent population with Bachelor's, Professional or Graduate Degree	0.34	0.40	0.04	0.02	0.03
SAMI Federal civilian spending	0.00	0.16	0.29	-0.33	0.13
SAMI State and local government spending	0.08	-0.20	-0.22	0.01	-0.03
Per capita Federal civilian spending	0.00	-0.07	0.00	0.03	0.02
Per capita state and local spending	0.06	-0.11	-0.10	-0.07	0.04
Murder	-0.31	0.32	0.13	0.24	-0.08
Burglary	-0.47	0.41	0.21	0.00	0.09

Testing for the coefficient of correlation between socioeconomic indicators and health indicators based on the 2012 BRFSS revealed a number of indicators that show weak or moderate relationships. Though one relationship, that between good or better health and the SAMI for median household income, revealed a moderate correlation, no other SAMI value showed a moderate correlation above a coefficient of 0.5, and the correlation between the SAMI for GDP and obesity and that between federal civilian spending and average air quality index offered the only other weak correlations (coefficient of correlation above 0.3).

Interestingly, means of transportation to work correlated strongly with overall health and obesity, but not with air quality or asthma rates. Asthma is least influenced by socioeconomic indicators, with the GINI coefficient having the strongest negative correlation (meaning that as inequity rises, asthma rates fall) and walking to work shows the strongest positive correlation (meaning that asthma rates rise as more people walk to work), although the coefficient is too low to be significant. The AQI was most strongly correlated with federal civilian spending; aggregate travel time per person and walking showed some correlation but were not significant.

Rates of depression were most strongly correlated with median household rents and bicycling to work, followed by median household income. Obesity and general good health correlated to a much wider range of socioeconomic indicators with some degree of significance. These include: educational attainment, means of transportation to work, rent and income.

Regarding the question of whether SAMIs correlate with the levels of environmental and health indicators that do not scale with city size, it appears that they do not generally serve as a significant tool for explaining the relationships between the factors studied. In the one case in which the SAMI showed a stronger correlation than per capita indicator levels, the difference between the SAMI MHI and MHI correlation to overall good health was 0.61 to 0.55. This shows that there is a stronger fit between the scale-adjusted MHI than MHI as a dollar value. Even the regional commodity price and inflation-adjusted real per capita income showed a weak correlation. However, the SAMI for MHI fails to correlate with other health and environmental indicators in the BRFSS, indicating that scale-adjusted income is moderately correlated with overall health, but not with particular ailments and not with air quality. The SAMI for MHI shows a more significant correlation with the larger health dataset from the 500 Cities project.

Correlations between health and socioeconomic indicators that have been observed and studied at the neighborhood level can also be observed at the metropolitan level.

The highest correlation between any socioeconomic SAMI and a health or environmental indicator was between MHI and good or better health. Represented visually, a general trend can be seen in which cities that perform poorer on MHI have a lower percentage of people reporting to be in good or better health. Though the pattern looks similar in each graph below (figures 1 and 2), there is a considerable difference between the SAMI MHI and MHI. Seven individual cities in the graphs below have been highlighted with distinct icons to demonstrate the change relative values. In general, very high or low SAMI cities also have high or low values even when scale is not accounted for. Accounting for scale does change how cities perform relative to each other. It is on this basis that adjusting for scale offers the possibility of revealing new relationships relative to performance on non-scaling indicators.

The SAMI MHI relationship with good or better health represents the highest coefficient of correlation recorded in this analysis (.61), while the unadjusted MHI also had a relatively high coefficient of correlation (.55), the difference in their representation indicates the difference between adjusting for scale and not adjusting for scale. The Y-axis in the SAMI graph represents the amount by which MHI exceeds or falls short of the expected MHI for a city based on its population, while the Y-axis in the MHI graph represents the dollar value of the median household income.

Compared with Figure 2, showing MHI to percent of the population reporting to be in good or better health, the trend lines in Figure 1 for the SAMI MHI have better fits, when calculated using cubic or linear functions, while MHI shows a better fit with a power law function.

Detroit, Dallas, Chicago and Los Angeles report lower scale-adjusted incomes than expected for their level of overall health, but slightly higher than expected incomes when not adjusted for scale. Los Angeles-Long Beach-Santa Ana has a relatively high median income, but for a city its size, it is below its expected income, while Berlin, NH, has a moderately low income, which is only slightly below the scale-adjusted expectation for a city its size. In these cases, it appears that adjusting for scale corresponds with health outcomes. However, in the case of Tampa, Kapaa and Torrington, the health outcomes are much closer to the trend line that does not adjust for scale than to the scale-adjusted trend line. Finally, Detroit, Berlin, NH, Dallas, Chicago and Los Angeles are on opposite sides of the trend lines depending on whether or not MHI is adjusted for scale. The coefficient of correlation with overall health in the BRFSS is stronger with scale-adjusted MHI (0.61) than with unadjusted MHI (0.55), which is reflected in the linear, but not the power law fit lines in Figures 1 and 2.

Figure 1

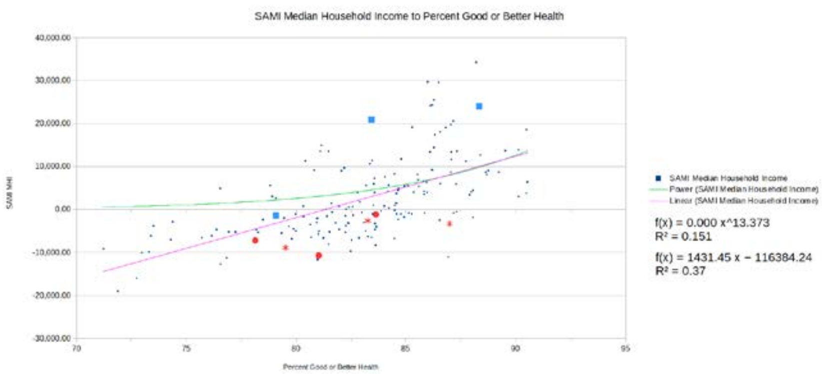
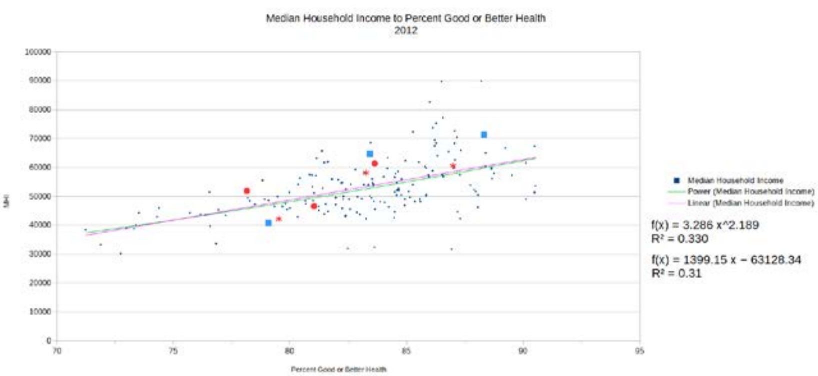


Figure 2



Red points from left to right represent: Detroit-Warren-Livonia, MI, Youngstown-Warren-Boardman, OH-PA; Tampa-St. Petersburg-Clearwater, FL; Dallas-Ft. Worth-Arlington, TX; Chicago-Joliet-Naperville, IL; and Los Angeles-Long Beach-Santa Ana, CA. The three enlarged blue points are Berlin, NH-VT; Kapaa, HI and Torrington, CT. BRFSS rates of good or better health correlated most highly with both median household income and scale-adjusted median household income (r= 0.55 and 0.61).

CDC 500 Cities

When looking at the correlations with another somewhat larger dataset from the CDC from their 500 Cities project, which includes 2015 data on 222 cities (compared with 173 Metropolitan Areas in the Behavioral Risk Factors Survey), correlations with socioeconomic indicators are considerably different. One possibility is that the difference is a result of the health data being three years more recent than the economic data, or from differences in how it was collected.

The 500 Cities Project uses both the Risk Factor Surveillance System (BRFSS) and the National Survey of Children's Health data as the primary sources of health data in its estimations. Data are calculated at the zip code and census tract level and averaged within cities to determine city level data. Though the two data sets are both based on the BRFSS data, the 2012 BRFSS data are disaggregated by MMSA, while the 500 Cities Project are disaggregated by city, meaning that in many cases, there are separate data for cities within the same MSA or there are data on cities not within an MSA.

In Table 3, below, using the 500 Cities data, only correlation coefficients for which at least one pair of indicators resulted in r = >0.3 are included.

Table 3: Coefficients of Correlation: 500 Cities Project Health Indicators to Socioeconomic indicators Red backgrounds are significant (r>0.3) positive correlations; green backgrounds significant negative (r<-0.3) correlations.												
500 Cities (n=222)	SAMI Income Deficit	MHI	AMI MHI	AMI Aggregate Rent	Median gross rent	Aggregate Travel Time per Person Age 16+	5 + with Bachelor's, Graduate or	Lowest Quintile-Share of Aggregate Income	Walked to work	Drove Alone to work	Bicycle to work	State and local spending per capita
Cancer	-0.09	-0.1	-0.02	-0.05	-0.18	-0.17	-0.22	0.36	-0.3	0.32	-0.14	-0.36
Chronic Obstructive Pulmonary Disease (COPD)	0.13	-0.37	-0.43	-0.23	-0.36	0.04	-0.37	-0.07	-0.37	0.54	-0.42	-0.3
Current Asthma	-0.03	-0.05	-0.12	-0.17	-0.15	0.15	0	-0.19	0.02	0.22	-0.15	-0.03
Diabetes	0.32	-0.28	-0.41	-0.16	-0.17	0.27	-0.46	-0.13	-0.49	0.41	-0.51	-0.35
High Blood Pressure	0.19	-0.29	-0.36	-0.21	-0.27	0.17	-0.45	-0.05	-0.52	0.55	-0.56	-0.35
Obesity	0.21	-0.35	-0.38	-0.33	-0.45	-0.03	-0.42	-0.06	-0.34	0.54	-0.52	-0.21
Sleep less than 7 hrs.	0.17	-0.08	-0.22	-0.07	-0.02	0.38	-0.27	-0.21	-0.23	0.26	-0.45	-0.17
Medication for high blood pressure	0.08	-0.29	-0.31	-0.2	-0.38	0.04	-0.32	0	-0.44	0.58	-0.5	-0.3
Stroke	0.17	-0.26	-0.35	-0.2	-0.22	0.19	-0.33	-0.09	-0.38	0.43	-0.4	-0.31
Phys Health not good >14 days	0.3	-0.33	-0.45	-0.15	-0.18	0.19	-0.47	-0.11	-0.41	0.37	-0.43	-0.34
Mental Health not good >14 days	0.18	-0.28	-0.4	-0.14	-0.12	0.22	-0.25	-0.29	-0.16	0.26	-0.26	-0.1
Binge Drinking	-0.22	0.3	0.36	0.08	0.16	-0.09	0.32	0.09	0.42	-0.34	0.39	0.28

Crime and Health

Three categories of crime showed significant correlations with health outcomes, with murder rates having the most significant correlation with health outcomes. Only cancer and asthma were insignificantly correlated with these three categories of crime. High blood pressure had the strongest correlations with all three crime categories. Obesity and high blood pressure both showed a significant correlation with the SAMI for murder, though in both cases it was weaker than the correlation with murders per 100,000 people (for both r=0.30).

Binge drinking was negatively associated with crime rates, with significant correlations with burglary and murder. This was a surprising result, since the CDC claims that binge drinking is associated with violence including homicide, suicide, intimate partner violence, and sexual assault (CDC 2018) (the coefficient of correlation with rape was 0.069, and although insignificant, this was the only crime indicator with a positive correlation with binge drinking).

The CDC also claims that binge drinking is responsible for 77% of the costs associated with losses in workplace productivity, health care expenditures and criminal justice costs, costing the country $191 billion (ibid). Binge drinking was found to be more common among people with higher incomes, and more frequent among people with lower incomes (ibid), yet the loss of productivity is not apparent on a metropolitan level, with positive correlations between binge drinking and income and attainment of higher education.

Table 4:
Coefficients of Correlation: Health outcomes and Crime

500 Cities and FBI data (n=190) Red=significant positive correlation; green=significant negative correlation	Violent Crime	Murder	Burglary
Cancer	0.01	0.02	0.18
Chronic Obstructive Pulmonary Disease	0.24	0.42	0.37
Current Asthma	0.05	0.22	0.1
Diabetes	0.32	0.52	0.31
High Blood Pressure	0.35	0.58	0.42
Obesity	0.26	0.44	0.33
Sleep less than 7 hrs.	0.26	0.45	0.19
Medication for high blood pressure	0.19	0.4	0.27
Stroke	0.18	0.4	0.29
Phys Health not good >14 days	0.28	0.42	0.32
Mental Health not good >14 days	0.21	0.35	0.23
Binge Drinking	-0.24	-0.33	-0.37

The SAMI for crime indicators are highly correlated with many SAMI values for socioeconomic indicators. Murder, robbery and burglary are most strongly associated with health outcomes, and these three indicators decrease most significantly with increasing SAMIs for rent, aggregate travel time to work, walking to work, and taking public transportation to work, while driving alone to work has a significant positive correlation. With the exception of aggregate travel time to work and the SAMI for rent, each of these indicators, when not adjusted for scale, is also significantly correlated with various health indicators.

The particularly strong correlations between SAMI values for crime, economic and transportation indicators and the SAMI values for crime show that there is a similarity in how cities perform relative to scale in multiple socioeconomic categories. This confirms previous findings by the Santa Fe Institute. Given that there were significant correlations between some categories of crime and health outcomes and between health outcomes and the SAMI for median household income, one surprising finding is that the SAMI for median household income showed no significant correlation with any scale-adjusted crime category, although it was significantly correlated with the per capita values for property crime and burglary.

Table 5: Coefficients of Correlation: SAMIs for crime to Socioeconomic indicators

Color Scale: significant negative correlations are green; significant positive correlations are red.	Violent Crime SAMI	nonnegligent Manslaughter SAMI	Rape SAMI	Robbery SAMI	Aggravated assault SAMI	Property crime SAMI	Burglary SAMI	Larceny-Theft SAMI	Motor Vehicle Theft SAMI
SAMI Aggregate Household Income	-0.19	-0.30	-0.30	-0.16	-0.26	-0.32	-0.53	-0.23	-0.30
SAMI Aggregate Earnings	-0.21	-0.30	-0.31	-0.22	-0.28	-0.35	-0.51	-0.27	-0.33
SAMI aggregate Rent	-0.32	-0.67	-0.44	-0.52	-0.22	-0.41	-0.40	-0.46	-0.22
SAMI Aggregate Travel Time to Work	-0.45	-0.39	-0.72	-0.67	-0.36	-0.68	-0.70	-0.60	-0.75
SAMI GDP	-0.26	-0.42	-0.41	-0.48	-0.25	-0.47	-0.47	-0.44	-0.44
SAMI State and Local Spending	-0.47	-0.62	-0.64	-0.68	-0.34	-0.6	-0.57	-0.59	-0.55
SAMI Federal non-military spending	0.12	0.17	0.36	0.54	-0.02	0.33	0.10	0.40	0.31
SAMI Bachelor's, Graduate or Professional degree	0.16	0.08	0.26	0.46	-0.01	0.26	-0.03	0.36	0.24
SAMI Walked to Work	-0.49	-0.59	-0.78	-0.85	-0.33	-0.81	-0.76	-0.77	-0.79
SAMI Drove alone to work	0.52	0.54	0.77	0.86	0.32	0.74	0.62	0.74	0.72
Took Public Transportation to work	-0.31	-0.34	-0.55	-0.58	-0.24	-0.53	-0.50	-0.51	-0.50
Drove alone to work	0.23	0.28	0.31	0.34	0.21	0.26	0.29	0.27	0.21

Health and Socioeconomic Indicators

Comparing the SAMI for median household income correlations with health outcomes, the SAMI values show higher correlations than median values, however, this is not true for median rent, which shows higher correlations with health outcomes than the SAMI value for rents. Portrayed in Figures 3 and 4 below are the median and SAMI values of two of the indicators most highly correlated with COPD: MHI and SAMI MHI. In Figures 5 and 6 are the indicators most highly correlated with obesity: median rent and SAMI rent.

Mode of transportation to work showed some of the highest correlations with health outcomes, with bicycling and walking having negative correlations to the prevalence of negative health outcomes and driving a car, truck or van alone to work with positive correlations. Crime rates, especially murder, had a significant influence on all health indicators except cancer and asthma, with the most significant relationship being with high blood pressure and diabetes.

Cancer prevalence was most closely associated with inequality and state and local government spending (see Table 3). Asthma was not strongly correlated with any socioeconomic factor, showing the greatest degree of correlation with driving alone to work (r=0.20). COPD showed weak or moderate correlations with rent, household income, educational attainment, state and local government spending and mode of transportation. Likewise, obesity, diabetes and high blood pressure showed weak or moderate correlation with several socioeconomic indicators.

The relationship between the SAMI for aggregate rent and obesity (Figure 5) is such that all of the high rent outliers are below the median obesity rate, while there is a tendency among low SAMI rent outliers toward higher rates of obesity. Scale-adjusted aggregate rent in most cities, however, is very close to its expected value and cities with rents near expected values span the range of obesity rates. Most of the data are close to the SAMI mean, and have considerable variance, yet the trend toward lower obesity with higher scale-adjusted rents is still apparent at values close to expected.

When correlations are ranked between socioeconomic indicators and each individual health outcome, mode of transportation to work has the greatest impact on health outcomes for COPD, high blood pressure, asthma, diabetes, obesity, binge drinking and sleeping less than 7 hours per night. Within this group of diseases, the number of people who drove alone was most determinant for obesity, asthma, and COPD, while the number of people who bicycle to work was most determinant for high blood pressure, diabetes, and sleeping less than 7 hours. Binge drinking was most highly correlated with the number of people who walk to work, and correlated with socioeconomic indicators in the opposite direction from other health indicators (See Table 3).

Cancer was most strongly affected by social inequality and government expenditure. The murder rate and the SAMI for MHI were the only indicators significantly correlated with mental health; Asthma, did not show a significant (r>0.3) correlation with any indicator studied (See Tables 3 and 4).

Figure 3

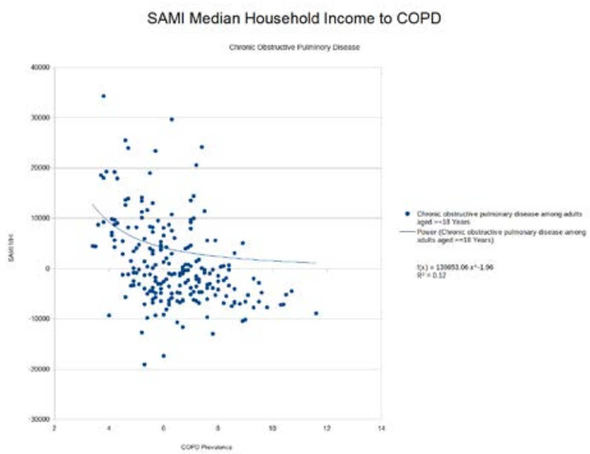
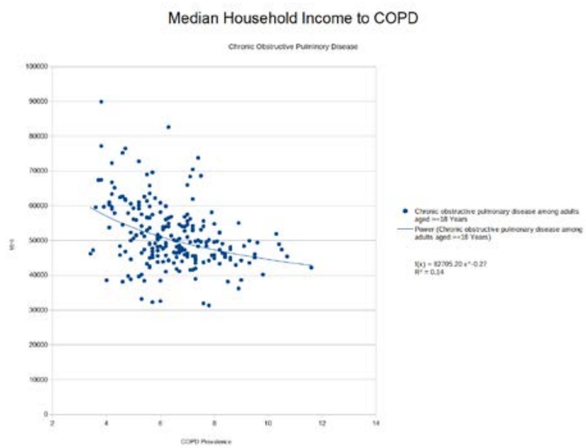


Figure 4



Although SAMI MHI shows a greater coefficient of correlation with COPD than MHI, the fit line for COPD to MHI shows a slightly stronger fit.

Figure 5

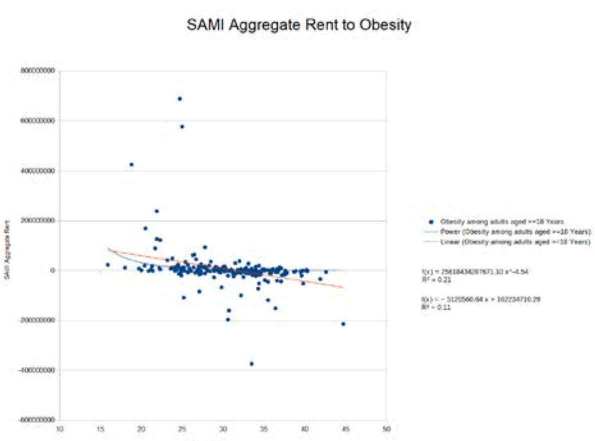
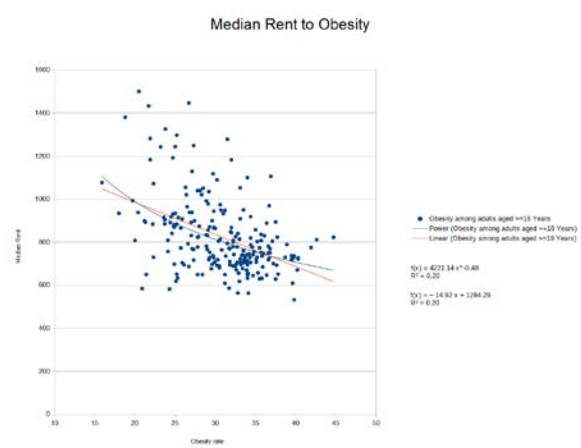


Figure 6



Health indicators can be placed into four non-exclusive categories based on how they relate to socioeconomic indicators (see Table 6, below). Rent and the SAMI for rent have significant correlations with COPD, obesity and taking medication for high blood pressure. Education and the SAMI for MHI correlate significantly with COPD, diabetes, high blood pressure, taking medication for high blood pressure, obesity, stroke, physical health and binge drinking. At least one mode of transportation showed a significant correlation with all indicators studied except asthma and mental health, and all indicators except asthma, mental health, cancer and sleep showed significant correlations with more than one transportation indicator. State and local government spending per capita showed significant correlations with COPD, diabetes, high blood pressure, stroke and physical health (See Table 6).

Diabetes was the only health indicator to show a significant correlation with the SAMI for income deficit. Travel time per person was the only indicator for which sleeping <7 hours a night showed a significant correlation, and cancer was the only indicator to show a significant correlation with the share of aggregate income going to the lowest quintile.

Table 6: Categories of disease based on association with socioeconomic indicators

Rent and SAMI Rent	Education and SAMI MHI	Mode of Transportation (Walked, Drove, Bicycle = W,D,B)	Gov't spending	Exceptions
COPD	COPD	COPD (W,D,B)	COPD	Diabetes-SAMI Income deficit
Obesity	Obesity	Obesity (W,D,B)	Medication for high blood pressure	Sleep-travel time per person
Medication for high blood pressure	Medication for high blood pressure	Medication for high blood pressure (W,D,B)	High blood pressure	cancer -lowest quintile share of income
	High blood pressure	High blood pressure (W,D,B)	Stroke	
	Diabetes	Diabetes (W,D,B)	Physical health	
	Stroke	Stroke (W,D,B)		
	Physical health	Physical health (W,D,B)		
	Binge drinking	Binge drinking (W,D,B)		
		Cancer (D)		
		Sleep (B)		

The information in Table 6 is portrayed again in Figure 7, in which it becomes visually clear that means of transportation is significantly correlated with all health indicators except sleeping less than seven hours a night and asthma.

Figure 7: Categories of disease based on association with socioeconomic indicators

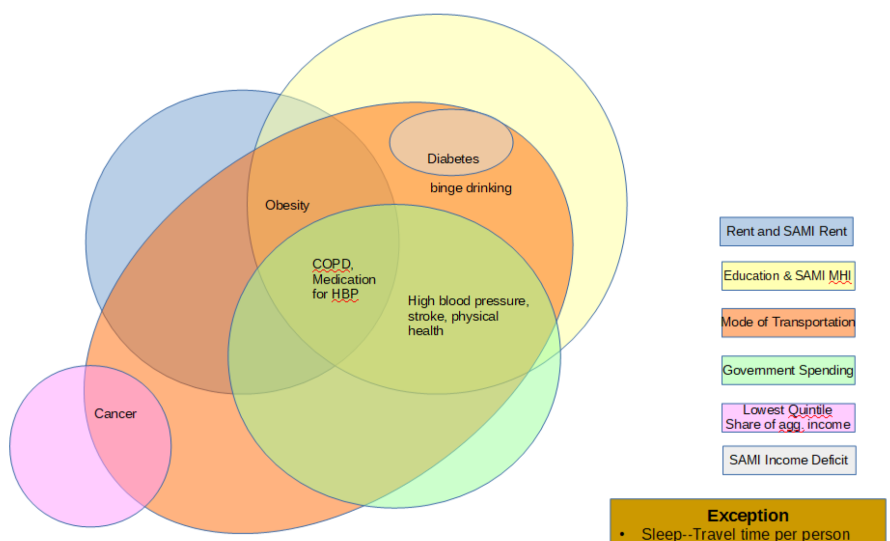


Table 7, below, shows the nine most and least healthy cities determined by the frequency with which they placed in the top or bottom 25 cities in each of the 12 CDC disease or risk factor categories analyzed from the 500 Cities project. Each city was in the top or bottom 25 cities (11% of the 222 cities analyzed) in at least 8 of 12 categories. Youngstown Ohio ranked poorly in 11 of 12 health indicators. The socioeconomic categories, except the GINI index, were chosen to represent those with the strongest correlations to the disease and risk factors.

What can be seen is that the healthiest cities are generally associated with high values in the SAMI for MHI, educational attainment, median rent, walking or bicycling to work, and state and local government spending; the healthiest cities also show low values for unemployment and driving alone and have low murder rates. The least healthy cities were characterized by generally high values in unemployment and driving to work and generally low values in the SAMI for MHI, median gross rent, higher education, walking and bicycling to work and state and local spending and high rates of murder. Poverty, the lowest quintile's share of aggregate income, and the GINI index had similar values among both the healthiest and least healthy cities (healthiest/least healthy averages of: 0.33/0.36, 3.47%/3.37%, 0.45/0.46 respectively), while per capita state and local spending was nearly 70% higher in the healthiest cities. Although police and criminal justice are typically a significant portion of state and local spending, there is no correlation between state and local spending and crime rates. People were more than twice as likely to walk to work in the healthiest cities and over ten times more likely to bicycle to work. There was a 14% difference in levels of educational attainment of a bachelor's or higher between the healthiest and least healthy cities. The healthiest cities generally had less tree canopy coverage than the least healthy (although ecozones are not accounted for here).

The average size of the healthiest cities was half that of the least healthy, and many of the healthist cities have a large college student population, with the healthiest cities having about 80% more college students on average than the least healthy (based on 2012 ACS estimates). Lawrence, KS, and College Station, TX, with 26% of the population in college, had the highest portion; while Plymouth, IN, and Cambridge, MD, also among the healthiest cities, had the lowest portion of college students. Among the least healthy cities, Dayton, OH, and Albany, GA, had the highest college populations at 10.8% and 9.2%, respectively. The percentage of people in college correlated most highly with mean usual hours worked (-0.60) and walked to work (0.60), and correlated negatively with cancer, diabetes, high blood pressure, medication for high blood pressure, stroke and physical health in the 500 Cities data, and with obesity in the BRFSS data.

Cities with higher rates of binge drinking tended to be among the most healthy cities while cities with the lowest rates of binge drinking tended to be among the least healthy. There was one exception, Provo-Orem, UT, which was both healthy and had a low rate of binge drinking. The factor most highly correlated with binge drinking is walking to work, a factor that shares a strong positive correlation with better health outcomes.

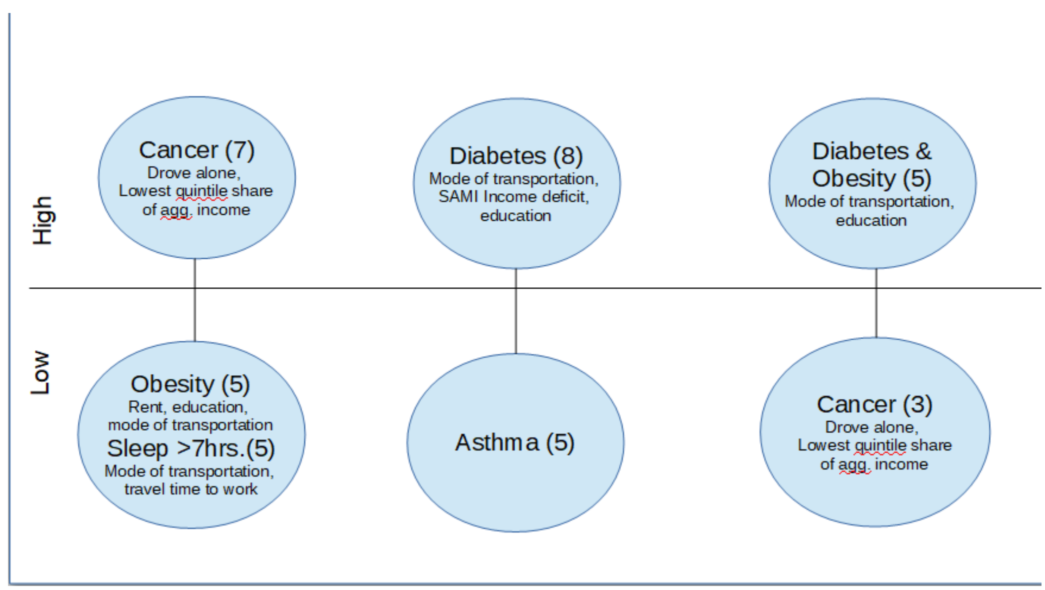
The most and least healthy cities for each disease or risk factor do not show any strong patterns in their air quality data (not shown in table). The healthiest cities for high blood pressure,

Table 7: Comparing the Most and least healthy cities															
Most Healthy Cities (in order by population size)	Population	SAMI Median Household Income	Estimate; Median gross rent	Percent population 25 + with Bachelor's, Graduate or Professional degree	Unemployment rate; Estimate; Population 16 years and over	% below 200% poverty level	Estimate; Gini Index	Estimate; Quintile Share of Aggregate Income: - Lowest Quintile	Total; Estimate; MEANS OF TRANSPORTATION TO WORK - Walked	Total; Estimate; MEANS OF TRANSPORTATION TO WORK - Car, truck, or van - Drove alone	Total; Estimate; MEANS OF TRANSPORTATION TO WORK - Bicycle	State and local spending per capita	Murder Rate per 100,000	Tree Canopy coverage
Seattle-Tacoma-Bellevue, WA Metro Area	3453748	9272	1072	37.52%	8.2	24.50%	0.44	3.66	3.6	69.7	1	6002	2.4	0.38
Provo-Orem, UT Metro Area	526804	8740	841	35.16%	6.9	34.27%	0.41	4.35	4.7	73.1	1.2	3058	0.8	0.10
Fort Collins-Loveland, CO Metro Area	300870	9067	934	43.55%	7.9	28.72%	0.45	3.46	2.7	75.4	4	5125	2.6	0.13
Boulder, CO Metro Area	297218	18586	1077	58.01%	6.9	27.26%	0.48	3.05	4.5	65.5	4.3	6372	0.3	0.24
College Station-Bryan, TX Metro Area	227843	-9300	809	33.96%	7.8	47.33%	0.52	1.97	3.7	76.7	1.6	9840	2.0	0.12
Burlington-South Burlington, VT Metro Area	211581	13956	993	40.06%	6	24.63%	0.43	3.84	6.7	72.6	1.5	5875		0.21
Lawrence, KS Metro Area	111073	2892	826	48.44%	7.2	36.08%	0.47	3	5.4	74.8	1.5	8967		0.12
Plymouth, IN Micro Area	47045	4423	650	15.85%	9.3	35.97%	0.4	4.45	3.4	79.8	1.7			0.13
Cambridge, MD Micro Area	32570	4513	808	17.94%	12.2	36.38%	0.45	3.48	2.2	80.3	0.3		1.1	
Average:	578,750	6,905	890.00	0.37	8.04	0.33	0.45	3.47	4.10	74.21	1.90	6,462	1.5	
Least Healthy Cities														
Detroit-Warren-Livonia, MI Metro Area	4304617	-7184	823	27.59%	13.9	33.01%	0.46	3.18	1.4	84.2	0.2	3099	9.0	0.16
Cleveland-Elyria-Mentor, OH Metro Area	2074824	-7134	736	27.77%	10.3	32.02%	0.47	3.15	2.1	82.1	0.3	4322	9.2	0.24
Birmingham-Hoover, AL Metro Area	1128093	-5277	775	27.05%	9.6	34.69%	0.48	3.22	1.1	84	0.1	4273		0.40
Dayton, OH Metro Area	842459	-5168	730	25.25%	10	33.55%	0.45	3.47	2.5	82.8	0.3	3778	7.9	0.13
Youngstown-Warren-Boardman, OH-PA Metro Area	564768	-8884	610	19.04%	10.3	36.92%	0.44	3.8	1.7	85.5	0.1	3419		0.29
Reading, PA Metro Area	411094	5060	812	22.35%	9.4	29.34%	0.43	4	3.1	80.2	0.2	4065	4.1	0.16
Canton-Massillon, OH Metro Area	404421	-4484	652	20.00%	10	34.00%	0.44	3.92	1.5	85.1	0.1	3200		0.22
Macon, GA Metro Area	232259	-7724	727	20.81%	10.2	43.93%	0.49	2.86	1.5	84.3	0.1	3815	7.8	0.39
Albany, GA Metro Area	157634	-10439	671	16.09%	13.5	49.07%	0.49	2.72	2.3	78.9	0.2	4326	11.8	0.21
Average:	1,124,463	-5,692	726.22	0.23	10.80	0.36	0.46	3.37	1.91	83.01	0.18	3,810	8.3	0.24
Ratio: average healthiest/average least healthy	0.51	-1.21	1.23	1.60	0.74	0.90	0.98	1.03	2.15	0.89	10.69	1.70	0.18	0.73

obesity and stroke tend to have a relatively high number of days with ozone and the cities with the highest rates of cancer tend to have a lower air quality index value (better air) and a higher percentage of good days and days with ozone as well as a lower percentage of days with unhealthy levels of PM5.

Among cities that ranked in the top or bottom 25 cities for each health indicator or risk factor, sixteen ranked in both the highest and lowest for different categories. Five of the seven cities with the highest cancer rates had the lowest rates of obesity and five had the lowest percentage of people sleeping less than seven hours per night. Five of the eight cities with the highest rates of diabetes had the lowest rates of asthma, and three of the five cities that had among the highest rates of both diabetes and obesity had the lowest rates of cancer.

Figure 8: Balances of health outcomes in most and least healthy cities



Among the cities that ranked in the top or bottom 25 cities, 16 ranked in both the top and bottom. Cities with the highest rates of cancer, diabetes and both diabetes and obesity also had the lowest rates of other health indicators. Underneath each disease/risk factor are the socioeconomic indicators that were most strongly correlated. Numbers represent number of cities with highest/lowest rates; for example, of the eight cities with the highest rates of diabetes, five had the lowest rates of asthma.

Mapping the healthiest/least healthy cities

(See maps of most/least healthy cities in Appendix A)

The business/housing mix and street intersection density for each of the healthiest and least healthy cities is mapped in Appendix A. The business/housing mix represents the agglomeration of people and economic activity, measured by the number of jobs and the diversity of employment types. while the street intersection density represents the physical density of the network, which will help to qualify commute times as relating more to distance or congestion. Both of these variables have been shown to have a strong influence on the walkability of a city (Ewing 2010), and both were expected to represent the presence or absence of physical barriers to social mixing because they represent shorter distances, and more possible connections between destinations.

In the maps in the appendix, there may be slightly more of a tendency for high intersection density not to overlap with high employment-housing mix scores in less healthy cities, though this pattern can also be seen in healthier cities.

Ultimately, there seems to be no clear pattern distinguishing healthier and less healthy cities in terms of employment housing mix and street intersection density using maps put out by the EPA. This is evidence that the structure of a city has little effect on health outcomes, and may also be evidence that social connectivity is either not a product of transportation costs or is not a factor in health outcomes.

In most cases, areas with high employment-housing mix scores are peripheral to concentrations of high intersection density. This is likely a slight drain on the potential benefit of economic agglomeration since productivity per capita declines with distance from urban cores (Fujita 2002), and the pattern shown in these maps is that areas of higher productivity (higher employment/housing mix) are not directly correlated with street density, and are not only clustered in centers. The lack of clear clustering of employment housing mix, even in healthier and wealthier cities seems counterintuitive in terms of scaling and agglomeration. The number of jobs that can be reached by people within a certain amount of time is a major factor in explaining labor productivity, with productivity falling at fairly low time cutoffs (Prud'homme 1999). Although the clustering of employment and housing does not, in many areas, coincide with the densest street networks, the distance between centers may not be so great as to significantly impact the interconnectivity, although Prud'homme (2011) shows that increasing speed in a city by 10% increases productivity by 2.9%, so even small deviations from clustering business and housing in dense centers likely has a significant impact on productivity and connectivity, though perhaps not on health.

In data analyzed in this study, density only showed a significant correlation with real aggregate personal income per person. Correlations with productivity, GDP or other measures of economic output were insignificant or did not exist.

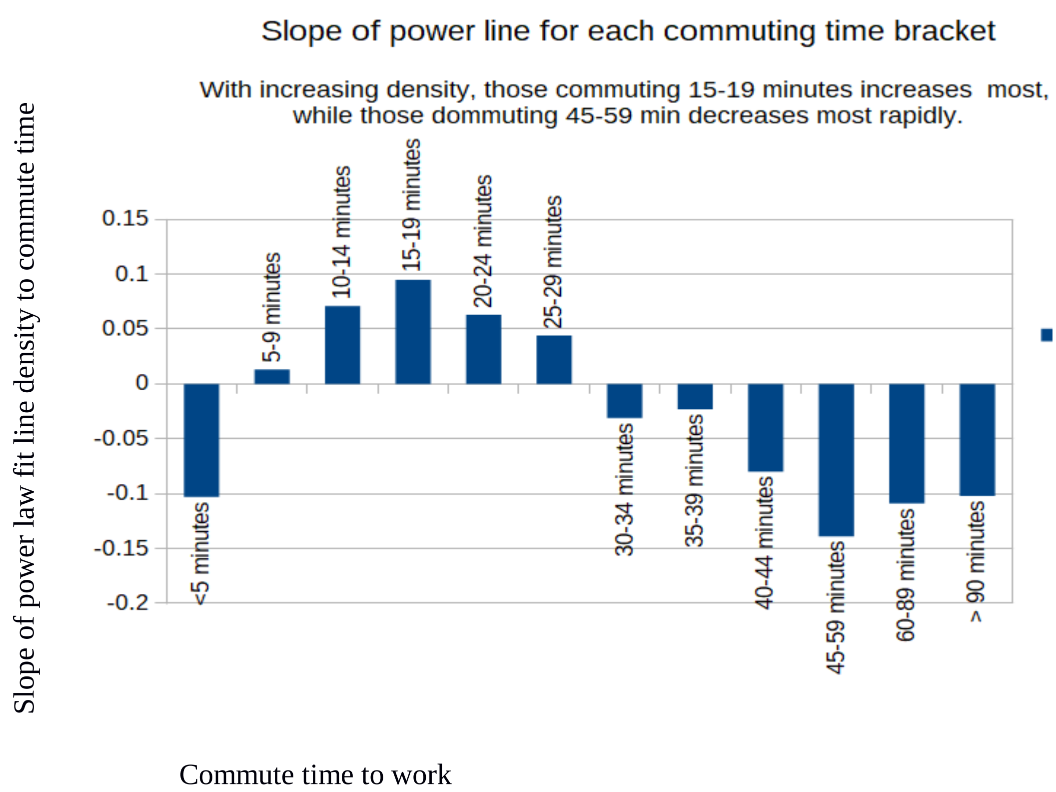
The visible patterns of peripheral areas showing higher scores for employment housing mix do indicate that commute times and the externalities of longer commute times such as lower rates of active transportation and higher pollution are likely to impact health within a city. Looking at the population density in New Mexico census tracts, there is a clear correlation between density and commute times. In Figure 9 below, the Y-axis represents the power-law exponent of the relationship between population density and the portion of commuters who fall into each time bracket, and it shows that as density increases the portion of commuters who commute between 5 and 29 minutes increases while the portion of commuters whose commute is longer than 30 minutes decreases. The negative trend with people commuting under 5 minutes is a reflection that the most densely populated tracts in New Mexico are not mixed use.

Holtzclaw (1994) also found that density has a significant negative impact on VMT. While those living in denser areas drive less and have shorter commutes, the attraction of the economy of agglomeration brings in more people from greater distances, which increases congestion and overall commute times.

For much of postwar history, vehicle miles traveled per person (VMT) has grown faster than GDP, but since 1996, GDP has grown faster than VMT. Between 1969 and 2001, VMT increased by 70%, but incomes for the bottom 60% of households only increased by 18% (Kooshian 2011), and since 1983, VMT have increased at eight times the rate of population (Speck 2012), which suggests that the finding on happiness correlates with a decreasing socioeconomic return on driven distances. The relative cost of urban travel has increased out of proportion to travel budgets for most people and represents a significant barrier to mixing, as "more driving to get to work and gain access to basic needs … does not proportionately contribute to quality of life in the form of increased incomes" (ibid 31).

Even if VMT were not related to income, it may still be considered a sort of socioeconomic metric; one study “found that a 23-minute commute had the same effect on happiness as a 19% reduction in income” (Speck 2012, 48), which may help explain some of the strong correlations between driving alone to work and health outcomes.

Figure 9: Commute Times to Residential Density in New Mexico Census Tracts



Civic engagement could be yet another indicator of the strength of social networks in a city, and to this consideration, time spent driving has a significant impact: “Each additional 10 minutes of daily commuting cuts involvement in community by ten percent,” and in another study, “Examining Walkability,” the walkability of neighborhoods had a significant impact on how much neighbors trusted each other, participated in community and volunteer projects and even impacted the portion of people who described TV as a major form of entertainment (Speck 2012, 49). If walking is positively correlated with social networks, the trend since the 1970s has been a precipitous decline in the portion of people who walk to work, from 1/10 to fewer than 1/40 today (Speck 2012, 103), (which means that technological advances are likely responsible for offsetting the decline in productivity that would be expected as a result of strained social networks). Sprawl "manages to combine the traffic congestion of the city with the intellectual culture of the countryside" (ibid, 59).[3]

Tree Canopy and Impervious Surface

Tree cover data were organized in two ways; first, as the average percent tree canopy coverage within all of the "places" within each metropolitan or micropolitan statistical area, and second as the average percent tree canopy within all of the places within all of the urban areas within each metropolitan and micropolitan statistical areas. The results were very similar, though the results from the former were consistently slightly stronger and it is these which are referred to throughout.

Considering the possibility that the very weak correlation between tree canopy coverage and economic indicators could be a result of the lag time between a tree sprouting or being planted and reaching maturity, tree cover in a given year may be a more meaningful reflection of the relative historical wealth or poverty of the city. To test this hypothesis, tree cover data from 2011 were tested for correlations with 1990 median household income and the SAMIs for median household income. The result was a significant decrease in the coefficient of correlation, with the 1990 economic data showing effectively no correlation with the 2011 tree data at the national level (See Table 8 below).

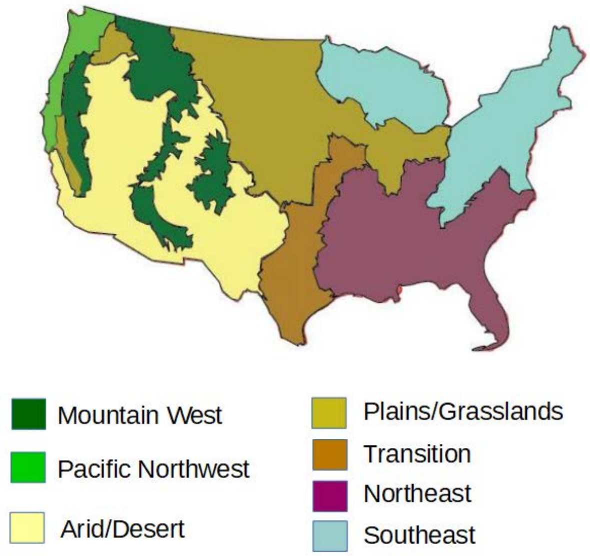
Table 8: Tree cover to current and historical income

Coefficient of Correlation	1990 Median household income	1990 SAMI Median household income	2012 Median household income	2012 SAMI Median household income
Tree Cover in Places within MMSAs	-0.05	-0.01	-0.08	-0.15
Tree cover in places within Urban Areas within MMSAs	-0.02	0.01	-0.07	-0.15

3 Speck is quoting Andres Duany, Elizabeth Plater-Zyberk, Jeff Speck, Suburban Nation, 7-12.

Figure 10

Figure 11



Ecozones created by National Land Cover Database.	Ecoregions referred to in this document created by merging similar ecozones.

Looking at tree and impervious surface data by ecoregion, however, reveals that there are correlations in some regions but not in others.

Both percent tree canopy coverage and percent impervious surface data were analyzed. Impervious surface coverage relates to tree cover in that it gives an indication of potential tree cover, exposed or undeveloped land and land with non-tree vegetation such as lawns and shrubs. Though impervious surface and tree canopy are, in some ways, inversely related factors, tree canopy and impervious surface show only weak correlations to each other in each ecoregion (ranging from -0.08 in the Plains and -0.38 in the Northeast) and do not consistently relate to health or socioeconomic factors in inverse ways.

Tree canopy coverage and percent impervious surface data were categorized by ecozone. Ecozones provided by the Forest Service were often too small to contain a significant number of cities for statistical analysis. In order to analyze the effect socioeconomic indicators on tree coverage, ecozones were combined to create six ecoregions (see maps in Figures 9 and 10: The first map contains ecozones provided by the National Land Cover Database and the second map shows the combined ecozones into ecoregions referenced in this paper).

Average tree canopy coverage within the urban areas within each MSA are displayed in Figure 12 below, showing that the ecozones have a strong influence on urban tree cover. In the maps below, green represents a low value for tree canopy and red a high value.

Figure 12

The maps of metropolitan and micropolitan statistical areas showing average percent impervious surface (Figure 13) and population density (Figure 14) are very similar.

Figure 13

Percent Impervious within Urban Areas within MMSAs

Legend

Figure 14

Population Density within Urban Areas within MMSAs

Legend

Table 9: Coefficients of Correlation between socioeconomic indicators and tree cover by ecoregion							
Color Scale: significant negative correlations are green; significant positive correlations are red.	Plains Tree Cover	Southeast Tree Cover	Northeast Tree Cover	Arid Desert Tree Cover	Transition Tree Cover	Mountain West Tree Cover	Pacific Northwest Tree Cover
N	N=140, 1990 N=43	N=252, 1990 N=70	N=206, 1990 N=68	N=111, 1990 N=29	N=63, 1990 N=1	N=37, 1990 N=10	N=23, 1990 N=4
1990 MHI	0.063	-0.115	0.137	-0.173	-0.091	-0.078	0.432
1990 SAMI MHI	0.076	-0.133	0.197	-0.187	-0.318	-0.129	-0.251
Estimate; Median household income in the past 12 months (in 2012 inflation-adjusted dollars)	-0.125	-0.079	-0.031	0.190	0.413	0.169	0.348
SAMI Median Household Income	-0.260	-0.081	-0.051	0.173	0.232	0.183	0.300
Estimate; Median gross rent as a percentage of household income	0.207	0.090	0.108	0.155	0.069	0.058	-0.511
Estimate; Median gross rent	0.084	0.035	0.090	0.216	0.327	0.498	0.082
Percent population 25 + with Bachelor's, Graduate or Professional degree	0.059	0.065	0.101	-0.256	0.444	0.187	0.020
Estimate; Gini Index	0.135	0.110	0.264	0.019	0.008	0.236	-0.363
Aggregate Travel Time per Person age 16+	0.316	0.114	0.147	0.090	0.248	0.095	0.452
Estimate; Quintile Share of Aggregate Income: - Highest Quintile	0.127	0.086	0.249	0.035	0.002	0.214	-0.325
Total; Estimate; MEANS OF TRANSPORTATION TO WORK - Public transportation (excluding taxicab)	0.075	-0.121	0.069	0.006	0.210	0.142	0.514
Total; Estimate; MEANS OF TRANSPORTATION TO WORK - Walked	-0.049	-0.071	0.096	0.167	0.020	-0.100	-0.312

The regional coefficients of correlation between socioeconomic indicators and modes of transportation to work are shown below. Because of the time it takes for a tree to grow to maturity, 1990 median household income and the SAMI for median household income were also included to test the correlation between tree cover and relative historical wealth or poverty[4]. Socioeconomic indicators included that were significant (r > 0.3) in at least one instance. In analyzing the relationship between tree cover and health data, there was a significant correlation in at least one region for each indicator in this study.

[4] 21 years is less than it takes most trees to grow to maturity. In 1990, The term "metropolitan area" (MA) was adopted by the US Census Bureau and it combined the categories of metropolitan statistical areas, consolidated metropolitan statistical areas, and primary metropolitan statistical areas. 1990 data were chosen because of the similarity in definition between metropolitan area and core based statistical area, which has been used since 2000.

Correlations between tree cover and socioeconomic performance show both positive and negative relationships (though only in the case of cancer were both significant) for most indicators depending on the region with only a single or a few regions showing significant correlations. The 1990 median household income data showed a positive correlation with tree canopy cover in the Pacific Northwest (though the sample size is so small -- only four cities with both 1990 MHI and tree canopy data -- that the results may not be significant) and a negative correlation with the SAMI for 1990 MHI in the Transition ecozone. Compared with MHI data from the year after the tree cover data were taken, there is a positive correlation between MHI and tree cover in the transition zone, suggesting that cities that were poorer in 1990 had more tree canopy in 2012 than cities that were wealthier in 1990, and poorer cities in 2012 had less tree canopy cover than wealthier ones in the Transition zone. At a national level, 1990 median household income and scale-adjusted median household income did not significantly correlate with tree cover, impervious surface or any other economic, environmental or health indicator in this study.

Most economic indicators were significantly correlated with tree cover in only one region, although median household income, median rent and aggregate travel time to work per person showed significant positive correlations with tree canopy in two ecoregions. The Pacific Northwest had the fewest cities (23 cities with 2012 economic and tree cover data and four cities with 1990 MHI data) and showed significant correlations with nine economic indicators, followed by the transition zone which showed significant correlations with four indicators. In the Southeast, Northeast and arid/desert regions, tree canopy cover was not significantly correlated with any economic indicators.

The SAMI for Median Household Income showed a significant positive correlation with tree cover in the Pacific Northwest and an insignificant negative correlation in the Plains regions, while median household income was significantly correlated with tree canopy in the Pacific Northwest and Transition regions.

Regarding the expected correlation between tree cover and socioeconomic output, the results are ambiguous. There are a few different pressures at play that may help to explain why the results do not affirm expectations. As cities get larger, they get both wealthier and denser, which will put competing pressures on tree cover, which correlates positively with wealth (Gerrish and Watkins 2018) and negatively with density (Nowak 2012). Therefore, cities that are both denser and poorer than expected or wealthier/more productive and more spread out than expected should show the lowest and highest relative tree cover within their respective ecological zones. This turns out to be somewhat true: among the most and least dense third of cities in the Northeast, those with a SAMI MHI beyond +/- $3000, 59% of those that were wealthy but not dense had more than 40% tree canopy coverage, though 35% had less than 30% tree canopy; only 6% of those that were poor and dense had more than 40% tree canopy and over half had below 30%.

When tree canopy coverage is analyzed for its correlation with health outcomes (Table 10 below), obesity showed a significant negative correlation with tree cover in three regions, as did poor physical health, sleeping less than seven hours in two regions (nearly three). High blood pressure showed a significant negative correlation in two regions and binge drinking a positive correlation in two regions. One region, the transition between plains/grasslands and forest stands out for having significant correlations between tree canopy and health outcomes in eight health indicator categories, followed by the Mountain West with four (and nearly five) indicators. The

Table 10: Coefficients of correlation between tree canopy coverage and health indicators

Color Scale: significant negative correlations are green; significant positive correlations are red.	Southeast Tree Cover	Northeast Tree Cover	Plains Tree Cover	Arid Desert Tree Cover	Transition Tree Cover	Mountain West Tree Cover	Pacific Northwest Tree Cover
N	54	45	39	31	16	15	5
Cancer (except skin) Crude Prevalence	-0.03	0.38	0.00	0.20	-0.17	0.16	-0.94
Chronic obstructive pulmonary disease among adults aged >=18 Years	0.03	-0.02	0.04	-0.18	-0.37	-0.13	-0.17
Current asthma among adults aged >=18 Years	0.20	-0.10	0.14	0.00	0.09	-0.16	0.61
Diagnosed diabetes among adults aged >=18 Years	0.22	-0.04	0.22	-0.28	-0.63	-0.28	-0.08
High blood pressure among adults aged >=18 Years	0.28	0.02	0.26	-0.37	-0.62	-0.08	-0.16
Obesity among adults aged >=18 Years	0.01	-0.16	0.19	-0.59	-0.77	-0.52	0.21
Sleeping less than 7 hours among adults aged >=18 Years (2014)	0.28	-0.04	0.27	-0.31	-0.32	-0.29	0.13
Taking medicine for high blood pressure control among adults aged >=18 Years with high blood pressure	0.14	0.15	0.28	-0.26	-0.37	0.04	-0.13
Stroke among adults aged >=18 Years	0.31	-0.09	0.26	-0.06	-0.26	-0.16	-0.03
Physical health not good for >=14 days among adults aged >=18 Years	0.10	-0.04	0.05	-0.07	-0.59	-0.30	-0.37
Mental health not good for >=14 days among adults aged >=18 Years	0.16	-0.04	0.03	-0.04	-0.21	-0.34	-0.06
Binge drinking among adults aged >=18 Years	-0.28	-0.28	0.11	-0.08	0.65	0.43	-0.19
Air Quality Index	-0.03	-0.24	0.13	0.07	0.29	-0.26	-0.01
PM2.5	0.12	-0.12	0.04	0.11	-0.55	-0.36	0.13
Ozone	-0.08	0.05	0.01	0.23	0.59	0.35	-0.16

arid and desert region and the Pacific northwest showed significant correlations with three indicators. Higher educational attainment was only significantly associated with tree cover in the transition zone.

At a metropolitan level, air quality was not correlated with tree canopy except in the mountain west and the transition ecoregions. The Air Quality Index correlated in opposite directions in the two regions, and tree canopy was correlated with lower particulate matter but higher ozone levels.

Impervious Surface

Impervious surface could be analyzed in two ways. On the one hand, it is a corollary to tree canopy in that it represents all of the areas from which trees or green spaces are excluded and relates to residents' potential exposure to nature, plants, landscaping, vacant lots or otherwise undeveloped space. If part of the explanation for trees correlating with health indicators is their calming, psychological effect on people, it is possible that permeable or undeveloped surface may have a similar effect. Through this lens, ecoregion may be important in determining what grows on undeveloped or open land.

On the other hand, percent impermeable surface within places within MMSAs is an indicator of the intensity and density of development, with more impermeable surface indicating cities with less dense or intensely developed land and could be considered without regard to ecoregion. In this sense, it is not surprising that travel time to work is negatively correlated with impervious surface in all ecoregions because less impervious surface indicates less distance between people and destinations. Percent impervious surface would be expected to increase as cities get larger, though the correlation is weak, only r=0.31. The correlation between impervious surface and aggregate real income was slightly stronger, r=0.36 (See Table 11 below).

At a national level, in addition to aggregate real personal income, the only socioeconomic indicators that showed a significant correlation with percent impervious surface were the SAMI for federal civilian, non-military spending and median household income (r= 0.33 and 0.35), while the only socioeconomic indicator that tree cover was significantly correlated with at a national level was aggregate travel time to work per capita (r=0.37).

Similarly to health outcomes, the relationship between health indicators and percent impervious in places within MMSAs differed by ecoregion (See Table 12). In the case if impervious surface, the Northeast and Mountain West showed positive correlations with most health indicators; the Plains and Transition zones showed positive correlations with the majority of health indicators, while the Arid/desert and Pacific Northwest regions showed mostly negative correlations with health outcomes. At a national level, impervious surface did not relate to health outcomes. The risk factor most highly associated with impervious surface at the national level was binge drinking.

Table 11: Coefficients of Correlation: Percent Impervious Surface to Socioeconomic Indicators by Region

Color Scale: significant negative correlations are green; significant positive correlations are red.	**Arid Desert Impervious**	**Mountain West Impervious**	**Northeast Impervious**	**Pacific Northwest Impervious**	**Plains Impervious**	**Southeast Impervious**	**Transition Impervious**	**Nationwide Impervious**
N	N=105, 1990 N=34, Fed spending N=52	N=41, 1990 N=12, Fed spending N=17	N=201, 1990 N=63, Fed spending N=91	N=25, 1990 N=5, Fed spending N=9	N=89, 1990 N=44, Fed spending N=57	N=231, 1990 N=75, Fed spending N=110	N=62, 1990 N=21, Fed spending N=22	N=942 1990 N=298 Fed Spending N=382
1990 MHI	0.014	0.118	0.051	0.554	-0.373	0.060	0.523	0.021
1990 SAMI MHI	0.083	0.277	0.064	0.792	-0.352	0.041	0.250	0.026
2012 Median household income	0.213	0.112	0.254	0.215	0.126	0.366	-0.055	0.35
2012 SAMI Median Household Income	0.094	-0.155	0.045	0.173	-0.060	0.192	-0.046	0.223
Median gross rent	0.290	0.085	0.228	0.082	0.328	0.479	-0.204	0.347
SAMI Bachelor’s, Graduate or Professional Degree	-0.312	0.020	-0.126	0.273	-0.053	-0.241	0.005	-0.123
Percent population 25 + with Bachelor’s, Graduate or Professional degree	0.056	0.029	0.188	0.519	-0.073	0.213	0.114	0.186
Aggregate Travel Time per Person age 16+	0.006	0.167	0.150	-0.234	0.385	0.116	-0.275	0.028
Quintile Share of Aggregate Income: - Lowest Quintile	-0.011	-0.041	-0.165	-0.395	-0.014	0.122	0.121	0.148
SAMI Drove Alone to Work	-0.384	-0.340	-0.148	-0.169	-0.316	-0.181	0.021	-0.142
Public Transportation to Work	0.190	0.311	0.250	0.225	0.254	0.269	0.058	0.266
Bicycle to Work	0.025	0.251	0.106	0.580	-0.072	0.223	0.120	0.186
SAMI State and Local Spending	0.096	0.347	-0.020	0.162	-0.132	0.088	-0.041	0.102
SAMI Federal, Non-military Spending	-0.034	-0.118	0.040	-0.131	-0.045	-0.072	-0.098	0.33

The Pacific Northwest showed significant correlations with eight health indicators, the Northeast with seven and the Mountain West and arid/desert regions with six each. The Plains and Transition ecoregions with two each and in the Southeast and at the national level there were no significant correlations (r>0.3) between health and impervious surface. There were also no significant correlations between tree cover and economic performance in the Southeast.

Table 12: Coefficients of Correlation: Impervious surface to health indicators and AQI by ecoregion

Color Scale: significant negative correlations are green; significant positive correlations are red.	Southeast Impervious	Northeast Impervious	Plains Impervious	Arid Desert Impervious	Transition Impervious	Mountain West Impervious	Pacific Northwest Impervious	National Impervious
N	54	45	39	31	16	15	5	
Cancer (except skin) Crude Prevalence	0.05	-0.31	-0.25	-0.53	0.36	0.21	0.19	-0.087
Chronic obstructive pulmonary disease among adults aged >=18 Years	0.00	0.23	-0.05	-0.41	0.06	0.15	-0.72	-0.026
Current asthma among adults aged >=18 Years	-0.28	0.44	-0.08	-0.34	0.13	0.52	-0.80	0.099
Diagnosed diabetes among adults aged >=18 Years	0.12	0.35	0.34	-0.03	-0.01	0.40	-0.36	0.081
High blood pressure among adults aged >=18 Years	-0.11	0.26	0.16	-0.38	0.09	0.26	-0.27	-0.053
Obesity among adults aged >=18 Years	0.03	0.34	-0.03	-0.20	0.10	0.47	-0.65	-0.008
Sleeping less than 7 hours among adults aged >=18 Years (2014)	-0.03	0.49	0.36	0.38	-0.40	0.01	-0.24	0.161
Taking medicine for high blood pressure control among adults aged >=18 Years with high blood pressure	0.04	-0.01	-0.01	-0.51	0.25	0.33	-0.43	-0.105
Stroke among adults aged >=18 Years	0.10	0.32	0.14	-0.25	0.09	0.28	-0.48	0.092
Physical health not good for >=14 days among adults aged >=18 Years	0.12	0.24	0.15	-0.05	-0.13	0.31	-0.49	0.079
Mental health not good for >=14 days among adults aged >=18 Years	-0.04	0.42	0.11	0.13	-0.25	0.33	-0.78	0.104
Binge drinking among adults aged >=18 Years	0.28	-0.17	0.09	0.25	-0.02	0.04	0.64	0.221
Air Quality Index (AQI)	0.15	0.41	0.34	0.18	-0.39	0.04	-0.15	0.17
PM2.5	0.01	0.32	0.15	0.25	0.42	0.14	0.36	0.17
Ozone	0.07	-0.28	-0.10	-0.06	-0.46	-0.09	-0.37	-0.08

Correlations between impervious surface and air quality are surprisingly strong in a few regions, but at a national level, there is no meaningful correlation. The Air Quality Index correlates positively with impervious surface (meaning worse air quality with more impervious surface) in the Northeast and plains, but negatively in the transition region. PM2.5 correlated positively with impervious surface in three regions, and ozone negatively in two.

Median usual hours worked and SAMI aggregate hours worked

When proportionally more hours are worked, it may be assumed that generally less time is available for pursuing intellectual and social interests, lowering both the level and value of social connectivity, and consequently having a negative effect on certain health outcomes. Indeed, longer hours working make a person less likely to put time into self-care in the form of physical exercise. Using BRFSS data, the correlation coefficient between travel time to work and getting adequate exercise

Table 13: Coefficients of Correlation: Hours Worked to Socioeconomic Indicators

Coefficient of Correlation	SAMI aggregate usual hours worked	Median usual hours worked
SAMI population below 100% poverty level	-0.77	-0.10
SAMI Bachelor's, graduate or professional degree	0.75	0.04
Percent population with bachelor's, graduate or professional degree	0.21	-0.38
SAMI Income deficit	-0.71	0.06
SAMI Median Household Income	0.67	0.00
Median Household Income	0.23	0.15
SAMI Aggregate Earning	0.64	0.02
SAMI Drove alone to work	0.45	-0.02
Drove alone to work	-0.02	0.19
Walked to work	-0.05	-0.45
Bicycled to work	-0.04	-0.38
Median rent as percentage of household income	-0.13	-0.55
Unemployment	0.17	0.20
SAMI unemployment	0.00	0.34
SAMI Federal Civilian Spending	0.50	-0.04

Aggregate hours worked does, in fact, scale slightly superlinearly (exponent = 1.015, R^2=0.99). (This could be a reflection of a larger percentage of the population being of working age in larger cities).

Analyzing the correlations between the scaling of aggregate hours and health outcomes reveals that the SAMI for aggregate usual hours worked correlated most strongly (r=0.3) with the number of people who are normal weight (BMI = 18.5-24.9) in the CDC's 2012 BRFSS. Accordingly, the SAMI for aggregate usual hours worked was also moderately correlated with physical activity and the likelihood of driving alone to work (r=.45). This could mean that when a greater portion of the population is working, people are more likely to drive alone to work, or it could mean that when people work longer hours, they are more likely to drive alone. There is some more evidence for the latter since the SAMI for usual hours worked was not correlated with bicycling and walking to work, but median usual hours worked showed moderate negative correlations (-0.38 and -0.45) with bicycling and walking as well as with normal weight and physical activity (-0.37 and -0.34).

Considering hours worked is a unit of economic value that can meaningfully compare output in cities of different sizes and levels of development, it has an interesting relationship with other economic outputs.

As more hours are worked relative to city size, correlations with SAMIs for federal civilian spending, higher education, aggregate earnings, household income and driving alone to work rise, while correlations with SAMIs for poverty, rent as a percentage of household income and income deficit fall. Thus, wealthier cities appear not to be more productive, but to be more efficient at extracting labor hours out of their populations. That increasing the quantity of work exerted in a city correlates with a decline in relative poverty, but not unemployment, it supports the notion that much poverty is connected to underemployment rather than unemployment. The rate of involuntary part time work spiked after the financial crisis of 2008, and has remained higher relative to unemployment due to structural changes in the economy. The most important structural changes that contribute to underemployment, measured as involuntary part time work are the rise of the gig economy, rising employment in the leisure and hospitality, education and health sectors, which tend to have high rates of part time employment (Valletta 2018). The increasing productivity of larger cities is, therefore, partly due to more full employment. However, as median hours worked rise, median rent as percentage of HHI, percent of population with higher education, as well as bicycling and walking to work fell with moderate to strong correlations.

Median hours worked showed significant and stronger correlations than the SAMI for hours worked in only a few cases: unemployment, median rent as a percentage of household income, percent of population who walked to work, and percent of people with higher education.

Though in the case of higher education and of driving alone to work, the SAMI values showed a far stronger correlation than the percent values.

With regard to educational attainment, opposite pressures are seen when considering the SAMI for higher education and median usual hours worked. The median hours worked correlates negatively with the percentage of people with bachelor's, graduate or professional degrees, while the SAMI for bachelor's, graduate or professional degree correlates positively with the SAMI for usual hours worked. More educated cities tend to work less per capita, but relatively more educated cities tend to work relatively more. Working more was also associated with making less healthy transportation choices.

While the relationship between education and hours worked appears complicated, it is interesting to note the intuitively strong negative correlation with the SAMI for poverty, and yet no correlation with the SAMI for unemployment, and a weak positive correlation with median usual hours worked. Longer hours may exacerbate unemployment, but unemployment seems to have no influence on the scale-adjusted aggregate hours worked. More hours worked on average also corresponds with more affordable rents.

Aggregate Travel Time to Work

An important part of scaling theory is based on the premise that cities exist in order to mix their populations, and that the increasing returns to scale for socioeconomic outputs are a reflection of the efficiency with which populations are mixed and social networks are maximized. This has been formulated in terms of the way travel costs impact a city, from explaining sprawl and polycentricity to defining factors such as $\bar{g}$ as a measurement of the value of the average social interaction. One way this has been demonstrated is through the sublinear scaling of infrastructure lengths with population. Indeed, if the length of roads per capita is smaller in larger cities, one would expect travel time to work to scale sublinearly to a similar degree, yet commute time is not proportional to infrastructure length and it scales superlinearly with an exponent of 1.078.

This can be explained in theory as travel expenditures are expected to increase with increasing income (Louf 2014), and aggregate MHI increases with nearly the same exponent (1.074) as aggregate travel time to work (1.078), and GDP with an even greater exponent (1.097). As the average distance between people decreases in larger cities, average travel times to work tend to increase proportionally with MHI, though there is much more variation in commute times (R^2=0.78 for aggregate commute times compared with 0.98 for aggregate MHI). While commute times to work may increase in larger cities due to traffic and the organization of the road network becoming more web-like and less dendritic and with a decrease in per-capita road capacity (Levinson 2012). Interestingly, both the highest and lowest SAMIs for aggregate travel time are for many of the largest cities in the country, but higher distances traveled are not connected to economic output at the metropolitan level (and are negatively correlated with economic output at a state level) (Kooshian 2011).

Though efficiency of mixing can be difficult to quantify, one indicator for which data are readily available that should represent average travel costs is the aggregate commute time to work. By adjusting for scale, both population and differences between commuting in large and small cities are taken into account. Cities with high SAMI values for aggregate time spent commuting to work should be cities that have either greater than expected barriers to mixing or higher travel budgets, while low SAMI values should correspond with cities with fewer barriers to mixing or lower travel budgets. Considering that there are more barriers to mixing than only travel costs, travel costs in terms of minutes of commuting either per capita or adjusted for scale should correlate with other socioeconomic outputs that scale with population.

In the case of travel time, the SAMI for aggregate travel time showed significant correlations with a number of socioeconomic outputs, while travel time to work per capita showed its strongest correlations (r=0.36) for both walked to work and for percent tree canopy cover (See Table 14).

Table 14: Coefficients of Correlation: Aggregate Travel Time to Work to Selected Socioeconomic Indicators

	Aggregate household income	Aggregate earnings	Walked to work	Drove alone	Public transit	SAMI GDP	Real per capita personal income
SAMI Aggregate Travel Time to Work	0.62	0.70	0.82	-0.73	0.66	0.66	0.66

This is in line with other studies that have found that as traffic worsens and commute times rise, fewer people drive - in fact, cities with the more congestion use less fuel per capita, and Speck suggests that congestion is one of the only factors that effectively reduces a city's dependence on driving (Speck 2012). Walking to work may not make commute times any shorter, but more people choose to walk as commute times rise. This is likely related to the association between decreasing density and increasing congestion and commute times - as most major cities in the world are trending toward less density, they are becoming less efficient and city dwellers' ecological footprint is rising (UN Habitat 2014). Additionally, it has been shown that at a regional level, greater congestion is associated with higher wages and productivity, though at a neighborhood level, the opposite seems to be true (Cervero 2000). It is also in line with findings that travel budgets rise with rising income demonstrated in the positive correlation between earnings, income and GDP with aggregate travel time to work. The association between increasing GDP or earnings with commute times can be explained as a willingness of people to sit in traffic or drive greater distances when the economic incentive of jobs in centers outweighs the disincentive of the extra commuting cost in time.

Given that longer commute times are often associated with denser and more walkable centers, the aggregate commute time in cities is a combination of distance and speed. In an international traffic study, drivers were found to spend around 9% of their time in congestion, moving at an average speed of under 9 mph (INRIX 2016), therefore, the meaning of aggregate travel time to work is ambiguous. A high travel time to work could represent a barrier to mixing, but some of the cities with the highest SAMI travel times are New York, Washington, DC, Chicago, Boston, Baltimore and Atlanta, which are all high-income cities with high rents and in which a larger than average portion of people who use public transit. It is likely that these longer commute times are a direct or indirect result of congestion. Los Angeles, CA, Detroit, MI, Phoenix, AZ and San Diego, CA, on the other hand, have exceptionally low SAMIs for travel time, but relatively high travel times per person. These cities are also characterized by high rates of inequality, and poverty and have low SAMIs for driving alone to work and aggregate real personal income. The loose association between income and scale-adjusted driving times suggests that the ability to afford longer commutes is an important factor in shaping collective behavior that results in longer or shorter commutes.

Aggregate travel time did not correlate strongly with health or environmental indicators. The strongest coefficient of correlation was 0.27 with rates of depression (BRFSS), followed by 0.24 with the average Air Quality Index and 0.22 with mental health not good for more than 14 of the past 30 days (500 Cities).

Multinomial Linear Regressions

A multinomial linear regression can help distinguish the influence of different socioeconomic variables on outcomes. The prevalence of different diseases shows correlations with multiple variables, suggesting that each exerts some influence. A multiple regression can shed light on the strength of influence of various variables in combination.

Because Chronic Obstructive Pulmonary Disease (COPD) has been described in the literature as particularly sensitive to socioeconomic status and because it shows significant correlations with a broad range of socioeconomic indicators, it will be analyzed below to demonstrate the variable strength of influence of various socioeconomic indicators. Indicators with which COPD correlated most strongly were chosen: SAMI MHI, median rent, educational attainment and driving alone to work (with r=-0.43, -0.36, -0.37, and 0.54 respectively). The formula comes out as:

COPD prevalence =-6.06406-0.000056*(SMHI)+0.00106*(Rent)-0.172182*(higher ed)+ 0.15386*(drove alone)

R^2=0.38

(calculated using “=linest” in Google Spreadsheet)

The coefficients from a multinomial regression represent how much one unit of change in the predictor variable will have on the outcome variable. Because the units and the magnitude of variation in the data for each predictor variable are different, the numeric value of the coefficients are not proportional to the significance of their influence. This makes intuitive sense, as a $1 change in the SAMI for MHI would be less significant than a $1 change in median rent, and far less significant than a 1% change in the percent of people with a bachelor’s, graduate or professional degree or a 1% change in the percentage of people who drove to work. The relatively low value of the R^2 is expected, because although socioeconomic conditions are a risk factor in the disease, they do not cause it and would not be expected to fully explain its prevalence. Interestingly, if the GINI coefficient is added to the mix, it exhibits an even greater influence than education, despite its relatively low correlation with COPD alone (r=0.12).

To demonstrate, I will apply this model to a healthy city, Burlington-South Burlington, VT, because it has a rate of COPD very close to the predicted value in this model: actual 4.70%, predicted 4.71%. If Burlington's percentage of people who drive alone to work were lowered 30%, predicted COPD value would be nearly zero, ceteris paribus, if the MHI were lowered by half, from $61,600 to $30,800, the predicted rate of COPD would rise to 6.4%, ceteris paribus. Halving the number of people with bachelor's, graduate or professional degrees would raise the expected rate of COPD to 5.0%, while increasing the median rent by 50% would raise the expected rate of COPD to 5.2%, ceteris paribus.

In this example, the hypothetical changes are drastic and at least some of the results are not realistic. For example, reducing the percentage of people who drive by 30% would nearly eliminate the illness. The whole range of values for percentage of people who drive to work falls between 50 and 86%, so moving Burlington's driving rate from 73% to 43% cannot be meaningfully calculated in this model since the hypothetical value is below the range of observed values. However, lowering the percentage of people who drive alone to work by 10% results is a predicted rate of COPD would fall by 1.1% to a rate of 3.6%, while raising the percentage of people who drive alone to 85% would bring the predicted rate of COPD to 6.5%. A great drop in the percentage of highly educated people (to 10%) or rent (to $450) would have a very modest impact on COPD rates (a rise or fall of 0.41% and -0.45%, respectively), while a great fall in income (to $25,000/yr.) would increase the risk of COPD by a modest 1.6%.

Interestingly, adding the Air Quality Index (AQI), created by the Environmental Protection Agency, to the equation gives AQI a low coefficient (0.0038) and does not improve the R^2 value. Altering the AQI in Burlington from its observed value of 36 to an unhealthy value of 96 would only raise the predicted incidence of COPD by 0.22%. However, considering the GINI coefficient as an additional variable slightly increases the R^2 to 0.4, thus increasing the level of inequality from its current 0.43 to 0.51 (the highest observed value) increases the predicted rate of COPD by 1%, while lowering it to 0.38 (the lowest observed value) lowers the predicted rate of COPD by .0.6%.

The following (Table 15) are the multinomial regression coefficients for socioeconomic variables with each health indicator. Included are only socioeconomic indicators with a correlation coefficient value above 0.3. Though no illness correlated that highly with any of the air quality data, a multiple regression was done with AQI including variables that showed correlations with values between 0.25-0.3. AQI was the only air quality indicator with significant correlations to any socioeconomic indicators, thus AQI is also analyzed below. The R^2 for AQI is so low that it shows that the variables analyzed do not explain air quality at a metropolitan level.

Table 15: Multiple Regressions: Impact of socioeconomic performance indicators on health outcomes

	R^2	AMI Income Deficit	AMI MHI	AMI Aggregate Rent	Median gross rent	Percent population 25 + with	Lowest Quintile- Share of Aggregate Income	Walked to work	Drove Alone to work	Bicycle to work	State and local spending per capita
COPD	0.37		-0.000056		0.001072	-0.1384			0.154406		
Cancer	0.25						0.598369		0.040470		-0.000062
Asthma	0.17						-1.475594		0.061440		
Diabetes	0.42	0.000000003	-0.000056			-0.3159		-0.486		-0.982316	
Obesity	0.42		-0.000120	-0.000000008	(-0.002073)	-0.5628		0.364			
High Blood Pressure	0.47	0.000000002	0.000092			-0.5345		-0.788	0.232833	-2.083708	
Stroke	0.27		-0.000022			.0047		-0.072	0.028398	-0.212207	
Physical Health	0.35		-0.000083			-0.4450		-0.327	-0.012720	-0.709793	
Mental Health	0.19		-0.000079				0.854153				
Binge Drinking	0.22			0.000092		-0.0816		0.639	0.038678	0.905879	

Table 16: Multiple Regressions: Impact of socioeconomic indicators on AQI and sleep

	R^2	SAMI Aggregate Travel Time	SAMI Federal Spending	real aggregate personal income	Aggregate travel time/person	Bicycle to work
AQI	0.04	-0.0000001	0.00031	-0.000060	0.88	
Sleep <7 hours	0.31				0.51	-2.54

Because the scale of each of these indicators is different and some of them are in terms of percent and others in terms of dollars, the following chart shows the scale range for each indicator as well as the impact on health outcomes in percent predicted change given an adjustment in the socioeconomic indicators.

Table 17: Impact of change of individual indicators based on multiple regression

	SAMI Income Deficit	SAMI MHI	SAMI Aggregate Rent	Median gross rent	Percent pop. 25 + with Higher education	Lowest Quintile-Share of Agg. Income	Walked to work	Drove Alone to work	Bicycle to work	State and local spending per capita
	Range -328,206,528 – 631,871,73	Range -19057 – 34325	Range -373,672,696 -688,560,714	Range 533-1501	Range 11-58%	Range 1.97 – 4.64	Range 0.7 – 8.6	Range 50.2 – 86.2	Range 0 – 4.8	Range 2,731 – 17,328
	impact of $10mil	impact of $10,000	Impact of $100,000,000	Impact of $100	Impact of 1%	impact of 1%	impact of 1%	impact of 1%	impact of 1%	Impact of $1000
COPD		-0.56		0.11	-0.14			0.15		
Cancer						0.60		0.04		0.00
Asthma						-1.48		0.06		
Diabetes	0.03	-0.56			-0.32		-0.49		-0.98	
Obesity		-1.2	-0.8	-0.21	-0.56		0.36	0.32	-2.01	
High Blood Pressure	0.02	-0.92			-0.53		-0.79	0.23	-2.08	
Stroke		-0.22			-0.005		-0.07	0.03	-0.21	
Physical Health		-0.83			-0.44		-0.33	-0.01	-0.71	
Mental Health		-0.79				-0.85				
Binge Drinking		0.92			-0.08		0.64	0.04	0.91	

Both the SAMI for MHI and the rates of driving alone, walking, or bicycling show moderate correlations with COPD, diabetes and other illnesses, yet the SAMI for MHI does not correlate with transportation mode choice. This indicates that that there are independent causes of the coinciding correlations with various illnesses. If the correlation with MHI is unrelated to transportation choice, both independently influence or are independently influenced by separate variables.

To test the influence of each variable on COPD independently, an interaction model was set up as a multinomial regression using first:

COPD = B0+B1*MHI+B2*driving alone*B3*MHI*driving alone

To isolate the effect of driving alone, the following equation was used:

Effect of driving alone on COPD = B0+B1*B3*MHI

=-7.127725+0.171109+-0.000008*MHI

And to isolate the effect of MHI:

Effect of MHI on COPD = B0+B2+B3*drove alone

=-7.127725+0.000546+0.000008*drove alone

To give some meaning to these results (Grace-Martin 2000B), the influence coefficients for several cities with high and low rates of COPD are displayed below (Table 18).

The influence model in Table 18 shows a very consistent influence of MHI, and a slightly more varied influence of driving alone to work. Driving alone to work seems to have slightly less than average influence in McAllen-Edinburg, TX, College Station-Bryan, TX and Youngstown, Warren-Boardman, OH-PA, and slightly more in Cambridge, MD, Plymouth, IN and Canton Massillon, OH. Though a similar percentage of people drove to work in all of the cities above (between 76.6% and 85.5%), the three cities above for which driving showed a stronger than average influence all had positive SAMI values for driving alone to work equal to more than 2%

of total population.[5] Among the cities in which driving alone showed a weaker than average influence, McAllen TX had a negative SAMI for driving alone equal to about 8% of the total population, while College Station and Youngstown both had SAMI Values close to expected. All of the other cities, for which the influence of driving alone to work had a SAMI value that was close to ±1% of total population. Perhaps, as more people than expected drive to work alone, the potential influence of a decrease in driving alone becomes greater.

Table 18: Influence Model Coefficients: Effect of median household income and drove alone to work on COPD

Cities with high and low rates of COPD

Low rates of COPD	COPD rate	Predicted COPD rate	Influence of MHI coefficient	Influence of drove alone coefficient	High rates of COPD	COPD rate	Predicted COPD rate	Influence of MHI coefficient	Influence of drove alone coefficient
Cambridge, MD Micro Area	3.4	9.2	-7.128	-6.991	Detroit-Warren-Livonia, MI Metro Area	10.3	3.1	-7.128	-6.901
Plymouth, IN Micro Area	3.5	9.2	-7.128	-6.991	Canton-Massillon, OH Metro Area	10.7	5.0	-7.128	-6.922
College Station-Bryan, TX Metro Area	4.0	0.7	-7.128	-6.885	Cleveland-Elyria-Mentor, OH Metro Area	10.4	2.9	-7.128	-6.902
Roswell, NM Micro Area	4.6	3.7	-7.128	-6.913	Dayton, OH Metro Area	10.5	4.1	-7.128	-6.917
McAllen-Edinburg-Mission, TX Metro Area	5.3	-4.3	-7.128	-6.810	Youngstown-Warren-Boardman, OH-PA Metro Area	11.6	2.6	-7.128	-6.888
Average			-7.128	-6.918				-7.128	-6.906

The relative influence of each variable can best be illustrated by using some imaginary cities given the extreme ends of the range of observed values for the SAMI for MHI and for driving alone to work. In the example below, it will be possible to more clearly evaluate the significance of the influence of each indicator. The range for the SAMI for MHI spans from $-19,057 to 34,325, and the range of observed values for drove alone to work spans from 50.2% to 86.2%. The average rate of COPD is 6.4%.

[5]That is, amount by which the number of people who drove alone to work exceeded the number predicted by scaling was more than 2% of the population.

- High SAMI MHI, high drove to work: predicted COPD rate: 2.4% (4% below mean).
- Median SAMI MHI, median drove to work; predicted COPD rate: 6.4%.
- Low SAMI MHI, low drove to work; predicted COPD rate: 5.7% (0.7% below mean).

The influence of MHI is very stable and has a slightly larger coefficient of influence than driving alone to work. There is more variance in the influence of driving alone to work on COPD. Likely this is due to variance in what other mode of transportation replaces driving alone in different cities.

Since high values for MHI and driving alone exert opposite pressures on the rates of COPD, with a high value for the SAMI for MHI associated with a lower rate of COPD and a high value for driving alone to work associated with a higher rate of COPD. When the two-factor regression is run with the maximum observed values for both, the influence of MHI seems to dominate, lowering expected COPD rates. When both factors are at their lowest observed rates, low values for driving to work alone seem to have a stronger mitigating effect on the pressure on COPD rates, such that even the poorest city, adjusted for scale, would have a lower than average rate of COPD if it also had among the lowest rates of driving alone to work.

These results, however, may be skewed because we are comparing the ends of a range spanning 36% and a range spanning over $49,000. If instead, the top and bottom quartiles for each influence variable are used (drove alone = 76.7 and 82.2; SAMI MHI = -4750 and 5209) the results show a stronger influence of driving alone.

- Top quartile SAMI MHI, first quartile drove alone; predicted COPD rate: 6.39% (0.03 below median).
- Median SAMI MHI and median drove alone; predicted COPD rate: 6.43%.
- Bottom quartile SAMI MHI, third quartile drove alone; predicted COPD rate: 6.67 (0.24 above median).

In this analysis, a 2.39% point below mean rate of driving alone to work is nearly enough to offset an income $5,662 below the mean SAMI or $4,750 below population adjusted

Table 19: Multiple Regression: Crime, Canopy and Socioeconomic indicators on health
Organized by R^2 for all five variables.
Grey columns represent results from different regressions.

N=168	Canopy	Canopy without crime	Higher Ed	SAMI MHI	burglary	murder	R^2	R^2 without crime	Difference with/without crime	R^2 Without canopy	Difference with/without canopy
HBP	12.85	15.22	-30.87	0.000034	0.002	0.697	0.55	0.38	0.17	0.46	0.09
Physical Health	3.15	3.56	-16.42	-0.000016	-0.001	0.156	0.46	0.41	0.04	0.42	0.03
COPD	3.18	3.55	-6.82	-0.000020	0.000	0.103	0.39	0.33	0.06	0.31	0.08
Obesity	4.39	5.79	-25.37	-0.000060	0.000	0.456	0.36	0.29	0.07	0.35	0.01
Sleep	12.72	14.27	-14.71	0.000015	-0.002	0.556	0.36	0.23	0.13	0.23	0.13
Taking Medication for High Blood Pressure	10.13	11.55	-15.42	-0.000046	0.000	0.454	0.35	0.25	0.10	0.27	0.08
Stroke	1.67	1.86	-3.37	-0.000005	0.000	0.061	0.32	0.26	0.06	0.24	0.08
Mental Health	4.40	4.68	-6.08	-0.000036	-0.001	0.108	0.30	0.27	0.03	0.19	0.11
Binge Drinking	-2.37	-2.98	11.14	-0.000002	-0.002	-0.113	0.23	0.18	0.05	0.21	0.02

There were only 168 metropolitan and micropolitan statistical areas for which crime, socioeconomic and tree cover data were available. Because of the relatively small number of cities, they were not analyzed separately by ecoregion.

expectations. From this, it can be calculated that a one percent change in people driving alone to work has the same effect on COPD as a difference in scale-adjusted median household income of $2,631, or about half a percent of average MHI.

Because the way people use green space in a city may be influenced by perception of the threat of crime, a multiple regression using the two most influential crime indicators, murder and burglary along with the two most influential socioeconomic indicators, the SAMI for median household income and the percentage of people with higher educational attainment and the percent tree canopy. The expectation was that when all of these factors were considered together, the fit line would be stronger for stress-related illnesses, and when crime was factored out, the R2 would fall and the strength of the influence of trees would fall. In fact, the fit line or R2 was substantially better when including canopy, but the coefficients for canopy without crime were slightly greater (Table 19). The decision whether to include crime was most influential for high blood pressure and sleep.

In order to see whether tree canopy improved predictions of health outcomes a regression was run with only the socioeconomic and crime indicators. The R2 fell slightly, the decision to include tree canopy was most significant for sleep and mental health. Percent canopy coverage is more predictive of mental health, while levels of crime are more predictive of high blood pressure, and sleeping less than seven hours a night is influenced about equally by both factors.

Discussion

Socioeconomic indicators have been shown to correlate with health and environmental indicators in previous studies, but these have been conducted at either the neighborhood level or by the socioeconomic class of individuals. Investigating this question on the scale of the metropolitan and micropolitan areas has shown consistency with previous findings on health outcomes (See House 1989), and weak or inconsistent correlations with environmental outcomes. Findings such as that trees improve mental health (Roseland 2012), reduce stress and stress-related illnesses (Tsai 2018) including asthma (Lovesai 2008), and by reducing the heat island effect, have a positive impact on respiratory illnesses. On a neighborhood level, trees have been shown to reduce local levels of pollution (Hirabashi 2016, Rao 2004). At a metropolitan level, however, Pilat et al. (2012), who found that there is no connection between levels of urban vegetation and air quality, which also aligns with the findings here.

Looking at aggregate metropolitan data has shown that scaling has limited applicability to this question, and although in many cases, scaling does not provide better correlations with health and environmental data, with the exception of the SAMI for median household income. This example may represent one new way to interpret the meaning of scaling and the connections between scaling theory and public health. Scaling aside, in the process of investigating this question, a number of unexpected insights into the relationship between certain illnesses and socioeconomic performance and behavior have come to light. Regarding tree cover, this study has shown that where there are correlations with economic or health outcomes, they are inconsistent across ecoregions. Air quality at the metropolitan level is unconnected with health and economic outcomes, which is surprising, since it suggests that the pollution, which has historically been associated with economic productivity is not a significant indicator of economic productivity or output, and it is not indicative of health outcomes, however measured.

This study has revealed certain patterns of correlation between socioeconomic indicators and health and environmental outcomes. The ultimate value of these correlations is that they suggest that certain kinds of illness may be connected to economic and behavioral trends at the metropolitan level, and therefore are the indirect consequence of economic and physical planning processes. Trends such as choice of mode of transportation may be directly linked with health outcomes, or they may be indirectly linked through common variables such as urban form and policy that both determine transportation options as well as influence other habits which impact health outcomes. The negative correlation between rates of driving alone to work and most health indicators represents both the individual impact of transportation choices on health as well as the externalities of those choices in how they affect the community. For example, each person's choice to drive contributes to air quality and safety and reinforces collective patterns of choices such as the need for parking lots, the market for patterns or types of development, and social bias toward mode of transportation.

Based on observations of the data, the diseases studied by the CDC's 500 Cities project can be grouped into four non-exclusive categories of disease - those with significant correlations to mode of transportation, those with significant correlations with education and the SAMI for median household income, those with significant correlations with median rent or the SAMI for median rent, and those with significant correlations with state and local government spending.

These socioeconomic data were found not to correlate strongly with air quality data, so although there are neighborhood and demographic discrepancies in exposure to air pollution, these do not appear significant at the metropolitan level of analysis. That is, while poorer neighborhoods may be more likely to experience higher levels of pollution, this is not the case for poorer metropolitan areas. Perhaps different types of chronic stress are both caused by aspects of urban living that are linked to urban indicators such as transportation choice, income, educational attainment, social inequality and local services, and how cities differ in these areas has distinctly different impacts on health. Stroke, high blood pressure and general physical health reported as not good for more than 14 of the past 30 days are connected to education and income levels. The stress of poverty and a lack of educational opportunity appear to exert a special stress on cardiovascular and physical health. Transportation habits, on the other hand are more associated with the health of the lungs and metabolism. This may reflect the stress of or inactivity associated with driving alone or the social isolation of cities in which driving alone is more necessary as a result of urban forms which encourage driving alone at the expense of more physically active or socially engaging modes of transportation. Cancer and mental health are most related to economic inequality and local expenditure on government services.

The correlations between health indicators and mode of transportation have multiple plausible interpretations. For each category of disease described above, I will try to explain the correlation through the lens of scaling theory. The data do not connect the health outcomes to the travel behavior of individual people, so it is not possible to say that, for example, people who drive alone to work are at higher risk of certain illnesses, but rather that people who live in cities in which a greater portion of the population drives to work alone are at a higher risk. Because choice of mode of transportation is strongly linked to physical and infrastructural characteristics in cities, the differences in health outcomes could be the product of factors which influence behavior and which result in both health outcomes and transportation choices, rather than transportation choices having a causative relationship on health outcomes.

Driving alone to work may have a direct or an indirect connection with health outcomes. A possible indirect connection would be in the form of emissions. Gasoline consumption, mostly for personal vehicles, is the largest single source of greenhouse gas emissions (TRB 2009), thus one externality of the driving habits of a city will be exposure to related pollution such as ozone and NO_2. Another indirect connection between metropolitan driving patterns and health would be a city form that either encourages or discourages more active forms of transportation.

A direct effect of driving on health would be the connection between driving, physical inactivity and obesity (Chen 2000). Commuting more than 10 miles by car has been linked to higher blood sugar, cholesterol, blood pressure, decreased happiness, worse cardiovascular fitness and not getting enough sleep (Kylstra 2014). Another study found that more time spent driving increased the risk of obesity, fair/poor quality of life, high/very high psychological distress, time stress, and having physical health or emotional problems that interfered with social functioning, with significant results above 60 minutes of driving a day and a marked increase in negative health impacts when more than 120 min per day was spent driving (Ding 2014).

Transportation behavior at the metropolitan level is more a product of policies implemented to create or promote options than it is a reflection of individual choices (Schoner 2014). Policies aimed at improving infrastructure such as bicycle, public transportation, increasing housing and employment density in order to promote walking, bicycling or public transportation would be likely to have a direct and positive and measurable impact on health outcomes. Indeed, places such as Amsterdam, in which 38% of all trips are made by bicycle, had similar rates of bicycle usage to the United States in the 1970s, but policies and activism that promoted the safety and use of bicycles changed behavior (Van der See 2015).

Driving alone to work is significantly and negatively correlated to the SAMI for state and local spending, median rent, educational attainment and for GDP, and the portion of people who use public transportation to get to work is significantly and positively correlated with many SAMI indicators (positive correlations with r>0.5 include: median rent, SAMI aggregate travel time to work, SAMI GDP, and SAMI state and local government spending, significant negative correlations with r<-0.5, with taking public transportation to work were found with most crime indicators as well as federal non-military spending), suggesting that public transportation most strongly improves social mixing. Higher density corresponds with increased proximity to destinations, therefore urban mixing can take place in central locations at a lower cost, although more concentrated centers result in longer commutes and higher rents. The data show no correlation between density in urban areas within MMSAs and travel times, but a strong correlation between the SAMI for higher aggregate transportation times and the SAMI for aggregate household income and aggregate earnings (r=0.62 and0.70), though the highest correlation is with the SAMI for the number of people who walk to work (r=0.82). This affirms previous findings that congestion is one of the strongest catalysts for altering transportation behavior (Speck 2012), and that density promotes economic development (UN Habitat 2014), though as more people walk or take other active means of transportation to work, the increase in commute times is reinforced due to the slow speed of walking or of not driving. Nevertheless, the high degree of correlation with walking (and the lack of correlation with percent rural population) suggests that commute times are most strongly influenced by a city's walkability, or rather the degree to which it is walked, and therefore its residential and commercial density.

The emergent patterns of behavior in a city, or the aggregate result of individual choices, is based on the options available to citizens. Thus, a city can be considered a sort of risk factor based on the influence on health outcomes of its economic outputs as well as its ability to generate certain collective transportation choices. Through the lens of complexity, the rules by which each agent operates generate large-scale emergent patterns. Considering that the physical, economic and social fabric of a city play an important part of establishing the basic rules, or options for individual decision making, the city itself can be seen as a risk factor. This consideration takes into account that personal choices that affect an individual's health are influenced by the social, physical or economic structure of a city (Ewing 2010), and that larger, societal structures, institutions, norms and policies contribute to health risks (Fan 2015). A person living in a sprawling city that lacks bicycle infrastructure may not individually choose to drive. Rather, that choice is collectively imposed upon each individual by the choices, policies and historical path that continually determine the emergent forms and patterns in each city.

Considering the city as a risk factor is supported by findings that the effect of the economic status of counties was found to have a significant effect on overall health outcomes related to arthritis, high blood pressure, obesity, physical activity, smoking, and overall health of residents (Shaw 2016).

Like studies that have striven to calculate the dollar-value of urban trees based on the services they provide, establishing the city as a risk factor in non-communicable diseases is a first step in calculating a dollar-value of investments in alternative transportation infrastructure, walkable neighborhoods and commercial districts, policies that redistribute wealth and alleviate poverty, provide decent wages and other services to residents, such as increasing the length and quality of education achieved by and available to residents of a city. Thinking of the city as a risk factor would help strengthen connections between socioeconomic status and the physical environment that contribute to health outcomes, and that impact the health of residents regardless of their individual status (Shaw 2016, Fan 2011).

Reflection on the Four Principles of Scaling

The correlations between the scale-adjusted values for MHI and health indicators imply that there may be some weak connection between scaling theory and a city's performance and health outcomes. Though this proposition is weakened substantially by the lack of correlation between health outcomes and other scale-adjusted socioeconomic indicators. It is more likely that the SAMI values for MHI represent well-being in a city in a way that other SAMI values do not, and that the correlation is not explained through the strength of social networks relative to city size. The role of social ties, networks and connectivity have been shown to have a significant influence on health outcomes, but the mechanisms by which health is impacted by social networks lead to patterns in health outcomes that do not correlate with the effect of social connectivity and networks on socioeconomic outputs.

The SAMI for median household income (MHI) measures the difference in an important aspect of the quality of life in cities of different sizes, and correlates with health outcomes better than the real individual income and the SAMI for the real aggregate income, a measure calculated by the Bureau of Labor Statistics to account for regional differences in prices. That the SAMI for MHI is significantly correlated with health indicates that the scaling of MHI may be a better indication of the relationship between income and well-being than adjusting for price and cost of living differences. The moderate correlations between the SAMI for MHI with physical health, diabetes, COPD and mental health (with the 500 cities data), and obesity and overall health (with the 2012 BRFSS data) suggest a practical, though very limited, application of urban scaling in connecting urban economics with metropolitan health outcomes.

Scale-adjusted values of an indicator measure the strength of agglomeration, and are theoretically proportional to the size of the social networks in a city (which would presumably be a product of transportation costs and travel budgets, as these determine the ease and degree of social mixing) and their quality (based on, for example, the level of educational attainment or discrimination, given that interactions between more educated individuals would have a higher average transactional value, $\bar{g}$, while discrimination or segregation would lower that value) (Bettencourt 2013, Galea 2011).

It would not be unreasonable to suspect that a lower than expected quality or quantity of social interactions could be connected to health outcomes such as COPD, diabetes or obesity. These illnesses have been shown in other studies to correlate with poverty as a result of the various disadvantages of poverty. Poverty, in this context, may serve as a proxy for less social connectivity due to transportation taking up a larger portion of the travel budget, especially the suburbs are where the majority of the urban poor live and are the fastest growing geography of poverty in the US (Kneebone 2013). During the 1990s, the number of poor individuals in suburbs grew at more than twice that in cities -- 19 percent compared with 8 percent. This trend accelerated strikingly in the 2000s. From 2000 to 2010, the nation’s poor population grew from 33.9 million to a record 46.2 million. As this occurred, the number of poor individuals living in the suburbs of the nation's largest metropolitan areas rose by more than half (53 percent), or 5.3 million. This was more than twice the rate of increase in cities, where the poor population grew by 23 percent, or 2.4 million. (Kneebone 2013, 17).

This trend is particularly disadvantageous for the poor, and reinforces the principle that there is a transactional aspect to interaction, and suggests that efficient mixing would lead to more equal opportunity and social equality, and that suburbs are designed to prevent this outcome. A more realistic consideration of cities may be one of largely separate social networks with limited mixing.

The World Economic Forum (2016) points to institutional silos as a significant barrier to effective planning, and given the importance of organizations and institutions in affecting certain social qualities, the degree of interconnectedness and collaboration between public entities, private interest groups and community organizations may a significant determinant of the strength and transactional quality of interpersonal networks. Not only interpersonal social mixing, but perhaps also institutional social mixing may be an important way in which the social connectivity of a city relates to economic output. If the degree of 'siloed-ness' could be indexed, I would predict that it would explain a great deal of the variance in how socioeconomic indicators scale with population.

Social connectivity has also been shown to have a significant impact on health outcomes. Therefore, it is possible and plausible that the social connectivity that explains the superlinear scaling of MHI could also influence the prevalence of certain illnesses. This would imply that high travel costs relative to travel budgets and barriers to mobility or low-density planning, the same factors that Bettencourt (2013) associates with cities that have a high or low G value, also influence behavior in ways that affect health or exposure to risk factors.

Considering that obesity and diabetes are associated with poverty in the United States and developed economies, but not in developing economies, this correlation would be unlikely to show a universal relation to scaling. However, considering how poverty and social mixing relate to each other in developed and developing economies might reinforce the connection and also help to explain the even higher correlations between disease and transportation choices. In the United States and many developed economies, the poor often live farther from the center and are compelled to drive to work and live in neighborhoods that lack transportation options and are separated from services and centers of social mixing (Kneebone 2013). Social mixing outside of the context of work may also differ in the United States and other parts of the world and especially the developing world, in that there is a greater dependence on spending money in order to participate in public life, raising an additional cultural/economic barrier to mixing. This is partly a function of so much public space being dedicated to cars that many social nodes are businesses primarily accessible by car. This barrier to social mixing along with increased burden of transportation costs relative to transportation budgets generally faced by the poor in the United States differs from the kind of economic segregation seen in many developing economies, in which the rich live in enclaves, are more dependent on driving and suffer from higher rates of obesity and diabetes.

The SAMI for MHI correlates strongly with the percentage of people living below 200% of the poverty line (r=-0.88) and moderately with the GINI coefficient (r=-0.52), although not with modes of transportation, so driving habits and MHI exert independent influences on health outcomes. As MHI and the SAMI for MHI rises, economic inequality decreases. The social barriers that arise as a result of economic inequality, and which distinguish developed from developing nations, are likely related to the causes of diabetes, obesity and COPD.

Cancer and Mental Health

Social inequality and state and local government spending impact health through investment in social services, exposure to violence, social norms and the effects of the built environment (Galea 2011), which all influence the nature and value of social mixing. A correlation with state and local government spending suggests that certain risk factors associated with cancer are related to services provided, the quality and size of infrastructure, and assistance to vulnerable parts of society. When this is less present, certain kinds of social stress become more likely, making cancer a slightly more likely outcome. With most state and local funding going into education, welfare, health and hospitals as well as police and corrections (US Bureau of Census), the social integration of the elderly and the youth would be most noticeably affected by such spending. Cities with higher state and local spending may be more likely to alleviate some of the acute stress of poverty with close to a fifth of state and local spending going to welfare programs (ibid), in addition to providing more of the long-term social uplift expected from higher per capita budgets for education.

State and local spending as well as the number of people in poverty both scale slightly sublinearly with population with decent fit lines. The R^2 for state and local spending is 0.88 and, depending on how it is measured, poverty also has a very strong power law scaling with population; measured as number of people below 100% poverty level, the R^2=0.93, while if measured as the number of people at or above 150% poverty level, the R^2=0.99. However, the poverty indicator that correlates best with cancer is the share of aggregate income going to the lowest quintile. This indicator does not scale with population, but like scaling, it represents a relative measure of inequality.

Dalton et al. (2008) found that people living in rural areas in Denmark were at a slightly higher risk for cancer. The data analyzed in this study showed that a higher percent rural correlates with lower income, lower rent, lower education, and AQI but did not correlate with cancer or any other disease measured by the CDC 500 Cities project. However, with CDC data from the 2012 BRFSS, depression was positively correlated with percent rural population (r=0.33), while obesity and asthma were very weakly positively correlated (r= 0.26 and 0.25). A SAMI for rural populations was not calculated because it scales very weakly with population size.

State and local spending, however, correlated with cancer and mental health in terms of dollars per capita, rather than in terms of a scale-adjusted indicator. While state and local governments in larger cities are more efficient with infrastructure spending per capita, this suggests that they are less "efficient" with providing social, welfare and health services, assuming that these costs offset scale-based savings on infrastructure. The health impact of state and local government spending per capita does not depend on population size.

This apparent inefficiency may be partially explained by the fact that unemployment scales sublinearly to the size of the labor force, but superlinearly relative to population. So larger cities are slightly better at putting people to work, but incur higher costs for providing for the increasing portion of the population that is not part of the workforce.

Mental health correlated most strongly with MHI and the portion of income earned by the lowest quartile. Mental health is associated with poverty-related stress (Murali 2004), and although low income people tend to live in neighborhoods with less tree canopy, mental health and stress levels in low income people are positively influenced to a greater degree than high income people by neighborhood green space (Dadvand 2014). One form of influence acute poverty, low levels of social cohesion and exposure to violence has on mental health in children is the result of chronic over-exposure to stress hormones that can affect their brain's development and have life-long effects (Jordan 2013). Exposure to green space has been shown to reduce levels of cortisol, and has been shown to improve children's performance in school, behavior, attention and memory and is a possible mechanism by which greenspace positively affects mental health (McCormick 2017). One expected result of childhood stress on adult mental health as well as of the effect of trees on mental health would be a correlation with historical median household income and mental health, but this is not the case. While poverty has been shown to put a strain in people, affecting their mental and emotional health (Jordan 2013), social stigma and discrimination against people with mental health issues may also be an explanation for the association (WHO 2018).

High Blood Pressure, Stroke, and Physical Health

The correlations between these diseases and the scale-adjusted MHI may be explained through another set of poverty-associated stresses, including a greater likelihood of exposure to toxins in lower-wage jobs and a greater likelihood for people in poverty to face inadequate living conditions, and financial stress. The financial stress of poverty has been linked to anxiety, depression and other mental health conditions (Engel 2017), and poverty has also been demonstrated as a risk factor in obesity and heart disease (ibid). Nearly 20% of people living in poverty were found by the CDC to be in poor overall health (ibid).

The fact that the SAMI for MHI correlates more strongly with these health outcomes than MHI per capita is a strong indication that poverty's impact on health is relative to its social context. The SAMI for MHI correlates more strongly with all illnesses studied than poverty rates or the SAMI for poverty suggests that average economic well-being is perhaps more important than the portion of people in poverty. This conforms with common sense that poverty is relative to its social context. Adding nuance to the analysis, it also indicates that the aggregate performance of a city relative to its scale-adjusted expected performance impacts overall health -- that is, income affects health in how it compares with other cities' performance on income. To explain this through the lens of scaling theory, the high or low SAMIs for MHI are the result of some fundamental barriers to social mixing, which would imply that in a low SAMI city, the stress of poverty is exacerbated by increased barriers to access to the whole of the city. Poverty by itself is less predictive of health outcomes than is the magnitude of the effect of agglomeration on average wages. A low performing city relative to size indicates that all parts of the economy are failing to capitalize on their proximity to each other and to labor, and health outcomes show that a low SAMI for MHI increases the likely prevalence of all of the illnesses and risk factors studied except cancer, asthma and not getting enough sleep.

A city failing to achieve the economic potential of the agglomeration of its population should signify that either travel costs are too high, infrastructure is inadequate, or clusters of industries and businesses are not dense enough to create an agglomeration effect. If this is what is happening economically, and the association between poverty and physical health is robust, it implies that the social impact of economic underperformance is particularly hard on the poor in low performing cities.

Obesity, diabetes, COPD binge drinking and sleeping less than 7 hours

This group of health indicators shows a strong correlation with choice of mode of transportation to work. Only in the case of general mental and physical health did the SAMI for MHI correlate more strongly with illness than transportation choices (diabetes showed the same correlation coefficient with SAMI MHI as to drove alone to work). In terms of scaling theory, the result raises some questions - one of the principles of the theory is that the purpose of a city is to mix its population, and if this explains the physical extent of the city as that which can be traversed by its average citizen in a day, then one would expect driving to work alone to maximize the potential social network of drivers, as they are more able to potentially meet anyone else in the city than non-drivers. However, as the portion of the population rises that uses other modes of transportation, which are far less efficient at ensuring that each person in an MMSA has an equal opportunity to potentially mix with anyone else in the MMSA, have noticeably better positive correlations with health outcomes. The value of active transportation modes is not clearly a function of distance, (if it were, driving alone would certainly trump walking), but is likely influenced by what social interactions result from a mode of transportation. For example, the possibility of talking with a passer-by in a car is near zero, but much higher on public transportation. Additionally, the types of neighborhoods or districts in which people take more active modes of transportation are more likely to be more efficient at facilitating social interactions.

There are several plausible explanations for why cities that promote other modes of transportation have better health outcomes that would not be ruled out by scaling theory. One is that scaling is not relevant since these strong correlations relate to the percentage of the population who make each transportation choice, and adjusting this for scale does not improve the correlation. However, these indicators scale very nearly linearly with population driving to work having an exponent of 1.018 and walking to work 0.987 (Walking to work scales slightly sublinearly to the size of the labor force, and slightly superlinearly to total population). Because transportation choices scale nearly linearly, scale-adjusted values do not provide a more meaningful description of the behaviors than a percent measurement.

Although population size exerts only a weak non-linear pressure on transportation choice, the transportation choices made by people in a city may be more indicative of significant differences between cities that affect social connectivity and mixing. For example, when a higher portion of people walk, bicycle or use public transportation, it can be explained through the presence of infrastructure that supports and enables it, whereas cities in which nearly all residents drive alone to work have also invested in the infrastructure that creates this outcome. Transportation mode is less a question individual choice than it is a result of infrastructure and spatial arrangement of destinations. These factors, though only indirectly represented in transportation modes, are directly connected to the principles of scaling theory. Driving alone to work is, in many cases, the most expensive mode of mixing in a city, (both to commuters and to cities in the form of increased road and highway maintenance (TRB 2009), safety costs, and the impact on air quality and climate) so in cities where this is less prevalent, the cost of mixing is lower for a greater portion of people and it also represents an infrastructure that is more likely to lend itself to more social interactions and a greater agglomeration effect, since the production advantages of agglomeration fall off over relatively short distances (Brinkman 2016). Because walking and bicycling are more sensitive to distance than driving (based on principle three, bounded human effort), their prevalence is connected to a greater density of residences and businesses, which further correlates with reduced transportation costs.

Principle four of scaling states that socioeconomic outputs are proportional to social interactions (Bettencourt 2013A), and the correlations with the SAMI for MHI with health indicators suggests that, while health is not sensitive to social interactions in the same way or to the same degree, it is nonetheless sensitive to it. That health outcomes are not as sensitive to the SAMI for aggregate income or GDP as they are to the SAMI for median household income suggests that there is something interesting about this SAMI.

The SAMI for MHI differs from many others in that it is not relative to total urban output, but a relative mean. The SAMI for aggregate household income has nearly the same scaling exponent, but far less variability. Comparing the R^2 of the SAMI for aggregate household income and for MHI gives values of 0.98 and 0.20, respectively. Population alone is a strong predictor of aggregate household income, but other factors including household size, age, and the distribution of income would skew the median household income, such that its variation relative to population more closely tracks disease outcomes.

Figure 15 Population to Median Household Income

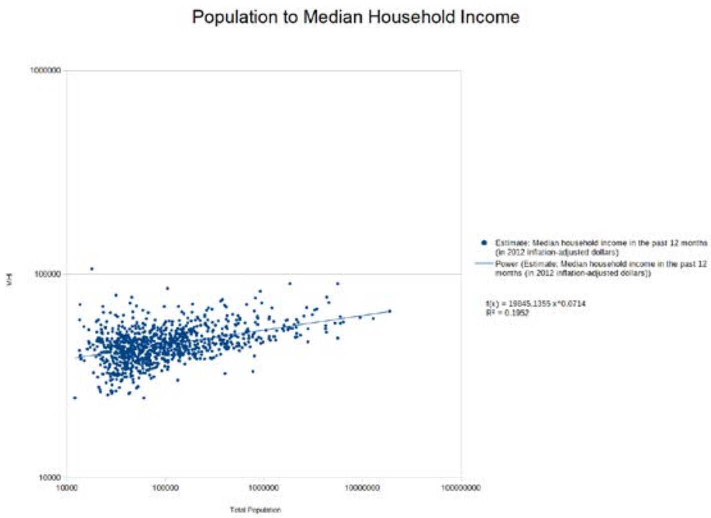
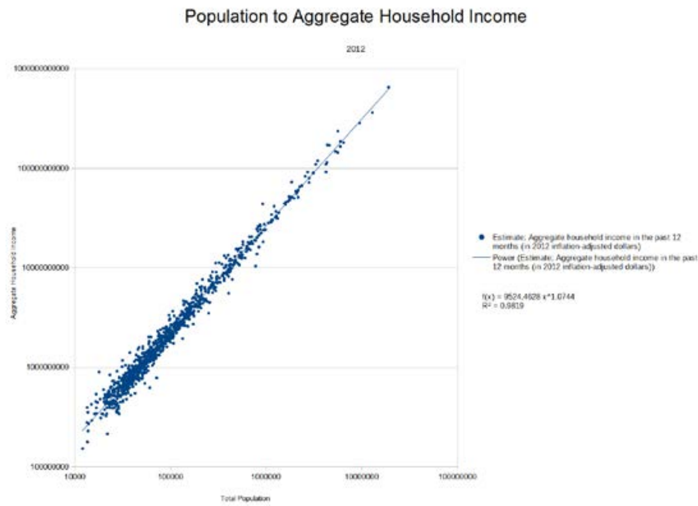


Figure 16 Population to Aggregate Household Income



Asthma

The severity of asthma has been shown to correlate with poverty and race, although the prevalence of asthma is less sensitive to socioeconomic conditions. This is also reflected in these findings in which correlations with asthma were weaker than with any other indicator studied. In the multiple regression, the influence of inequality as measured by the share of aggregate income earned by the lowest quintile was such that a 1% increase would predict a 1.5% decrease in asthma rates. The R^2 in the multiple regression model was also lowest for asthma, and was so low (0.17) that the implied strength of the influence of inequality is unreliable.

Camacho-Rivera (2014) and others have found that increases in neighborhood violence or perceptions of neighborhood violence can have a direct (stress-induced) or indirect (from over-exposure to indoor allergens, for example) influence on the early onset of asthma, but there was no apparent connection between crime rates and asthma at the metropolitan level.

COPD and Transportation

The strength of the correlation between COPD and poverty has been noted in other studies and is reinforced in this study. However, mode of transportation to work also showed a significant influence on COPD outcomes. Because MHI and modes of transportation are not significantly correlated with each other, the correlations of each to COPD are likely independent influences.

The results of the interaction model assume the independence of the influences and show that the influence of driving alone relates to the influence of income such that a one percent change in driving alone to work is equal to a $2,631 change in the SAMI for MHI in terms of their respective impacts on COPD (See Table 18). In terms of policy, this could be used to connect health outcomes to campaigns aimed at both changes in transportation infrastructure and changes in wage laws.

Both the SAMI for MHI and choice of mode of transportation to work reflect social networks in different ways. Social networks related to the workplace, and professional connections, are more influential on income and are themselves largely defined by social status such as level of education and wealth. The prevalence of people driving alone to work is more indicative of the physical space of social networks. Walking is very sensitive to distance and perceptions of safety, so if destinations are not close together, or if there is not a culture or aesthetic that promotes walking, fewer people will walk. Data on the number of people who walk to work reflect more than just the number of people who walk to work; they represent the agglomeration of businesses, density of housing, and the relative transportation cost for social interaction. Walking may also reflect an increased likelihood for particular types of social connections, that is, non-professional, non-work related social interactions. People in a city in which a greater portion of people drive alone to work would be more dependent on work and professional connections for their social network than people in a city in which business and residential densities are high enough to reduce travel costs enough to enable a greater portion of people not to drive alone. This kind of connectivity is less connected to economic outcomes than professional connectivity, but it is a more important risk factor for COPD and other illnesses than strictly economic status.

Safety may also be a significant factor in COPD. Rates of burglary and murder correlate with COPD (r=0.37 and 0.42), which confirm findings that perceptions of neighborhood safety impact respiratory disease, and suggest these findings are also applicable with regard to COPD, but not with asthma, at the metropolitan level (Camacho-Rivera 2014).

Tree Canopy and Impervious Surface

Tree canopy associates with health outcomes in different ways depending on the region. Most health indicators show significant correlations with tree canopy in only one, two or three of the seven ecoregions. Some regional tendencies can be seen, with mostly positive, but weak, correlations in the Southeast and Plains, with the strongest being stroke in the Southeast and taking medication for high blood pressure in the plains. Correlations between tree cover and health were mostly negative in the arid/desert, transition and mountain west regions, with obesity showing strong correlations in all three regions. Sleep, high blood pressure and physical health showed significant correlations in two regions, and binge drinking showed significant positive correlations in the transition and mountain west regions. Part of this could be due to cultural differences in how people relate to trees in each region. In naturally forested areas, urban trees may be seen as a nuisance for growing in unwanted places, and may be associated with poverty. In regions where trees require intentional planting, upkeep and maintenance, they may be more associated with wealth because of the additional water and economic costs associated with urban trees (Nowak 2012). In forested areas, urban areas tend to have less tree cover than rural areas, in grassland cities, more, and in desert cities there is no change or a slight decline (ibid). In the Southeast and plains ecoregions, the correlation between tree canopy coverage and income is generally negative, while in the arid, Transition and Mountain West ecoregions it is positive. The regional differences in the relationship between tree canopy and health, if they are robust, suggest that the effect of trees on human health is not related to the measurable services trees provide such as filtering the air and water, and cooling urban heat islands, because these services would not differ by region. Rather, the service of trees to human health must be in the form of more intangible differences, such as regional perceptions of trees.

Some studies have shown that there are regional differences in how urban trees are valued or perceived. Ordoñez (2017) showed some qualitative differences in the values and attitudes toward urban trees among residents of Canadian and Colombian cities, while Morré (2014) showed that people of different ethnic or racial backgrounds in Oregon had different attitudes and view of forests and urban forests. A German study found that migrants whose cultural narratives about urban forests and forests showed distinct attitudes and patterns of use and visitation to forested areas. A study in the United Kingdom found that tree cover had a positive impact on birth weight among white participants, but not among Pakistani participants (Dadvand 2014). The suggested explanation is that participants use green spaces in different ways (ibid). Differences were also noted based on both country/region of origin, gender and on generational status (Jay 2009) and have even been noted to change throughout a person's lifespan (Astell-Burt 2014). Attitudes toward trees have also been shown to be affected by the presence or absence of trees in people's neighborhoods, with residents of neighborhoods with high tree cover being less likely to want to see more trees near where they live than people living in neighborhoods with negligible tree cover, and were more likely to say that trees around apartments cause too many problems (Johnston 2012). In poor neighborhoods, especially in regions with high levels of natural tree cover, lack of maintenance is more likely to result in trees growing along fence lines, at the base of buildings and with branches that damage roofs, leading to property damage and a negative perception of trees (Schwarz 2015).

The impact of tree cover on health has also been shown to be stronger among more socioeconomically disadvantaged populations (WHO 2016). This may be because poor neighborhoods are exposed to more air pollution and the presence of trees has a greater impact in poorer neighborhoods (ibid), or it may be because wealthier people have more access to transportation and have more access green spaces even when they are not in the immediate residential vicinity (Dadvand 2014). Social capital was found to have a significant positive correlation with tree cover, but the type of green space as well as how the space is used mattered — having a green yard did not correlate with social capital and the presence of parks in a neighborhood also did not, presumably because parks in some neighborhoods are unappealing, unused or are perceived as sites of crime (Holtan 2015). Percent canopy coverage is more predictive of mental health, while levels of crime are more predictive of high blood pressure, and sleeping less than seven hours a night is influenced about equally by both factors.

While I was unable to find evidence describing regional differences in how Americans perceive trees, these findings show that cultural, regional and contextual differences in how people perceive trees exist. This implies the possibility that regional differences in the association between tree cover and health outcomes may be related to differences in cultural meaning ascribed to trees in different regions or by different populations. These findings are grounds to speculate that regional differences in how people perceive urban trees could be part of the explanation for why there are regional differences in the relationship between tree canopy coverage and health outcomes.

Regional and demographic differences have also been observed in health outcomes as well. Low income counties in the Northeast showed 19% lower mortality rates than low income counties in the West, likely due to a higher likelihood of access to effective health care (Cheng 2012). Different ethnic and demographic groups showed different sensitivities connecting health outcomes such as high blood pressure to socioeconomic status (Fan 2015). It is plausible, furthermore, that different demographic groups and geographic identities may have differing perceptions of trees, and trees may, therefore impact people's health differently based on these differences.

Similarly, the regional differences in correlations between impervious surface and health outcomes do not seem clearly related to differences such as rainfall or tree canopy (see Table 12). The relationship between impervious surface and health may also be related to cultural or planning traditions in different regions that shape attitudes toward development. Generally, a higher percent of impervious surface is associated with increased density, with the Northeast having the highest percent of urban impervious surface, while increased density is associated with a lower per capita amount of impervious surface (Nowak 2012). It could also be that the composition of impervious surfaces differs by region, for example, surface parking lots may make up a larger portion of impervious surface in the West than in the Northeast, or the size of yards or emphasis on parks and whether permeable surfaces tend to be public or private[6] may vary significantly by region.

Considering regional differences in correlations with impervious surface, one limitation of this study is that there is no reason to suppose that meaningful patterns in the composition of impervious surface correlate with the ecoregions which were chosen to approximate natural differences in tree canopy. However, regional differences, not only in the strength, but in the direction of the correlations with tree canopy and impervious surface suggest that the relationship between these indicators and health outcomes is not directly related to the ecosystem services or impacts of trees or impervious surface, as these would not vary by region. How people in different regions attribute meaning to trees and land use is a more likely explanation. The connection between impervious surface and population density may also imply that population

[6] Gerrish and Watkins (2018) found no significant impact of public versus private land on the correlation between poverty and tree cover. This does not imply that there would be no difference on the relationship between tree cover and health depending on whether canopy is on private or public land, but the lack of a connection with poverty makes it less likely.

density has an impact on health, possibly as a result of how it impacts social connectivity and mixing.

One limitation of this thesis is that It considers only average tree canopy, with no consideration for the types of green space amenities such as parks, public open space, or private yards or vacant lots. This also does not consider the accessibility or patterns of use of green space. The models used here also do not account for spatial autocorrelation, which would take into account that factors are not randomly distributed across space.

Deming, Gallup and Los Alamos: *New Mexico's high and low SAMI outliers*

Deming and Gallup are the MSAs with the lowest incomes in New Mexico, both in unadjusted income and income adjusted for scale, while Los Alamos is the MSA with the highest scale-adjusted income in the country. If urban scaling theory is to have a practical value at the local level, the it should be possible to explain the particularly low performance of these two cities, as well as the high performance of Los Alamos in terms of the theory.

The poor performance of Deming and Gallup on socioeconomic metrics is the expected result of barriers to mixing, including inefficient connectivity, high transportation costs and low jobs and population density. One common barrier these two cities have is their relatively great distance from larger, wealthier urban centers.

Deming was the headquarters for the state's largest cattle operation in the 1880s, with rangelands stretching from Truth or Consequences to the Mexican border (Sanchez 2013). During the world wars, Deming hosted large military bases and between the wars the base served as a sanatorium (ibid). The historical explanation for Deming's weak performance on wages may be the closing of military bases, the decline of the importance of passenger rail and the increasing automation of the mining industry. However, the resilience, or lack thereof, of the city to perform as it would be expected to in the context of the United States economy should also be reflected in the efficiency of the infrastructural network and the proclivity of the city to mix its population and foster social interaction.

Agriculture and mining are still important to Deming's economy, and although the military bases have closed, Homeland Security operates a border patrol operation out Deming and is one of the city's largest employers. Most of the major employers are government agencies, including border patrol, the city, county, and the public schools. The Mimbres Memorial Hospital is another of the largest employers, and the rest of the major employers are generally low-wage retail, including Wal Mart and some grocery stores. Manufacturing comprises just over 7% of the economy (Deming-Luna County Business Resource Guide 2010). The city's website does not include a master plan or planning documents beyond the city's ordinances and a water plan.

Gallup's economy is also heavily dependent on government jobs, a hospital and low-wage retail. The city is also particularly diverse, with large native populations from different tribes, as well as a mix of Hispanic, white and a considerable Arab population. In the case of Gallup, this diversity may act as a barrier to mixing resulting from cultural silos. The railroad and highway also cut through the city, creating a north-south divide.

The Gallup MSA ranks particularly low on the national SAMI for income. Part of this low performance may be explained by the fact that Gallup is surrounded by reservation land that has a low level of services. The city of Gallup serves as a regional center to which people come for many basic services such as groceries, shopping, medical care, etc. The city, with a population of just under 20,000, is able to perform about as expected, while the MSA, with a population of just under 75,000, contains a large rural population that is poorly connected to services and that depends on long commutes to the city. The contrast between the availability of jobs and services in the MSA inside and outside the city means that the majority of the residents of Gallup MSA face a significant barrier to mixing. The difference between the median household income in the city and the MSA in 2015 was $17,695. The percentage of people living at or below the poverty level in the city and MSA were 76% and 66%, respectively in 2010 (75% and 63% in 2015) (Census Bureau). Educational attainment tends to have a high correlation with income, and the percentage of people in the city and MSA, respectively, with a bachelor's or higher was 21% and 11% of pop over 25 years old in 2015 (ibid). The University of New Mexico has a branch campus in Gallup, but it does not offer any bachelor degree programs (only associate's and certificate programs) (www.gallup.unm.edu).

Gallup implemented a growth management master plan in 1999, and has updated the plan in 2009 and 2016. The plan includes recommendations such as increasing connections between local and collector streets, promoting affordable housing and multi-modal transportation, mixed use development and incentives for higher density development. Gallup also has a downtown redevelopment plan and an Arts and Culture Development Plan. Despite the implementation of this plan, incomes in Gallup have not been keeping pace with population growth. Since 2009, the population has been slowly and steadily rising by around 1.4 percent per year, yet after a peak in 2011, incomes have been steadily declining by around 1.9 percent per year (Census Bureau). Bettencourt et al. have noted the persistence over time of a city ranking high or low on a SAMI distribution (2010), and although growth management plans are promoted as a tool for sustainable economic stimulation (UN Habitat 2004) and have stimulated economic growth in other cities (TRB 2009), Gallup seems to persist as a low-wage city.

While growth management does have the potential to stimulate economic activity, it is most effective when implemented from a state level (rather than at the city level) (Carruthers 2002). Two states that have used such state-level growth management plans to successfully lead population growth to outpace the growth of urbanized land, are Arizona and Nevada (ibid, 1966). Not only are state-level growth management policies more effective at curbing sprawl, they also mitigate the upward pressure on housing prices that can stem from local restrictions (Downs 2003).

In a diagnostic report, the City of Gallup (2017) explains that the mixed-use zoning has not been taken advantage of because of high costs of development. It also states that C-1 zoning, which is for small retail spaces is also not being developed. The relatively low rate of population growth, which amounts to around 300 people per year, may make plans and policies adopted by the city very slow to generate impactful change and may help explain the persistence of Gallup's low performance on income despite efforts at intervention.

On City-data.com's crime index, Gallup ranks in the 98th percentile of cities for high crime rates. Deming was in the 86th percentile.

Both Deming and Gallup are geographically isolated and lack a close connection with a nearby, high-performing city. Deming, like many of the lowest performing cities is near the Mexican border[7], and Gallup is surrounded by Native American reservation land. Many high-performing cities, including Los Alamos, are geographically near other high-performing cities,

[7] The lowest performing three cities for scale-adjusted median household income are McAllen-Edinburg-Mission, TX, Rio Grande City, TX, and Brownsville-Harlingen, TX. All three are along the Mexican border and adjacent to each other. 13th lowest is Raymondville, TX, which is adjacent to McAllen-Edinburg-Mission and Brownsville-Harlingen. When adjusted for scale, Gallup is the 17th worst-performing city and El Paso is 21st and Deming 29th.

suggesting that the structure of a city or the degree to which it supports or impedes mixing is only part of the explanation for its performance, and regional trade or traffic play an important role in determining performance (supported by Jacobs 1984). In many structural ways, the city of Espanola resembles the city of Gallup, but performs about how it would be expected to perform, but its juxtaposition between the two high-performing cities of Santa Fe and Los Alamos cannot be ignored as a factor.

At the top of the SAMI for how income scales with size is Los Alamos, NM. The presence of the national labs, which are the foundation for the city's economy and its *raison d'etre* explain its high incomes, with over $2.5 billion of annual federal funding flowing into the small city to support the labs. For this reason, the logic connecting urban form as a tool for mixing to socioeconomic output is not relevant in Los Alamos in the same causal way as it would be in a city whose economic performance reflects the trading and interaction of its citizens. Los Alamos might reveal if there is a causality in the relationship between urban form and economic output. That is, if Los Alamos, because its population does not need to mix in order to maintain economic productivity, would be less likely to develop in a way that encourages mixing than other cities with similar SAMI rankings but different historical explanations for their higher incomes. If the nature of the physical network impacts the social network, and thereby leads to economic output, Los Alamos, because its economic output exists independently of the quality or nature of local social networks, should not necessarily exhibit features that would normally contribute to high economic performance in other cities. If Los Alamos does resemble other high-SAMI cities, it might challenge the theory of causation stemming from social networks as they emerge from physical space.[8]

Yet, the uniqueness of Los Alamos could also be interpreted as a city with two exceptional social-networking advantages. One advantage is that a large portion of people work in more-or-less the same place in related fields. So, while housing patterns or street networks could be efficient or inefficient, the centrality of the labs in people's lives may facilitate an exceptional degree of mixing. A second advantage is that the transactional value of social interactions is likely to be higher because the city that is more highly educated than most, with over 66% of the

[8] This is not necessarily the case. The city could have been planned in a way that resembles other high-performing cities, in which case, the structural influence on social networks could reinforce the already strong productivity.

population over 25 years old holding a bachelor's or higher (Census Bureau). That is, the information that residents of Los Alamos can exchange is likely to be of higher value. Among socioeconomic qualities that correlated most closely with scale-adjusted income was educational attainment.

Another outlier city, McAllen, TX, tops the list among 100 of the largest American MSAs in meeting four Sustainable Development Goals. The Sustainable Development Goals in which McAllen did exceptionally well were on the amount of public recreational and open space; Sustainable Cities and Communities, which is a measurement of a broad range of indicators including those related to public transit, walking, access to parks and housing affordability; Responsible Consumption and Production, which measures the amount of toxic waste released into the air, water or land per square mile of MSA, and Climate Action, which measures tons of carbon produced in each zip code within the MSA.

In some way, the relative poverty of McAllen may explain its high rankings in these categories. A lack of polluting industries could be a better explanation for the relative poverty than the presence of well-regulated industries that pollute less than their counterparts in other cities. The ample amounts of open space that give it a high rating in one category may reflect either weak planning or modernist, auto-oriented planning that could also be a contributing factor to the lack of economic development, either of which would describe a lower population density compared to other MSAs, which could give each zip code a lower carbon footprint, even if it is a higher per capita footprint.

Urban containment strategies like the one in Gallup should improve city performance on CO2 emissions (UN Habitat 2014) as well as on economic indicators (Cervero 2000). Urban peripheries are often the site of lower wages and higher property values, which put a particular burden on municipal governments that provide services for, but miss out on the taxes from peripheral residents. Expanding the limits of the incorporated city would have the immediate effect of lowering the median wages, however, the long-term impact could be beneficial to these residents and to the city as a whole. In Santa Fe, for example, several large businesses such as Wal Mart and an outlet mall have established themselves just outside the city limits such that they serve and employ the residents of the city but are not bound by the city's higher minimum wage laws and they pay county, but not city taxes.

The issue of urbanized areas developing outside of the administrative boundaries of a city shows a lack of coordination in planning between cities and counties, which permits and perhaps encourages patterns of sprawl and building just outside city limits. The World Economic Forum (2016) recommends that administrative boundaries should reflect the ecological and economic footprint of cities. A serious problem that arises out of the disconnect between urbanized areas and administrative boundaries is that of taxation. Residents outside the urban boundary are likely dependent on many of the services the city provides as a result of its economy of agglomeration, job opportunities, infrastructure and services, yet those services are paid for by taxes collected within the city limits. This may happen simply because it is cheaper to develop land outside the city limits, or it could reflect other sorts of motivations such as residents' desire to form separate school districts to avoid their kids mixing with the whole city population, which would mean that the city school districts, especially in places where they depend on property taxes, will be underfunded compared to nearby suburban districts -- a pattern that will reinforce the formation of social enclaves rather than encourage social mixing throughout a city.

Conclusions

This thesis set out to investigate the question of whether the scaling of socioeconomic outputs correlated to health and environmental outcomes within metropolitan and micropolitan statistical areas. Only the scale-adjusted values for median household income consistently showed stronger correlations than unadjusted median household income with multiple health outcomes. In the case of every other socioeconomic variable considered, when there was a significant correlation ($r>0.3$) per capita values showed stronger correlations than scale-adjusted values.[9]

While the scaling of all socioeconomic indicators with population is explained with the same overarching theory of networks and social mixing, median household income is unique in that it scales in a way that correlates with city performance on most health indicators. Because the explanations for the mechanisms that connect health and socioeconomic status relate to stress,

[9] The few singular exceptions to this were the SAMI for federal non-military spending with AQI and impervious surface;

lack of social integration or lack of access to educational opportunities, scale-adjusted MHI is likely a better representation of these factors than unadjusted median household income. This is likely because median household income corresponds with general well-being and the average citizen's ability to achieve adequate social integration in a way that other scale-adjusted indicators do not.

High values for scale adjusted median household income had strong positive correlations with the lowest quintile's portion of aggregate income, educational attainment, median rent, and real income per capita. Scale adjusted median income showed strong negative correlations with income deficit per capita, the highest quintile's portion of aggregate income, and the GINI coefficient of inequality. The SAMI for MHI, therefore, relates to well-being as a measure of inequality, deficit, better education and higher rent. In terms of scaling theory, the SAMI for MHI likely represents a significant measure of the presence or absence of barriers to certain types of social mixing and it likely also relates to the average value of social interactions. The kind of social barrier the SAMI for MHI might represent is an expanded potential social network due to less extreme class divisions, and material access to participate in more of the trade that goes on in a city.

Because median household income was the one exception in which the scale-adjusted values correlated more strongly with health outcomes, it casts some doubt on whether scaling theory has any explanatory value for a city's performance on health indicators. However, there is substantial evidence in the literature that connects health outcomes to social connectivity and social mixing that it is plausible that scaling theory has some explanatory power. The strength of correlations between modes of transportation and health outcomes at the metropolitan level further strengthens the likelihood that social networks and mixing help explain health outcomes, as transportation mode is a reflection of the physical network within which social networks are nested. Although there is substantial evidence that forms of transportation impact health directly, it is also possible that analyzing trends at the metropolitan level accounts for the positive and negative externalities of transportation choices as well as the infrastructural differences between cities that explain the transportation choices.

One result of this study has been a broad analysis of the various ways urban socioeconomic performance, including transportation behaviors, relate with health and environmental outcomes. The prevalence of asthma was the indicator least influenced by socioeconomic performance and was one of only three health indicators to show a significant correlation with tree canopy at the national level (along with sleeping less than 7 hours and mental health). BRFSS data (but not 500 Cities data) on asthma showed a significant negative correlation with the share of aggregate income going to the highest quintile and a nearly significant correlation with the GINI index, such that as a greater share of income goes to the highest quintile or as inequality rises, asthma rates fall. This is surprising, given that it suggests the opposite of what other literature has shown on how social segregation increases the risk of asthma (Rona 2000).

Asthma has been found to be influenced by perceptions of safety (Camacho-Rivera et al. 2014), but it did not significantly correlate with any crime indicators. Also surprising was that asthma did not correlate with any of the measures of air quality. One limitation of this study is that it considers only the prevalence of asthma and not its severity.

Means of transportation, whether walking, bicycling or driving alone to work, correlated strongly with all health outcomes except mental health and asthma (500 Cities data), but the portion of people taking public transportation did not. BRFSS data on obesity correlated significantly and negatively with use of public transportation. Means of transportation to work did not, however, correlate with any measurement of air quality, which is surprising due to the substantial contribution of automobiles to air pollution and the negative correlation between density and commute times.

Findings in this study confirm previous findings that higher transportation times are related to congestion and higher incomes. Scale-adjusted commute times to work correlated very strongly and positively with scale-adjusted rates of walking to work (0.82), and scale-adjusted economic performance and negatively with scale-adjusted crime rates. In general, scale-adjusted socioeconomic values showed a high degree of correlation with other scale-adjusted socioeconomic values.

Literature on the effect of socioeconomic performance of neighborhoods or counties or the socioeconomic status of individuals on health outcomes has shown similar connections to these findings at the metropolitan level (with the exception of asthma). These findings may prove useful is in considering cities as a risk factor, by showing the collective impact of factors such as poverty, income, crime, government spending, crime, education and transportation behaviors on health. Likewise, these findings may lead to a quantification or a health impact cost of policies that would affect a city's socioeconomic or transportation performance.

The healthiest and least healthy cities had several things in common. The healthiest cities had high scale-adjusted median household income, educational attainment, median rent, state and local government spending and more people walked or bicycled. Healthy cities also had low values for unemployment, driving alone and murder rates. The least healthy cities were characterized by high values in unemployment and driving to work, and generally low values in the SAMI for MHI, median gross rent, higher education, walking and bicycling to work and state and local spending and high rates of murder. Both the healthiest and least healthy cities had similar rates of poverty, share of aggregate income earned by the lowest quintile, and GINI coefficient. In the healthiest cities, the percentage of people who bicycled to work was ten times higher, while the percentage who walked to work was twice as high than in the least healthy. There was a 14% difference in attainment of higher education and the healthiest cities generally had *less* tree canopy coverage than the least healthy.

Another surprising finding was that binge drinking, considered a risk factor by the CDC, was generally correlated with better health and better socioeconomic performance. Binge drinking correlated in the opposite direction from other health indicators in relation to nearly every socioeconomic indicator, and was most strongly and positively correlated with walking to work. It is possible that the strong correlation with walking may indicate that walking to work explains the positive correlations with health outcomes. Although violence is attributed to binge drinking by the CDC, binge drinking has a significant negative correlation with murder and burglary.

Among cities that ranked in the top or bottom 25 cities for each health indicator or risk factor, sixteen ranked in both the highest and lowest for different categories. Five of the seven cities with the highest cancer rates had the lowest rates of obesity and five had the lowest percentage of people sleeping less than seven hours per night. Five of the eight cities with the highest rates of diabetes had the lowest rates of asthma, and three of the five cities that had among the highest rates of both diabetes and obesity had the lowest rates of cancer.

While state and local governments in larger cities are more efficient with infrastructure per capita, these results suggest that they are less "efficient" with providing social, welfare and health services. There should be no assumption that infrastructure costs offset by scale-based translate into more cost-effective provision of other services. The health impact of state and local government spending per capita does not depend on population size.

Regarding the expected correlation between tree cover and socioeconomic output, the results are ambiguous. When broken down by ecoregion, some significant correlations between socioeconomic indicators and trees could be seen, but there were no correlations that were similar in all regions. The most consistent correlation was with obesity, which was significantly and negatively correlated in three regions, followed by high blood pressure, sleeping less than seven hours a night and physical health. Binge drinking was positively correlated in two regions, and the transition ecoregion between the woodlands of the east and the desert and plains of the Midwest showed the greatest number of significant correlations with health indicators as well as with socioeconomic indicators. At a national level, mental health was significantly correlated with tree canopy, but was only significantly correlated in the mountain west when data were separated by ecoregion.

Impervious surfaces also show a similar degree of regional disparity, although there would be no reason to assume that ecoregion would have any relationship with impervious surface, the differences do suggest that the relationship between health and impervious surface is influenced by cultural, demographic or regional variables.

Overall, correlations between socioeconomic indicators and tree cover are very weak. Even when historical median household income and scale-adjusted median household income from 1990 were tested for the effect of historical wealth or poverty on tree cover 21 years later, each showed significant correlations in only one region. The Pacific Northwest showed the greatest number of significant correlations with tree canopy, but it also had a very small sample size.

One likely explanation for the regional differences in how tree canopy and impervious surface relate to health is that the effect is more psychogenic than related to the specific ecosystem services provided by trees or the ecological impacts of impervious surface. That is, the effect of tree canopy on health is related to regional or cultural attitudes or views of trees which differ from place to place. For example, heat island effects, caused by impervious surface and mitigated by tree canopy, would be expected to exacerbate respiratory disease. More tree cover would both reduce the heat island effect and remove pollutants from the air, yet COPD is only significantly correlated with tree canopy in the transition zone and with impervious surface in the arid/desert and Pacific Northwest zones. And air quality and tree canopy show no correlation.

It has been claimed that one of the benefits of trees is that they increase connections between people, improve social cohesion and foster social capital (Holtan 2015). If this is true, scaling theory would predict that more tree cover would correlate with increased economic productivity, yet this is not visible in the data.

The influence of trees on health and their relationship with economic performance may be weak and inconsistent across regions because it is a weak influence and its influence is drowned out in the complexity of other influences, and the nuance of cultural, regional or other idiosyncratic conditions that affect how people use or perceive green space may mean that a simple measure of average canopy is an inadequate measure against which to compare health results or economic performance.

COPD was significantly influenced by driving alone to work and median household income. It was calculated that a one percent change in driving alone to work has the same effect on COPD as a change in scale-adjusted income of $2631, or about half a percent of average MHI.

One factor that influences the health benefits of trees is how people use green space in a city. The use of green space has been shown to be sensitive to perceptions of crime and safety. A multiple regression showed that, when factoring in crime and the SAMI for median household income, tree cover was significantly predictive of mental health, but crime had a greater influence on high blood pressure. Both crime and tree canopy exerted an equal influence on sleep.

Whether or not scaling theory has meaningful descriptive value, metropolitan socioeconomic performance on certain indicators has been shown to correlate with health outcomes nationwide, and with tree canopy in different ways in each ecoregion.

References Cited

1. Alexander, C. (1987). A new theory of urban design. New York: Oxford University Press.

2. Astell-Burt, T., Mitchell, R., & Hartig, T., University of Western Sydney, (2014). *The association between green space and mental health varies across the lifecourse : a longitudinal study*. U.K., B M J Group.

3. Batty, M. (2013). The new science of cities. Cambridge, Massachusetts : MIT Press.

4. Bettencourt, L. (January 01, 2014). The Kind of Problem a City Is: New Perspectives on the Nature of Cities from Complex Systems Theory.

5. Bettencourt, L. (2013 A). The Origins of Scaling in Cities. Science, 340(6139), 1438-1441.

6. Bettencourt, L. (2013 B). Urban truths Luis Bettencourt explains the hidden laws that govern our cities. (December 14, 2013). *New Scientist*, 2947, 30.

7. Bettencourt, L., Lobo, J., Strumsky, D., & West, G. (2010). Urban Scaling and Its Deviations: Revealing the Structure of Wealth, Innovation and Crime across Cities. Plos One, 5(11)

8. Bettencourt LMA, Lobo J, Helbing D, Kühnert C, West GB (2007) Growth, Innovation, Scaling, and the Pace of Life in Cities. Proc Natl Acad Sci U S A 104: 7301–7306.

9. Bettencourt, L., Lobo, J., Strumsky, D., & West, G. (2010). Urban Scaling and Its Deviations: Revealing the Structure of Wealth, Innovation and Crime across Cities. Plos One, 5(11)

10. Brinkman, J. C. (2016). *Congestion, agglomeration, and the structure of cities*. Philadelphia, PA: Research Department, Federal Reserve Bank of Philadelphia.

11. Bureau of Economic Analysis. Gross Domestic Product by Metropolitan Area; State and Local Government Spending by Metropolitan Area; Federal Civilian Non-Military Spending by Metropolitan Area.

12. Business Insights: Essentials. (4-21-2017). Study findings from Department of Pulmonary Medicine broaden understanding of Chronic Obstructive Pulmonary Disease (prevalence and etiological profile of Chronic Obstructive Pulmonary Disease in nonsmokers). Health and Medicine Week: 10628. Accessed 10 Apr. 2018.

13. Camacho-Rivera, M., Kawachi, I., Subramanian, S. V., & Bennett, G. G. (January 23, 2014). Perceptions of neighborhood safety and asthma among children and adolescents in Los Angeles: A multilevel analysis. *Plos One, 9,* 1.)

14. Center for Disease Control. (2017) 500 Cities: Local Data for Better Health, 2017. <>. Accessed 3/6/18.

15. Center for Disease Control. Behavioral Risk Factor Surveilance System, 2012 Release.

16. Center for Disease Control. (2018). Fact sheets-Binge Drinking.

17. Cervero, R. (2000). Efficient Urbanization: Economic Performance and the Shape of the Metropolis, Lincoln Institute of Land Policy Working Paper.

18. Chen, D. T. (2000). The Science of Smart Growth. *SCIENTIFIC AMERICAN -AMERICAN EDITION-*, (6). 84.

19. Cheng ER, Kindig DA. Disparities in premature mortality between high- and low-income US counties. Prev Chronic Dis 2012;9:110120. DOI: .

20. City of Gallup (August 22, 2017). Land Development Standards, Diagnostic Report. Dekker Perich Sabatini and Bohannan Huston.

21. Clark, I. I. W. (2016). Smart Green Cities. Taylor and Francis. Print.

22. Cookson, G. and B. Pishue, "Inrix global traffic scorecard," 2017.

23. Cremeens, J.L. (2009) Sociodemographic Differences in Binge Drinking Among Adults --- 14 States, 2004. Center for Disease Control, MMWR. April 3, 2009 / 58(12);301-304 4/2/2009

24. Dadvand, P., Wright, J., Martinez, D., Basagaña, X., McEachan, R. R. C., Cirach, M., Gidlow, C. J., ... Nieuwenhuijsen, M. J. (October 01, 2014). Inequality, green spaces, and pregnant women: Roles of ethnicity and individual and neighbourhood socioeconomic status. *Environment International, 71*, 101-108.

25. Environmental Protection Agency. Air Quality Survey. Annual AQI by CBSA.

26. Environmental Protection Agency. 2018. Heat Island Effect. www.epa.gov/heatisland/index.htm. Accessed March 5, 2018.

27. Environmental Protection Agency. Smart Location Mapping. National Walkability Index. (Sources: 2010 Census TIGER/Line shapefiles, 2010 Census Summary File 1, Census LEHD 2010, InfoUSA 2011, NAVTEQ NAVSTREETS 2011, GTFS data for 228 transit agencies, Center for Transit Oriented Development TOD Database 2012). https://www.epa.gov/smartgrowth/smart-location-mapping#walkability

28. Dalton, S.O., Schuez, J., Engholm, G., Johansen, C., Kjaer, S.K., Steding Jessen, M., Olsen S., Olsen, J.H. (2008). Social inequality in incidence of and survival from cancer in a population-based study in Denmark, 1994-2003: Summary of findings. European Journal of Cancer 44(14):2074. Elsevier Ltd 20080901. 0959-804

29. Deming Luna County Business Resource Guide. (2010). Deming Luna County Economic Development, Inc. The Arrowhead Center.

30. Ding, D., Gebel, K., Phongsavan, P., Bauman, A. E., & Merom, D. (2014). Driving: A Road to Unhealthy Lifestyles and Poor Health Outcomes. *PLoS ONE, 9*(6), e94602.

31. Downs, A. (May 29, 2003). Growth Management, Smart Growth, and Affordable Housing. Brookings.edu.

32. Eng PM, Rimm EB, Fitzmaurice G, Kawachi I. (2002). Social ties and change in social ties in relation to subsequent total and cause-specific mortality and coronary heart disease incidence in men. Am J Epidemiol. 2002;155(8):700–709 []

33. Ewing, R. & Cervero, R. (2010) Travel and the Built Environment. Journal of the American Planning Association, 76:3, 265-294, DOI: 10.1080/01944361003766766

34. Fan AZ, Strasser SM, Zhang X, Fang J, Crawford CG. (2015). State Socioeconomic Indicators and Self-Reported Hypertension Among US Adults, 2011 Behavioral Risk Factor Surveillance System. Prev Chronic Dis 2015;12:140353. DOI:.

35. Federal Bureau of Investigation, Uniform Crime Reporting Program. Crime in the United States 2012.

36. Fragkias, M., Lobo, J., Strumsky, D., & Seto, K. (2013). Does Size Matter? Scaling of CO_2 Emissions and US Urban Areas. Plos One, 8(6),

37. Fujita, M. (2002). Economies of Agglomeration, Cities, Industrial Location and Regional Growth. Cambridge, UK. Cambridge University Press.

38. Gaffield, S. J., R. L. Goo, L. A. Richards, and R. J. Jackson. 2003.. *American Journal of Public Health* 93(9):1527–1533.

39. Galea S, Tracy M, Hoggatt KJ, Dimaggio C, Karpati A. (2011). Estimated deaths attributable to social factors in the United States. Am J Public Health 2011;101(8):1456–65.

40. Gerrish, E., & Watkins, S. L. (February 01, 2018). The relationship between urban forests and income: A meta-analysis. *Landscape and Urban Planning, 170,* 293-308

41. Glaeser, E. (2008). Cities Agglomeration and Spatial Equilibrium. New York. Oxford University Press.

42. Graziella, F., H, R. A., Richard, A., Gary, F., Inga, M., & Heather, W. (December 01, 2014). Traffic-related pollution and asthma prevalence in children. Quantification of associations with nitrogen dioxide. *Air Quality, Atmosphere & Health, 7,* 4, 459-466.

43. Grace-Martin, K. (2000A). StatNews #39: Interpreting Regression Coefficients. Cornell Statistical Consulting Unit. Cornell University.

44. Grace-Martin, K. (2000B). StatNews #40: Interpreting Interactions in Regression. Cornell Statistical Consulting Unit. Cornell University.

45. Guarnieri, M., & Balmes, J. R. (May 03, 2014). Outdoor air pollution and asthma. *The Lancet, 383,* 9928, 1581-1592.

46. Hirabayashi, S., & Nowak, D. J. (August 01, 2016). Comprehensive national database of tree effects on air quality and human health in the United States. *Environmental Pollution, 215,* 48-57.

47. Holtan, M. T. (2012). *Social life under cover: Tree canopy and neighborhood social connections*.

48. Horn, C. A City at a Crossroads. American Forests Magazine, Winter/Spring 2016. Accessed 2-2-2018. <>

49. House, J. S., Landis, K. R., & Umberson, D. (1989). Social relationships and health. *Economic Outlook USA*, (4). 6.

50. Jacobs, J. (1984). *Cities and the Wealth of Nations: Principles of Economic Life*. New York: Random House.

51. Jay, M., & Schraml, U. (January 01, 2009). Understanding the role of urban forests for migrants - uses, perception and integrative potential. *Urban Forestry & Urban Greening, 8,* 4, 283-294.

52. Johnston, M., Percival, G., Great Britain., Institute of Chartered Foresters (Great Britain), & Great Britain. (2012). *Trees, people and the built environment: Proceedings of the Urban Trees Research Conference 13-14 April, 2011*. Edinburgh: Forestry Commission.

53. Jordan, R. (2013). Poverty's Toll on Mental Health. Urban Institute.

54. Kneebone, E., & Berube, A. (2013). *Confronting Suburban Poverty in America*. Washington, D.C.: Brookings Institution Press.

55. Kooshian, C., Winkelman, S. (2011). Growing Wealthier: Smart Growth, Climate Change and Prosperity. Center for Clean Air Policy. 01329688.

56. Kylstra, C. (February 26, 2014). 10 Things Your Commute Does to Your Body. Women's Health. Time Health. <http://time.com/9912/10-things-your-commute-does-to-your-body/ > Accessed 4/14/18.

57. Levine, J. A. Diabetes. 2011 Nov; 60(11): 2667–2668. Published online 2011 Oct 17. doi: 10.2337/db11-1118. PMCID: PMC3198075

58. Levinson, D. (2012). Network Structure and City Size. PLoS ONE, 7(1), e29721. http://dx.doi.org/10.1371/journal.pone.0029721

59. Lovasi,G.S., J.W. Quinn, K.M.Neckerman, M.S. Perzanowski, and A. Rundle. 2008. Children living in areas with more street trees have lower prevalence of asthma. J. Epidemiol. Community Health 62: 647–649.

60. Louf, R. & Barthelemy, M. (2014). How congestion shapes cities: from mobility patterns to scaling. Sci. Rep. 4, 5561; DOI:10.1038/srep05561

61. Maas, J., Verheij, R.A., Groenwegen, P.P., de Vries, S., Spreeuwenberg, S. (2006).Green space, urbanity, and health: how strong is the relation? Journal of Epidemiology and Community Health (1979-), Vol. 60, No. 7 (July 2006),pp. 587-592. BMJ Stable URL: Accessed: 25-04-2018 19:43 UTC

62. McCormick, R. (November 01, 2017). Does Access to Green Space Impact the Mental Well-being of Children: A Systematic Review. *Journal of Pediatric Nursing*, *37,* 3-7.

63. Mills, G.B. and Zhang, S. (2013). Social Support Networks and Their Effects on Hardship Avoidance among Low-Income Households. Urban Institute.

64. Milman, O. (2017, September 14). US people of color still more likely to be exposed to pollution than white people. The Guardian. <>

65. Morré, S.K. (2014). Connecting Local Values and Forest Policy to the International Forest Sustainability Dialogue. A Dissertation submitted to Oregon State University, May 15, 2014.

66. Murali, V. and Oyebode, F. (2004). Poverty, social inequality and mental health. Advances in Psychiatric Treatment. Vol. 10, 216-224.

67. Nowak, D. J., Rowntree, R. A., McPherson, E. G., Sisinni, S. M., Kerkmann, E. R., & Stevens, J. C. (January 01, 1996). Measuring and analyzing urban tree cover. *Landscape and Urban Planning, 36,* 1, 49-57.

68. Nowak, D. J., & Greenfield, E. J. (July 01, 2012). Tree and impervious cover in the United States. *Landscape and Urban Planning, 107,* 1, 21-30.

69. Oliveira, E. A., Andrade, J. S., & Makse, H. A. (May 28, 2015). Large cities are less green. *Scientific Reports, 4,* 1.

70. Olpadwala, P. (2017). (Scaling and) The Systemic Secrets of Cities [Draft].

71. Ordoñez, C. (2017). Public values associated with urban forests: Synthesis of findings and lessons learned from emerging methods and cross-cultural case studies. Urban Forestry and Urban Greening 25:74. 1618-8667

72. Ortman, S.G., Andrew H. F. Cabaniss, Jennie O. Sturm, Luís M. A. Bettencourt. (2015) Settlement scaling and increasing returns in an ancient society. Science Advances 20 Feb 2015: e1400066

73. Pataki DE, Carreiro MM, Cherrier J, Grulke NE, Jennings V, Pincetl S, et al. (2011). Coupling biogeochemical cycles in urban environments: ecosystem services, green solutions, and misconceptions. Front Ecol Environ. 9: 27–36.

74. Prakash, M., Tekoz, K., Sachs, J., Shank, M. and Schmidt-Traub, G. (2017). Achieving a Sustainable Urban America, The U.S. Cities Sustainable Development Goals Index 2017. Sustainable Development Solutions Network.

75. Prud'homme, R. and Chang-woon, L. (1999). Size and Urban Efficiency of Cities. Carfax Publishing. Urban Studies, Vol 36, no. 11, 1849-1858.

76. Rao, P.S., A.G. Gavane, S.S. Ankam, M.F. Ansari, V.I. Pandit, and P. Nema. 2004. Performance evaluation of a green belt in a petroleum refinery: A case study. Ecol. Eng. 23:77–84.

77. Robinette, J.W., Charles, S.T. & Gruenewald, T.L. J Community Health (2017) 42: 865.

78. Roman, C.G. and Moore, G. (2004). Measuring Local Institutions and Organizations: The Role of Community Institutional Capacity in Social Capital. Urban Institute, Justice Policy Center.

79. Rona, R. J. (2000) Asthma and Poverty. Thorax. 55:239-244.

80. Saunders, D. (2011). Arrival city. London: Windmill. Print.

81. Schlaepfer, M. Luis M. A. Bettencourt, Sebastian Grauwin, Mathias Raschke, Rob Claxton, Zbigniew Smoreda, Geoffrey B. West and Carlo Ratti. (9 Jul 2014) The scaling of human interactions with city size. arXiv:1210.5215v3 [physics.soc-ph].

82. Schoner, J. E., & Levinson, D. M. (November 01, 2014). The missing link: bicycle infrastructure networks and ridership in 74 US cities. *Transportation, 41,* 6, 1187-1204.

83. Schwarz, K., Fragkias, M., Boone, C. G., Childers, D., Zhou, W., McHale, M., Grove, J. M., ... Cadenasso, M. L. (April 01, 2015). Trees grow on money: Urban tree canopy cover and environmental justice. *Plos One, 10,* 4.)

84. Shaw KM, Theis KA, Self-Brown S, Roblin DW, Barker L. (2016). Chronic Disease Disparities by County Economic Status and Metropolitan Classification, Behavioral Risk Factor Surveillance System, 2013. Prev Chronic Dis 2016;13:160088. DOI:.

85. Speck, Jeff. Walkable City. Farrar, Straus and Giroux. New York, 2012. Print.

86. Storper, M. (2013). *Keys to the city: How economics, institutions, social interactions, and politics shape development*.

87. Strano E, Sood V (2016) Rich and Poor Cities in Europe. An Urban Scaling Approach to Mapping the European Economic Transition. PLoS ONE 11(8): e0159465. doi:10.1371/journal.pone.0159465

88. Transportation Research Board (2009) Special Report 298: Driving and the Built Environment, The Effects of Compact Development on Motorized Travel, Energy Use, and CO_2 Emissions.

89. Ulmer, J. M., Wolf, K. L., Backman, D. R., Tretheway, R. L., Blain, C. J. A., O'Neil-Dunne, J. P. M., & Frank, L. D. (November 01, 2016). Multiple health benefits of urban tree canopy: The mounting evidence for a green prescription. *Health and Place, 42*, 54-62.

90. UN Habitat (2004). Urban Patterns for a Green Economy, Leveraging Density. UNON, Publishing Services Section, Nairobi, ISO14001:2004-certified

91. UN Habitat (2014). Urban Patterns for a Green Economy, Clustering for Competitiveness. United Nations Human Settlements Program. UNON, Publishing Services Section, Nairobi ISO14001:2004-certified

92. US Bureau of the Census, Survey of State and Local Government Finance, 1977-2015, accessed via the Urban-Brookings Tax Policy Center Data Query System, October 16, 2017

93. United States Department of Agriculture Forest Service Northern Research Station Unit 08. (2011) Urban and Community Forests of the Continental United States. (accessed 3/25/18).

94. Valletta, R.(2011). Involuntary Part-Time Work: Yes, It's Here to Stay. Federal Reserve Bank of San Francisco. SF FED Blog.

95. *Van der See, R. (May 5, 2015). How Amsterdam became the bicycle capital of the world. The Guardian.*

96. Wagai, J. (2016). The Street Connectivity Index (SCI) of six municipalities in Jalisco State, Mexico. United Nations Human Settlements Program (UN Habitat). Nairobi, Kenya.

97. West, Geoffrey B. Scale, the universal laws of growth, innovation, sustainability and the pace of life in organisms, cities, economies and companies. Penguin Press. New York, 2017.

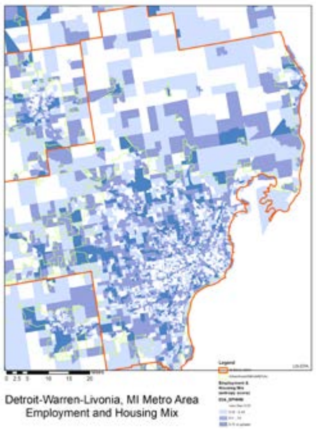
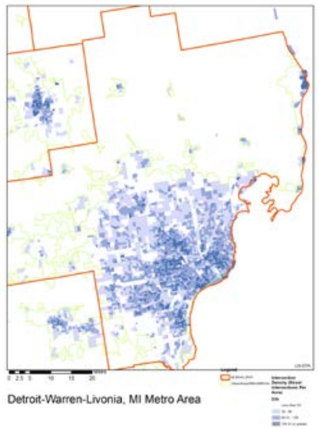
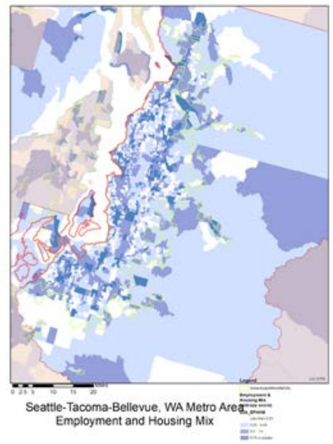
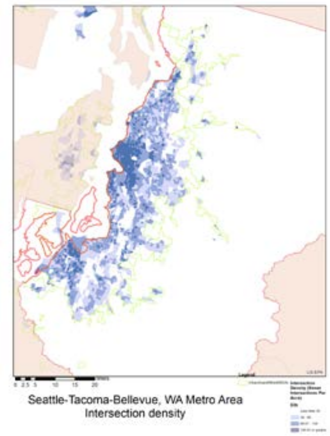
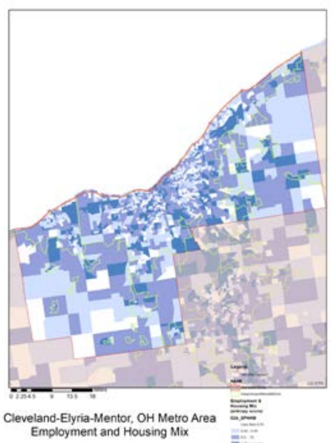
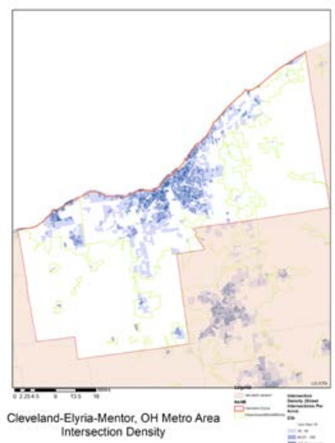
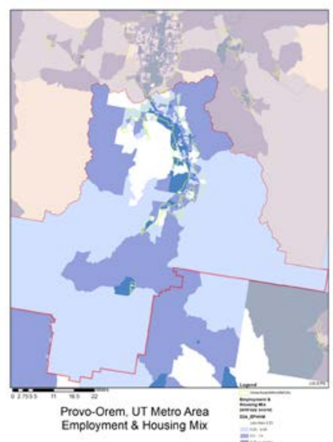
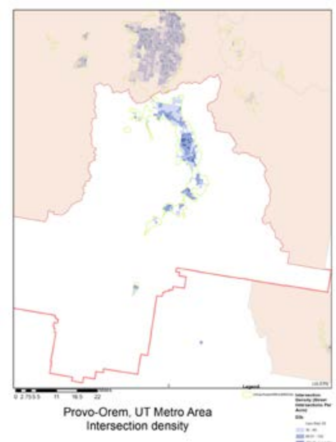
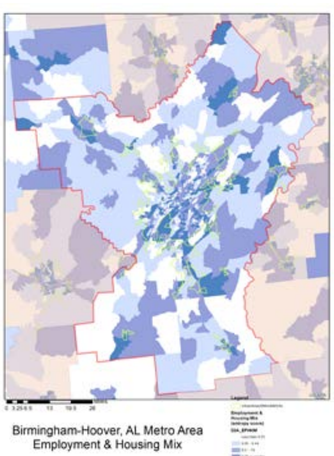
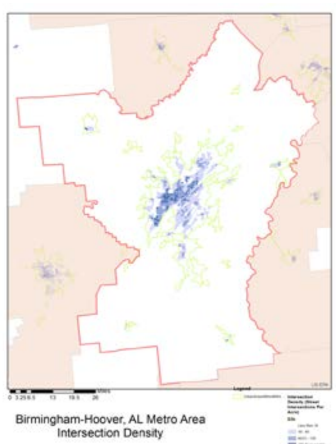
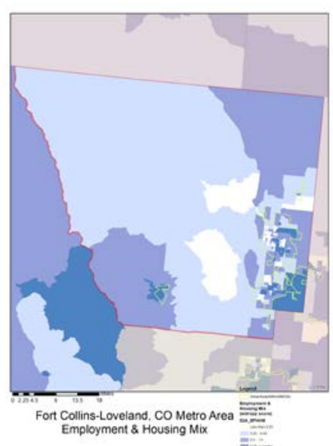
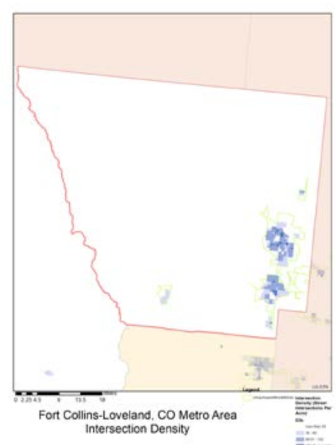
98. World Economic Forum. (2016). Inspiring Future Cities and Urban Services, Shaping the Future of Urban Development and Services Initiative. World Economic Forum, 2015. Ref: 250815.

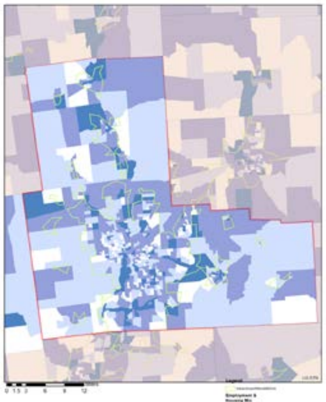
99. WHO (2016). Urban Green Spaces and Health – A Review of Evidence. World Health Organization, European Regional Office, Copenhagen, Denmark.

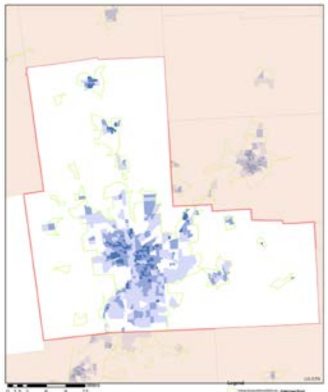
100. WHO (2018). Mental Health, Poverty and Development. Mental health policy and service development. Department of Mental Health and Substance Abuse, World Health Organization Geneva.

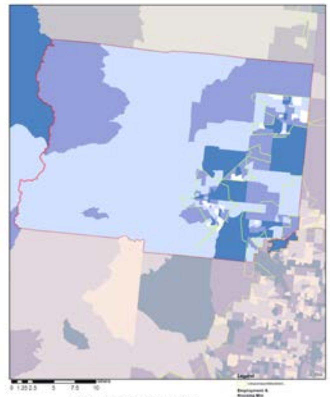
Appendix

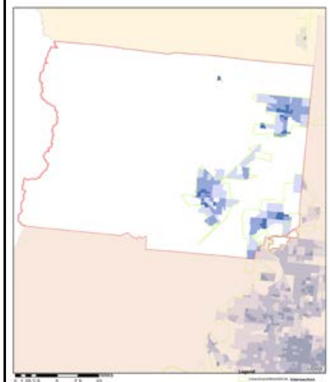
A.

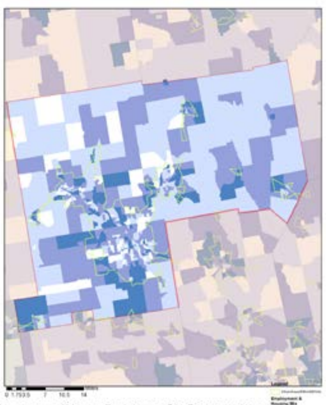
Maps of healthiest and least healthy cities showing Intersection Density and Employment-Housing Mix			
Least Healthy Employment-Housing Mix	**Least Healthy** Street Intersection Density	**Most Healthy** Employment-Housing Mix	**Most Healthy** Street Intersection Density
			
			
			

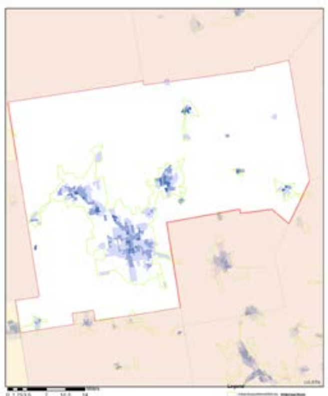

Dayton, OH Metro Area
Employment & Housing Mix

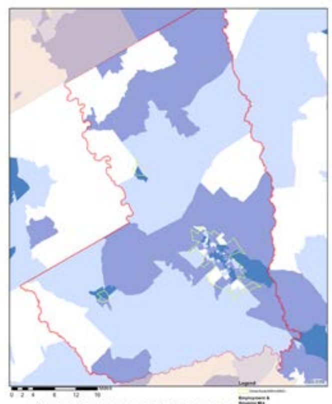

Dayton, OH Metro Area
Intersection Density

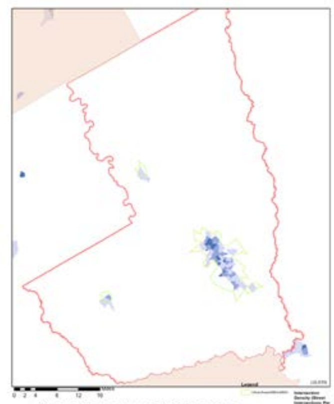

Boulder, CO Metro Area
Employment & Housing Mix

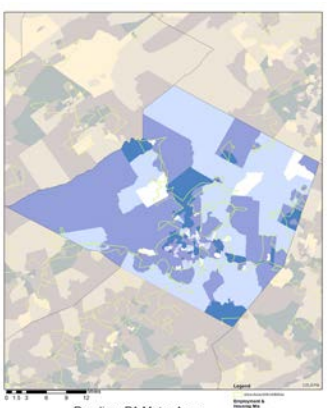

Boulder, CO Metro Area
Intersection Density


Youngstown-Warren-Boardman, OH-PA Metro Area
Employment & Housing Mix

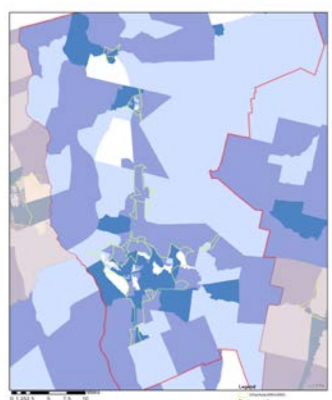

Youngstown-Warren-Boardman, OH-PA Metro Area
Intersection Density

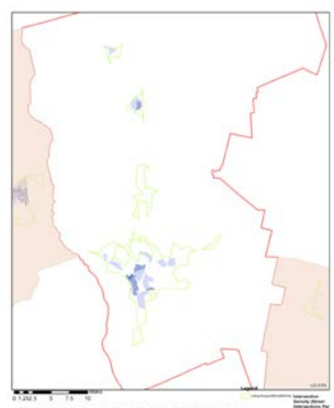

College Station-Bryan, TX Metro Area
Employment & Housing Mix

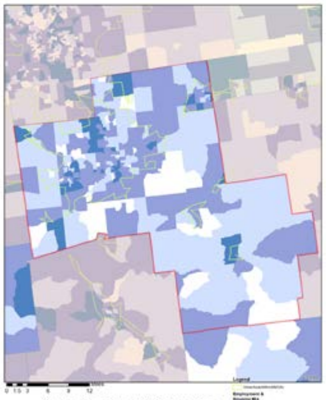

College Station-Bryan, TX Metro Area
Intersection Density


Reading, PA Metro Area
Employment Housing Mix

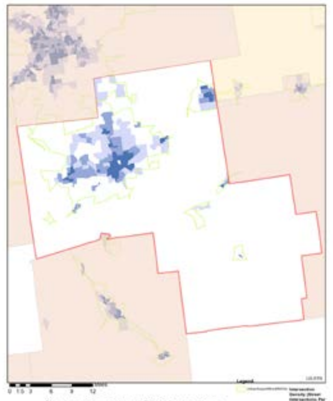

Reading, PA Metro Area
Intersection Density


Burlington-South Burlington, VT Metro Area
Employment and Housing Mix

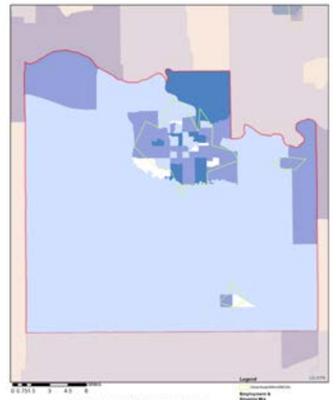

Burlington-South Burlington, VT Metro Area
Intersection Density



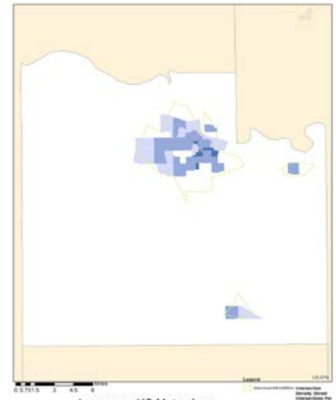
Canton-Massillon, OH Metro Area
Employment & Housing Mix



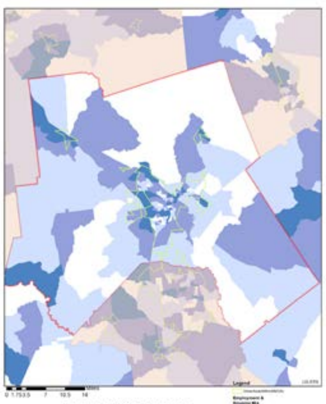
Canton-Massillon, OH Metro Area
Intersection Density



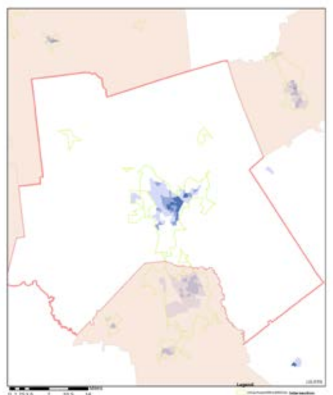
Lawrence, KS Metro Area
Employment and Housing Mix



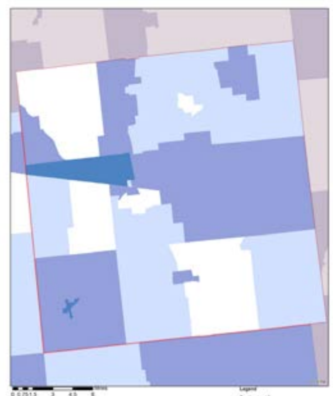
Lawrence, KS Metro Area
Intersection Density



Macon, GA Metro Area
Employment & Housing Mix

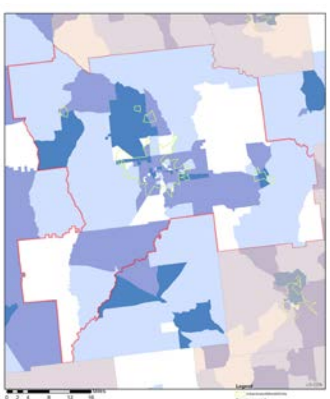


Macon, GA Metro Area
Intersection Density

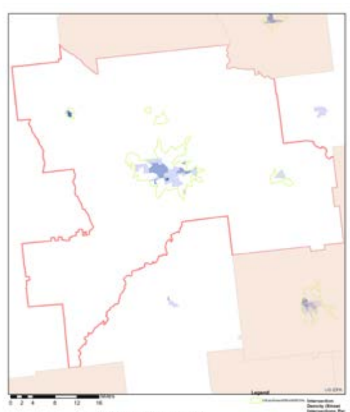


Plymouth, IN Metro Area
Employment Housing Mix

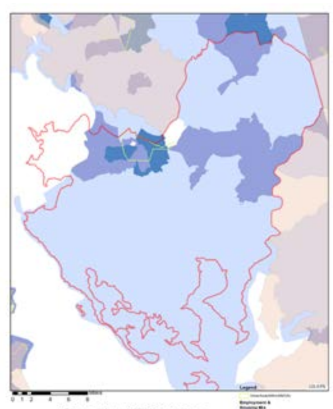
Plymouth, IN Metro Area
Intersection Density



Albany, GA Metro Area
Employment and Housing Mix



Albany, GA Metro Area
Intersection Density



Cambridge, MD Metro Area
Employment and Housing Mix

Cambridge, MD Metro Area
Intersection Density

B.
SAMI distribution of all MMSAs with labels on New Mexico MMSAs.