Cities as Six-By-Six-Mile Squares

Thomas J. Holmes Sanghoon Lee

November 28, 2007

Preliminary Draft for NBER Agglomeration Conference

1 Introduction

The urban economics literature examining the size distribution of cities generally takes the data as it comes from the Census Bureau, which typically bases its reporting units on arbitrary political boundaries. The Census Bureau has numerous constituents it must satisfy with its reporting. It is unlikely that in its reporting of city sizes, the Census Bureau would place high priority on what would be best for urban economists undertaking scientific study of the size distribution.

Scientists examining insect populations construct squares in fields and count bugs in samples of squares (Beall (1939)). In this paper we partition much of the United States into a grid of squares and count people in the squares. We count other measures of economic activity in the squares as well.

Rather than draw our own grid, we tie our hands by using the grid pinned down in the early 1800s by the Public Land Survey System (PLSS). As the United States expanded beyond the thirteen original colonies, the federal government assumed ownership of the additional lands. For the purpose of selling it, the new lands were partitioned into a relatively uniform grid of six-by-six-mile squares called *townships*. It is important to note that this grid was laid out essentially before the areas were settled by the new inhabitants, as land sales were a precondition to settlement. Figure 1 illustrates the states that are part of the PLSS grid. Figure 2 provides a close-up of a particular state (North Dakota) and this figure makes clear that the PLSS is a relatively clean grid of squares. Note, the original colonies were not surveyed because the Federal Government never owned this land. Likewise, Texas was not surveyed because the state government of Texas retained land ownership as part of the agreement of annexation. See Linklater (2003) for more about the PLSS.

The existing literature on the city size distribution using Census-based definitions of cities has established that the distribution is Pareto, at least in the right tail, and that the Pareto coefficient is remarkably close to one. This is Zipf's law and is often called the rank-size rule. See Gabaix (1999) and for a survey Gabaix and Ioannides (2004) The papers in this literature typically focus only on urban areas, with the smallest cities and rural areas left out. In our analysis of square townships, we look at all of the distribution, including the most remote areas. We determine the distribution of township sizes within groups of states, within individual states, and even within some metropolitan areas. The distribution of population across townships within these sets is certainly not Pareto over the entire range of township sizes. However, we do find regularities. In particular, the distribution is *piecewise* Pareto with two kinks. The first kink occurs at a log population density of 2.7 (15 people per square mile) and the second kink at log population density of 7.4 (1,500 people per square mile). (We uses logs as units in discussion because patterns are clearer when we think in these terms.) Between kink one and kink two, the slope is less is approximately .8, not Zipf's law, where the slope is one, but perhaps not too far off. So in this range things seem as least qualitatively similar to standard results in the literature. Below the first kink, the slope is very small (approximately .25) and above the second kink the slope is quite big (approximately 2). This basic pattern is quite robust in the data. It holds across the entire population of townships and it holds as well within individual states and even within large metropolitan areas. An analogous pattern holds for employment.

Kink one is at an extremely low population density. Places below this kink are extremely isolated and only a very small percent of the population lives in such places. We don't investigate why such a kink might exist but it does not surprise us that when things get so isolated that we might bang into some kind of corner where whatever forces are at work above this corner no longer apply. Kink two is much more interesting because 39 percent of the population resides there. Our paper sheds some light on what is going on here by relating our finding to results from the literature on the density gradient within urban areas. In most of the literature, variations in population across the country and variations within a given metro area are treated in different papers, but in this paper we do them both together.

An issue that is closely related to the size distribution of cities is the distribution of types of industries across cities. A property of the distribution of industries that is closely related to Zipf's law is the *Number Average Size Rule* (Mori, Nishikimi and Smith (2007)). We look at this property using townships as our geographic unit. We obtain similar results as Mori, Nishikimi and Smith. We also note that dartboard kinds of structures (Ellison and Glaeser (1997)) can deliver this empirical property.

1.1 Why Squares?

The typical unit of analysis for examining the size distribution of cities in the United States is the Metropolitan Statistical Area (MSA). These are defined as having "At least one urbanized area of 50,000 or more population plus adjacent territory that has a high degree of social and economic integration as measured by commuting ties." (Office of Management and Budget, 2006, p.2.) MSAs are composed of counties. The smallest MSAs are typically just one county, the largest MSAs like New York are comprised of many counties. For the 2006 MSA definitions, about 83 percent of the population live in one of 393 MSAs and 1,092 counties out of 3,141 are part of MSAs.

There are several limitations in using MSAs to examine the size distribution of cities. The first limitation is simply that the county is an irregular building block. On the east coast, counties tend to be small, with average sizes being on the order of hundreds of square miles. But in the west, counties are huge, typically in the thousands of square miles. So a typical county classified as an MSA in the West will include areas in the county not integrated with other parts of the county (e.g. in commuting).

A second limitation is that the very largest MSAs are typically in dense areas where the counties used to define MSAs can form contiguous blocks with the counties of other MSAs. The boundary of the New York MSA meets the boundary of the Philadelphia MSA, which in turn touches up against the boundary of the Baltimore MSA. It inevitably is a tough call to decide where it is in New Jersey that Philadelphia ends and New York begins. And it is a tough call whether or not to treat Oakland and San Francisco as one unit. The federal government has rules to make these calls. But there is no reason to think these rules are the ones we would want to use to determine whether, for example, the rank size rule holds at the top of the distribution.

The third limitation of looking at MSAs is that it forces us to leave out of the analysis the significant fraction of the population who live outside of MSAs. The 2006 MSA definitions leave out 17 percent of the population; the earlier MSA definitions used in previous studies leave out 20 to 25 percent of the population.

Eeckhout (2004) has recently argued that by leaving out the bottom tail of the distribution, we draw erroneous conclusions about the shape of the city size distribution. With the Pareto distribution, if we look at any two city sizes, say 10,000 people and 20,000 people, there are always more of the smaller size than the larger. (In particular, if the Pareto coefficient is one, there are four times as many of the smaller city size as the twice as large city size.). With a log normal distribution, things are different. Yes, the distribution is skewed like Pareto, but the difference is there is a modal city size, such that if we go below the modal size, we start seeing fewer cities of a given size, not more.

Eeckhout argued that we should look at data on Census places rather than MSAs to examine the size distribution of cities as we would have a better shot of seeing what is happening at the bottom tail. A Census Place is something that is either legally incorporated or is some other place that the Census thinks would be useful to designate as a place. By looking at places, Eeckout picks up small towns that are not part of MSAs. But in limiting attention only to places, he loses people living in rural portions of MSA counties that are not considered places. In fact, only 74 percent of the 2000 population lives in a place. Eeckhout found that the distribution of place sizes is log normal. In other words, at the very bottom of the place distribution, as we decrease size, there are actually fewer places. Definitely not Pareto at the bottom.

Our problem with using place data is that the definition of a place is likely to be just as arbitrary as a MSA. We make our point by showing in Table 1 the very smallest places from the 2000 Census, those with population five or below. This is the data Eeckhout used. There happen to be only two places with a population of a one person, including Lost Springs, Wyoming, and two places with two people, including Hove Mobile Park City, North Dakota. By contrast, there are four places with three people. Looking at slightly bigger places, we find there are twelve places with eleven people. The modal place size is a population of 86 for which there are 31 places. This discussion gives a hint for why Eeckhout's found that a log normal distribution fit the distribution of places sizes relatively well. Clearly the frequency count is increasing at very small sizes. But it should also expose a concern. What exactly is a place with a population count equal one or two? Remember, 26 percent of the population, mostly rural, are not considered living in a place. If we can call Lost Springs a place, with its one resident, why can't we call some farm house with a family of five a place? And if we were free to do that, we might very well find more places with five people than what would otherwise be the modal place of 86 people. It would seem that having an official Census place with less than ten people might depend upon legal particulars about incorporation or perhaps the hutzpuh of the residents. But neither consideration is likely something we want in our analysis of city size distributions.

The Census has recently extended its data reporting to something called "Micropolitan

Statistical Areas" and these have been included in some studies of the size distribution. These are counties that contain a city in the 10,000 to 50,000 range. Basically, rural counties have complained that they are being left out in data products, so the Census has now added some more counties that are treated like MSAs for reporting purposes.[1] The Census adding Micropolitan Statistical Areas gives no new tool to economists looking at the size distribution of population.[2] Researchers can always choose to treat rural counties as cities, since population data is readily available at the county level. (But other data products besides population are not available at the county level so that is why getting classified as a Micropolitan area makes a difference). Given the crude and arbitrary nature of what a county is and the way that size varies across different parts of the country and even within a state, treating counties as a unit of analysis doesn't seem like a particularly good idea.

So why squares? By using the Public Land Survey System grid, we get to determine in a precise way the extent that population density varies across space. We don't have to depend upon decisions the Census makes as to what is or what is not a city. Most importantly, we can go as far down in the size distribution as we want; we can go to the most rural of rural. We never have to get into the issue of whether or not we should treat a farmhouse with five people as a place.

Why not just look at density? Ciccone and Hall, for example, argue the focus should be on density. In our analysis, with land the same across locations, with each township approximately 36 square miles, density is the same as population. So we are looking at density. One needs to be careful with the "bandwidth" over which one calculates density because the choice will affect the answers. We can imagine that if we used a very narrow bandwidth, say twenty-by-twenty-feet squares instead of six-by-six-mile squares, we would get very different answers. (At 20 feet, most parts of metro areas (roads, stores, backyards,

[1] The Census cannot simply extend the reporting even further to the remaining rural counties. With the small populations of the remaining rural counties, the cost per person for collecting information gets very high. (Large sample sizes are needed to get the law of large numbers working). And issues of disclosure of confidential information come into play low populated areas.

[2] There are actually a few micropolitan areas thare are composed of more than one county but there are not very many of them.

etc., would have population density zero.) Since bringing in rural areas is of particular interest for this study, we think a six-by-six-mile square is useful for our purposes.

Through use of our six-by-six mile grid, we are able to get an analysis of the distribution of population *within* metropolitan areas in the same way we look across the country as a whole. The largest metropolitan areas are typically made up of many townships. Chicago, for instance, is made up of 233 townships. This is more than the number of MSAs used in a typical analysis of the city size distribution. We find that the pattern for the United States as a whole continues to hold within individual metropolitan areas like Chicago.

We are not trying to argue that squares are the best way to look at the size distribution. Rather, our point is that the use of MSAs has particular limitations, especially, for looking at rural areas, so let's see what we get from using squares where we don't get these particular limitations. Not surprisingly, squares have their own limitations. In particular, one way that cities grow is through adding more land. In a model where each individual has inelastic demand for a unit of land, then the size distribution of cities is all about the size distribution of land and nothing about density. So this analysis leaves out something important. But to its credit, the analysis it clear about what's in and what's out.

Finally, why use townships rather than just draw an arbitrary grid of six-by-six mile squares and go from there? Actually, though we haven't tried it yet, we expect that our results would be very similar if we were to have created a six-by-six mile grid from scratch. The inspiration for this paper was seeing a map of the township grid which we found to be intriguing and it lead us to think about how population might be distributed these squares. Something good to be said about using the township grid is that it ties our hands in a first foray into these issues. When our hands are not tied and we get to pick, we need to determine both the grid size as well as where to anchor the grid. We expect that our results would change very little if we were to move the grid around, holding the size fixed. We expect that things would be very different if we changed the size.

1.2 Relation to the Literature

Given the centrality of the issue to the field, there is a large literature in urban economic on the distribution of activity across space. It has many branches.

One branch focuses on whether nor not the size distribution of cities satisfies Zipf's law and what theories might account for why Zipf's law would apply. Gabaix and Ioannides (2004) survey this literature. As discussed earlier, in this literature, populations vary across location units both because density increases and because the area of the units differs. Here we use six-by-six mile squares at location units.

A second literature looks at detailed data of where people live and work. See Anas, Arnott, and Small (1998) for a survey of some of this literature. Among the topics considered in this literature is how to identify clusters of activity from raw data collected at a fine level of geographic detail. The earliest literature focus on one cluster, the central business district. A later literature allows for the possitibility there may be multiple subcenters of employment within a metropolitan area and designs empirical techniques to find them (e.g. McMillen and McDonald (1998)). Our work is like this in that the raw data we use is at a very fine geographic scale. The difference is we aggregate this to a grid, while the other literature's main interest is identifying boundaries. More generally this literature focuses on what is going on within a city and our paper is different in discussing what is going on within metro areas and across the whole country in the same paper.

A third literature focuses on the fractal nature of the location of activity. See Battey and Longley (1994).

The empircal regularity of Zipfs law is more than just a curiosity. Large issues in urban economics are at play here. Justifiably, much work has been directed at explaining this regularity. Most of the literature has focused on random growth (Gabaix (1998)). More recently, Hsu (2007) has developed a formalization of central place theory which delivers a Pareto distribution as an implication. The implications of these two separate lines of research are very different from each other. Our analysis of squares looks at the issue from a different angle than has previously been done. The regularities we establishment about squares and kinks can potentially help us distinguish between alternative theories of cities.

2 The Public Land Survey and Population Data

2.1 Public Land Survey Data

There are 29 states in the Public Land Survey System (PLSS).[3] The first issue we need to deal with is that for some states, part of the land in the state is not part of the public land survey. Across the 29 states, 93.9 percent of the land is part of the survey. There are four main reasons why some land in these states is not part of the survey and these are listed in Table 2.

The first reason is that some of the land has not been surveyed (2.66 percent). This land is typically part of a National Park and has never been prepared for sale. This is important in states like Idaho and Montana, as can be seen in Table 2.

The second reason is when Federal government assumed ownership, in some cases there was land that had previously been granted that the Federal government did not take away. This category is called "Land Grants, Civil Colonies." This is particularly important in New Mexico and California. In these states there were important settlements before the Federal government took over. In these cases, we do not typically get six-by-six-mile squares. We are wary about leaving in parts of New Mexico and California that are surveyed and excluding the parts that are from the land grants because we would be selecting by which area was initially the most desirable and attracted people.

There remaining categories are "Indian Lands" and "Private Survey" (Much of Ohio is not part of the PLSS because it was surveyed before the PLSS system was set up.)

Table 2 sorts the states in descending order of the percent of the land in the state that is part of the PLSS.[4] Given our intent to provide a clean analysis, we want to focus on those

[3] We exclude Allaska.

[4] One issue is that townships cross bounderies in some cases. To construct Table 2, we allocated land to

states where virtually all of the land is part of the system so we have six-by-six-mile squares throughout. We draw the line at 99 percent coverage. Our focus will be on the 15 states for which the PLSS land accounts for 99 percent of the land (excluding land marked as water for the purposes of the survey.) We call these the 15 complete PLSS states. We also run an analysis where we include the 14 remaining PLSS states and it turns out not to make a difference.

There are 32,254 townships in the 15 complete PLSS states. Table 3 provides summary statistics. The mean township sizes is 33 square miles. There are a number of small townships at corners, but they don't take up much area. In most of what we do, we weight by area. When we weight by area, we see that the average township has 35.3 square miles and the distribution is tight around this point. The 10th percentile is 34.4 and the 90th percentile is 37.0.

2.2 Population

We use the 2000 Census of Population. The finest level of detail reported by the Census is at the level of the Census Block. There are 8 million Census Blocks, so it typically is very fine in geographic detail.

The Census does not provide polygon boundary files for blocks, just a longitude and latitude of a point in the block. We proceed under the fiction that all of the people in the block group live at the longitude and latitude reported by the Census. We then use GIS software to assign each block group to a township. The population of a township is the sum of the population across the Census blocks assigned to the township.

Table 3 shows the distribution of counts of number of block across townships. There are some townships with no blocks at all. The 10th percentile location has 6 blocks, the 25 percentile has 19 blocks.

In dense areas, blocks are a fine grid. So our procedure is a very good estimate for the

---

the state with the highest share of land in the township.

population within the township. In the most sparse areas, blocks are coarse. Our basic analysis will focus on those parts of the country with at least one person per square mile. These places will typically have many block groups, so the approximation is good.

We next calculate the population density of each township. The median township (weighting by area) has only 5 people per square mile. Define the variable *lnpopden* to be log of population density. Patterns are more clear when we discuss things in logs. For the median township, *lnpopden* equals 2.1. The maximum population density is in downtown Chicago equals 22,000 per square mile or 10.0 in logs.

The maximum population is obtained on a parcel that happens to be 15 square miles because part of the township square extends into Lake Michigan. Nonetheless, a square of 15 miles is the same order of magnitude as a full 36 square mile township. At a much smaller scale, the maximum population density tends to be much higher. If we were to look more narrowly at census blocks, in downtown Chicago they are defined at the level of buildings. The maximum density takes place in highrise building with 1,000 people using only 10,000 square feet of land. On a per mile basis the density is 2.9 million per square mile and log density is 14.9. With the large grid that we are using, we don't get *lnpopden* that high, 10 is where things stop.

2.3 Employment and Business Activity

We also look at the distribution of employment by place of work from the 2000 Census. This is available only at the tract level. (About 10 percent of employment cannot be classified at the track level and so is dropped here.) We convert tract-level data to township-level data as follows. For each track we proceed under the assumption that employment is uniformly distributed over the land within the track and in this way allocate employment by place of work to census blocks. We then aggregate the Census blocks to townships as we do for population.

The bottom of Table 3 contains the summary statistics for the employment variable.

Note that on average employment density is on the order of half of population density. Interestingly, employment is skewed. Whereas for most of the distribution, employment density is much lower, at the very top, employment density is actually higher than population density. This is the case for downtown Chicago where more people commute into work than live.

We have available data on business locations from the ReferenceUSA data set. This commercial source keeps track of 14 million business establishments in the United States and for each has the longitude and latitude, an estimate of employment, and the industry code. We use the geographic coordinates to assign establishments to townships. Our resources are limited in acquiring this data so we only look at two cases, North Dakota and the Chicago MSA. These polar extremes bracket the data.

3 Size Distribution of Squares

To describe the distribution of cities, the typical exercise is to sort the cities by descending size. Then the log of rank is regressed on the log of population. The coefficient on the log of population is Zipf coefficient, the parameter of the Pareto distribution. This is a *Zipf plot.*

We proceed in a similar fashion only we look at log population density. We sort the townships by descending *lnpopden.* We create a variable *cumulative land* which is the total land up to an including this township in the sorted list. Now if each township were exactly 36 square miles and if our units of total land were a 36 mile parcel, then cumulative land would exactly equal rank. It isn't exactly this because all townships are not exactly 36 miles. But it is very close. We take logs and get *lncumland*

Figure 3 is a *Zipf Plot* of *lncumland* on *lnpopden.* We are only plotting those points above *lnpopden*≥ 0, or equivalently one person per square mile. This figure is certainly not a straight line. But it looks like it could be well approximated by the piecewise linear function with two kinks. The first kink around when the log of population is a little less than 3. The second kink is just above log population of 7.

Table 4 fits a linear regression on this figure as well as a piecewise linear function with two kinks. We weight by area. The R^2 of the linear regression is .912. Allowing two kinks and using nonlinear least squares raises the R^2 to .998. The kinks are *lnpopden* equal to 2.74 and 7.37. This corresponds to population densities of approximately 15 per square mile and 1,500 per square mile. Now the townships between the kinks account for about 30 percent of the land and 57 percent of the population. Then townships above kink 2 account for about .7 percent of the land and 39 percent of the population.

Above the second kink, the slope is 2.1 in absolute value, which is much greater than the one in the rank-size rule. A high slope like 2.1 means the tail is not so fat, like with the rank-size rule. The dense areas drop off quickly. When one works with MSA data, slope coefficients as high as this are not obtained. The difference can be readily explained. The biggest cities get bigger on two margins, they add population density and they add area. But here area is fixed at 36 square miles. The top seven townships by density are all pieces of the Chicago metro area. There is not a rank size rule *within* metropolitan areas.

Between the first and second kink the slope is around .8 in absolute value. This is well less than 1. This is different from the rank-size rule, but now in the opposite direction. This is not inconsistent with previous findings because the township is a different unit. Among other things, there are pieces of Chicago in this range as well.

It has been noted before with MSA data that a Zipf plot has deviations for linearity at the end points giving it a concave shape. What is particular interesting here is the regularity about where the kinks are. We make this point in two ways. First, we redo this with a completely different set of townships, the ones for the 14 states where there is a portion of the state not part of the PLSS survey. When we do this in Table 4 we see we get results that are strikingly similar to the results with original 15 states. As before, the fit is not so good with just a linear relationship. Going to the piecewise linear with two kinks raises the R^2 to .9980. Moreover, the estimated locations of the kinks are very close to the kinks in the original estimate.

The second thing we do is be examine the distribution of townships within individual states. Table 5 reports the results at the state level that are analogous to Table 4. First we do a linear regression. For a few rural states, the fit of the linear regression is quite good. This is probably due to the fact that these rural states don't have much land above the second kink. We use the same kinks as estimated above (although the results are similar if we estimate separate kinks for each state). Across all fifteen states, the piecewise linear function at the given kinks fits extremely well. Moreover, at the top bracket the slope is well above one and in the bottom bracket the slope is quite small, generally less than a half. In the intermediate range, the slope is less than or equal to one.

Consider next the distribution of employment across squares. Figure 4 plots *lncumland* against *lnempden* in an analogous fashion as Figure 3. The pattern is similar. The relationship is flat until a kink around *lnempden* equal to around 7. Note downtown Chicago is an outlier here, in a way that is not the case for the population variable. Table 6 presents the analogous regression results to table 4. Substituting the employment variable for population makes little difference.

3.1 MSA-Level Analysis

Above kink 2, population density is above 1,500 per mile or 54,000 in the township. This is the population threshold for MSA status. Hence all of the townships above kink 2 are in MSAs. So to understand what is going on for this portion of the relationship, we need to know what is going on in MSAs. Now the very highest population density townships in Figure 3 are from Chicago. So to understand what is going on here, we should look at Chicago. So here we will look at the same figure as in Figure 3, but only for Chicago. We look at other large MSAs as well. Analogous to what we did with states, we only focus on MSAs where virtually all of the land in the MSA is part of the Public Survey System (We use a 99% threshold). We focus on the 11 MSAs in this set with a 2000 population above one million. These are listed in Table 6. Typically there are on the order of 200 townships for each of these cities.

Figure 5 contains Zipf plots for each of the 11 large MSAs. The plot is illusrated in red. (We explain the blue later.) We also illustrate lines in each figure where we get the two kinks. There are several interesting things to note here. First, the pattern found in Figure 3 when we combine all townships continues to hold in each MSA individually. Note in particular what happens at *lnpopden* equal to 7.4 which is kink 2 from the aggregate case and is highlighted here with a line. To a remarkable degree, there is a kink in each MSA at this point. It is though *lnpopden* is some kind of magic number where something new starts happening.

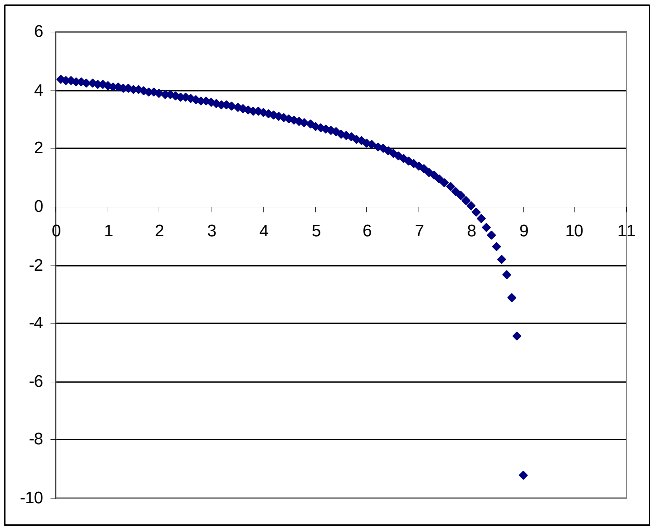
The second thing to note is that between the kink markers (which are set to be the same for all the MSAs), the Zipf function is relatively flat.

A third thing to note is the regularity that the maximum *lnpopden* for all MSAs is in the range of 8 to 9 with Chicago equaling 10 being an outlier. Now of course in levels this is a lot of variation. But the regularity of the maximum when taken in logs is intriguing. Note this regularity is dependent on the grid size. These cities are very different and are each made up of many townships and there is great variation within each MSA in *lnpopden.* But the maximum *lnpopden* varies very little.

In Table 7 we run the regressions for each individual MSA. Allowing the piecewise linear regression with the earlier kinks fits the data extremely well.

This discussion suggests that if we want to get a handle on why at the aggregate level there is a kink between 7 and 8, a good place to start is to try to understand why we might get such a kink within a metropolitan area. More generally, why might we get a concave Zipf plot within metro areas?

The most basic things we know in urban economics about the distribution of population *within* a metro area are: the monocentric model of the city as a theory and the density gradient as descriptive tool. In the analysis, there is some central location, call it c. This is



the location with maximum population density. Let x_i be the distance of township i from the center (or maximum population location). Typically the density gradient is specified in the following log linear form

$$lnpopden_i = \alpha - \gamma x_i$$

where α is the maximum density and γ is the density gradient. We can readily link this to a Zipf plot. Assuming $\gamma > 0$, the center location c has the highest density and so is the right-most point. Working in a continuous space for now (as opposed to dealing with squares), the cumulative area going out to a distance x is of course $A = \pi x^2$. So

$$\begin{aligned} \ln A &= \ln \pi + 2 \ln x \\ &= \ln \pi + 2 \ln \left(\frac{\alpha - lnpopden_i}{\gamma} \right) \\ &= \ln \pi + 2 \ln (\alpha - lnpopden_i) + 2 \ln \gamma \end{aligned}$$

Let is look at the function $\ln(\alpha - lnpopden)$. Setting $\alpha = 9$, the function looks as illusrated below.

A kink at *lnpopden*$= 8$ is readily evident. This function is not literally piecewise linear. But it is very close. More generally, whatever the maximum *lnpopden* α is, there is a kink at $\alpha - 1$. The MSAs have a maximum *lnpopden* in the range of 8 to 9, so the kink should be in the range of 7 to 8.

That is some theoretical reasoning to guide things. How well does the moncentric theory work with our data? For each of the 11 MSAs, we picked the center to be the maximum *lnpopden* of all townships and set its *lnpopden* to equal α. We used OLS to estimate the rent gradient γ. We took the fitted values of *lnpopden* for each MSA and created a Zipf plot from the fitted values. These are the blue lines in Figure 5. There is a strong correpondence between the blue and red lines. We take this as strong evidence that a density gradient of the kind considered here is the key explanation for the pattern we see. We don't expect things would turn out so cleanly if urban spatial patterns did not exhibit a monocentric property.

In looking at the theoretical figure above, the kink property is immediate. So if to have a metro area where the density gradient took a close approximation to the loglinear form, then we see that a Zipf plot would get a kink of the form above. What we find interesting is that when we aggregate the form is retained. It is important to note that the aggregate is not linear. Of course if we were to vertically sum these things than piecewise linear retains piecewise linear. But to the contrary, the aggregation is nonlinear and complicated. How things aggregate depend upon ratios of size types across cities.

4 The Number - Average Size (NAS) Rule

An issue that is closely related to the size distribution of cities is the distribution of types of industries across cities. A property of the distribution of industries that is closely related to Zipf's law is the *Number Average Size Rule* (Mori, Nishikimi and Smith (2007)). The NAS rule states that there is a strong negative log linear relationship between the number and the average population size of cities where an industry is found. For each industry, count the number of cities where the industry can be found and calculate the average size of these cities. If you plot the count and average size of the cities, there is a strong log linear relationship. Using the RefUSA data set we look at this property using townships as our geographic unit. As was in the Zipf's law, we use average township density instead of average city size in order to examine the NAS rule. In order to prevent small random outliers from driving our result we focus on industries which have presence in at least 20 townships.

Figures 6 and 7 show that the NAS rule holds quite generally: both for sparsely populated areas and densely populated areas (North Dakota and Chicago) and for different industry classification levels (NAICS 2 digit industries and 4 digit industries). The negative log linear pattern is strikingly clear. For example, in figures 6.a and 7.a, the slope coefficients for North Dakota and Chicago are -0.7 and -.065 with R^2s equal to 0.95 and 0.97. The NAS rule may not hold well for some industries though, like agriculture.

Mori, Nishikimi and Smith (2007) theoretically show that the central place theory and the Zipf's law lead to the NAS rule. The central place theory implies that large cities have the superset of industries: industries found in small cities can also be found in large cities. Thus, industry which exist in small number of cities are more likely to be in large cities and this lead to the negative relationship between the average number of the cities where an industry can be found and the average city size. The Zipf's law leads to the log linear relationship between them.

We find that this structure can be delivered by dartboard kinds of structures (Ellison and Glaeser (1997)). We randomly relocate all the establishments in Chicago with the probability of locating in a township proportional to its population size and examine the NAS rule. Figure 8 shows one realization of this simulation. This random process seems to generate the NAS rule better than the real data. For example, when we ran this simulation 100 times, the mean NAS rule regression coefficient for 2 digit industries was -0.79 (0.018) with mean R^2 equal to 0.99 (0.004).

However, this should not be taken as evidence against the role of the central place theory generating the NAS rule because the location pattern generated by this random process is also consistent with the central place theory. When we relocate establishments according to population size, the townships with more population are more likely to have more variety of industries.

References

Anas, Alex; Arnott, Richard; Small, Kenneth A..“Urban spatial structure” Journal of Economic Literature, Sep98, Vol. 36 Issue 3, p1426, 39p; (AN 1138346)

Geoffrey Beall, "Methods of Estimating the Population of Insects in a Field," *Biometrika*, Vol. 30, No. 3/4. (Jan., 1939), pp. 422-439.

Eeckhout, Jan (2004), Gibrat’s Law for (all) Cities", American Economic Review 94(5), 2004, 1429-1451.

Gabaix, Xavier. "Zipf’s Law for Cities: An Explanation." Quarterly Journal of Economics 114, no. 3 (August 1999): 739-67.

Gabaix, Xavier and Yannis M. Ioannides (2004), “The Evoluation of City Size Distributions," *Handbook of Urban and Regional Economics* Handbook of Regional and Urban Economics,

Linklater, Andro, *Measuring America: How the United States was Shaped by the Greatest Land Sale in History,* Plume Publishing: New York, 2003.

Table 1: Census Places
2000 Census Population Less Than Or Equal to Five

Place	Population
New Amsterdam town, IN	1
Lost Springs town, WY	1
Hove Mobile Park city, ND	2
Monowi village, NE	2
Hobart Bay CDP, AK	3
East Blythe CDP, CA	3
Hillsview town, SD	3
Point of Rocks CDP, WY	3
Flat CDP, AK	4
Blacksville CDP, GA	4
Prudhoe Bay CDP, AK	5
Storrie CDP, CA	5
Baker village, MO	5
Maza city, ND	5
Gross village, NE	5

Table 2
Land Classification in States Covered by PLSS

	Public Survey	Unsurveyed Area	Land Grants, Civil Colonies	Indian Lands	Private Survey
All	93.92	2.66	1.71	0.98	0.64
AR	100.00	.	.	.	.
IA	100.00	.	.	.	.
KS	100.00	.	.	.	.
MN	100.00	.	.	.	.
ND	100.00	.	.	.	.
NE	100.00	.	.	.	.
OK	100.00	.	.	.	.
SD	100.00	.	.	.	.
OR	99.99	0.01	.	.	.
WA	99.99	0.00	0.01	.	.
NV	99.77	0.23	.	.	.
WI	99.75	.	0.12	0.13	.
MI	99.53	.	0.47	.	.
IL	99.48	.	0.52	.	.
AL	99.18	0.01	0.81	.	.
UT	98.61	0.76	.	0.64	.
MS	98.43	0.01	1.57	.	.
IN	97.98	.	2.02	.	.
MO	97.10	0.07	2.83	.	.
CO	95.00	2.97	1.80	0.23	.
WY	92.22	7.78	.	.	.
MT	91.23	8.77	.	.	.
NM	88.57	0.97	9.38	0.64	.
FL	87.63	6.75	5.61	.	.
CA	87.34	3.81	8.84	.	.
ID	86.07	13.93	.	.	.
LA	84.98	7.66	7.37	.	.
AZ	71.73	8.50	0.43	18.07	.
OH	64.26	0.03	0.27	.	35.05

Table 3
Summary Statistics
15 Complete PLSS State
(32,254 townships)

	Mean	min	10%	25%	50%	75%	90%	max
Area in sq miles (unweighted)	33.0	0	23.0	35.1	35.9	36.3	36.9	58.2
Area in sq miles	35.3	0	34.4	35.4	36.0	36.4	37.0	58.2
Number of Census Blocks	81.7	0	6	19	45	84	152	4,241
Population Density (persons per square mile)	58.5	0	0	.6	5.0	22.6	78.0	22,012.6
log of population density	1.9	.	.	-.7	2.1	3.4	4.5	10.0
Employment Density (workers per square mile)	24.9	0	0	.2	.9	4.0	19.0	30,398.0
log of employment density	-.1	.	.	-3.1	-.0	1.4	3.0	10.3

Townships are weighted by area except for the first row where it is unweighted.

Table 4
Regression Results
Log Cumulative Area on Log Population Density
(All squares have Population Density >= 1 per square mile)

	15 States Where Virtually All Land is Part of Public Survey		14 Public Survey States with Land Not Part of the Survey	
	Linear	Piecewise Linear, Two Kinks	Linear	Piecewise Linear, Two Kinks
Constant	14.093 (.004)	13.5787 (.0008)	13.688 (.004)	13.2101 (.0009)
Slope (linear)	-.597 (.001)	.	-.5261 (.0014)	
Slope 1 (below kink 1)	.	-.2779 (.0005)		-.2531 (.0005)
Slope 2 (between kink 1 and kink 2)	.	-.7990 (.0005)		-.7442 (.0005)
Slope 3 (above kink 1)	.	-2.0993 (.0064)		-2.5760 (.0082)
Kink 1	.	2.7450 (.0020)		3.1382 (.0025)
Kink 2	.	7.3748 (.0037)		7.5675 (.0029)
R^2	.912	.9980	.894	.9980
N	21,960	21,960	16,311	16,311

Table 5
Regression Results
Individual States

	Linear Regression		Piecewise Linear Kink1 = 2.7450 Kink2 = 7.3748			
State	Slope	R^2	Slope1	Slope2	Slope3	R^2
AL	-0.718	0.884	-0.058	-0.910	-5.007	0.979
AR	-0.710	0.898	-0.263	-0.987	-12.714	0.992
IA	-0.894	0.966	-0.543	-0.987	-4.492	0.995
IL	-0.659	0.955	-0.235	-0.677	-1.420	0.997
KS	-0.647	0.968	-0.485	-0.819	-2.566	0.998
MI	-0.590	0.820	0.028	-0.723	-2.511	0.982
MN	-0.684	0.949	-0.394	-0.847	-1.723	0.995
ND	-0.909	0.995	-0.877	-0.999	*	0.996
NE	-0.718	0.967	-0.568	-0.933	-1.332	0.992
NV	-0.473	0.971	-0.367	-0.504	-1.211	0.994
OK	-0.606	0.907	-0.296	-0.892	-2.930	0.998
OR	-0.522	0.932	-0.281	-0.671	-2.232	0.995
SD	-0.806	0.984	-0.708	-1.000	-1.290	0.995
WA	-0.508	0.907	-0.287	-0.554	-2.679	0.993
WI	-0.709	0.897	-0.117	-0.892	-1.870	0.997

Table 6
Regression Results
Log Cumulative Area on Log Employment Density
(All squares have Employment Density >= 1 per square mile)

	15 States Where Virtually All Land is Part of Public Survey		14 Public Survey States with Land Not Part of the Survey	
	Linear	Piecewise Linear, Two Kinks	Linear	Piecewise Linear, Two Kinks
Constant	13.322 (.002)	13.180 (.0006)	13.077 (.003)	12.9150 (.0008)
Slope (linear)	-.627 (.001)	.	-.581 (.001)	
Slope 1 (below kink 1)	.	-.4928 (.0005)		-.4551 (.0005)
Slope 2 (between kink 1 and kink 2)	.	-.6680 (.0003)		-.6632 (.0007)
Slope 3 (above kink 1)	.	-2.087 (.0046)		-2.077 (.0066)
Kink 1	.	1.9077 (.0053)		2.8560 (.0074)
Kink 2	.	6.8758 (.0025)		6.8051 (.0038)
R^2	.978	.9992	.963	.9982
N	15,527	15,527	12,954	12,954

Table 7
Regression Results
Metropolitan Areas with More than One Million in 2000 Population

	Linear Regression		Piecewise Linear Kink1 = 2.7450 Kink2 = 7.3748			
State	Slope	R^2	Slope 1	Slope 2	Slope 3	R^2
Birmingham-Hoover, AL	-0.725	0.899	0.518	-0.769	-3.088	0.961
Chicago-Naperville-Joliet, IL-IN-WI	-0.465	0.790	0.052	-0.277	-1.521	0.989
Denver-Aurora, CO	-0.389	0.935	-0.294	-0.334	-1.483	0.990
Indianapolis, IN	-0.674	0.938	0.000	-0.567	-2.578	0.989
Kansas City, MO-KS	-0.562	0.924	-0.149	-0.564	-2.577	0.993
Las Vegas-Paradise, NV	-0.339	0.925	-0.274	-0.269	-1.131	0.972
Milwaukee-Waukesha-West Allis, WI	-0.737	0.915	0.000	-0.563	-1.318	0.963
Minneapolis-St. Paul-Bloomington, MN-WI	-0.641	0.939	1.555	-0.534	-1.845	0.991
Oklahoma City, OK	-0.564	0.939	-0.196	-0.578	-2.933	0.996
Portland-Vancouver-Beaverton, OR-WA	-0.443	0.850	-0.060	-0.422	-1.868	0.978
Seattle-Tacoma-Bellevue, WA	-0.417	0.668	0.006	-0.274	-2.360	0.970

Figure 1: The Public Land Survey System Area

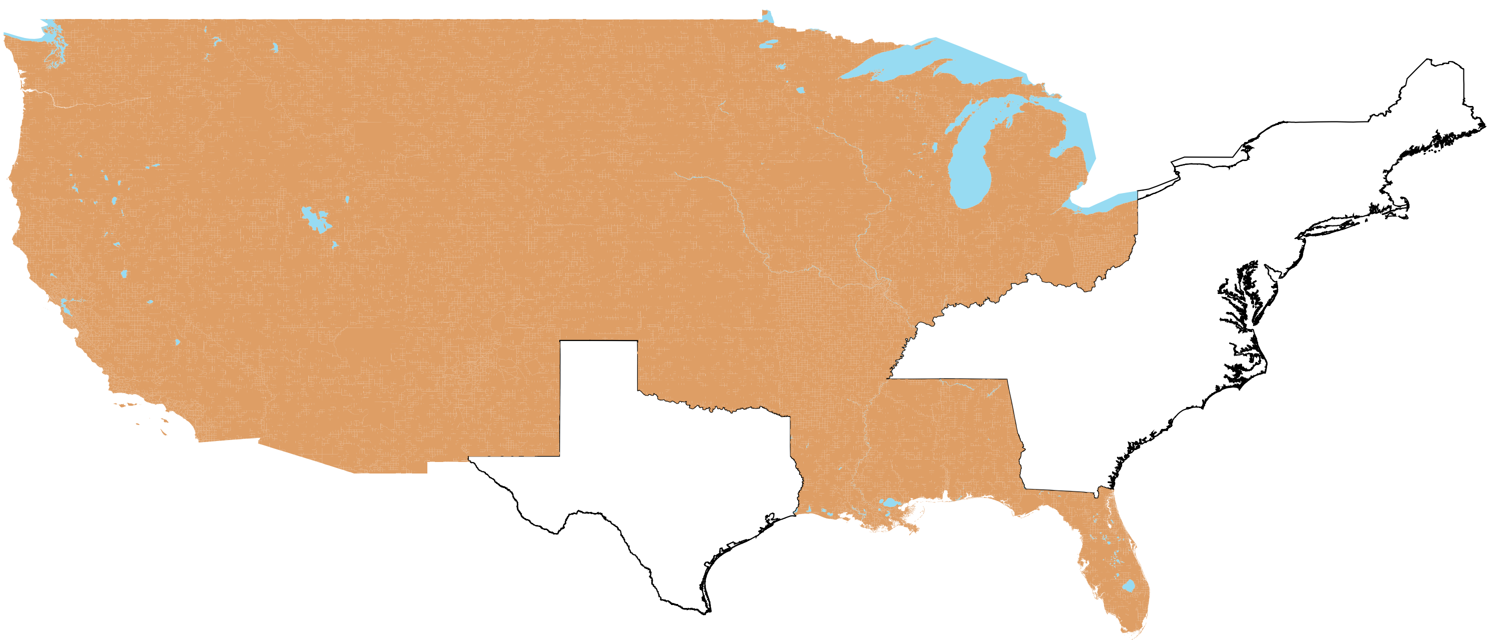


Figure 2: Townships (6*6 mile squares) in ND

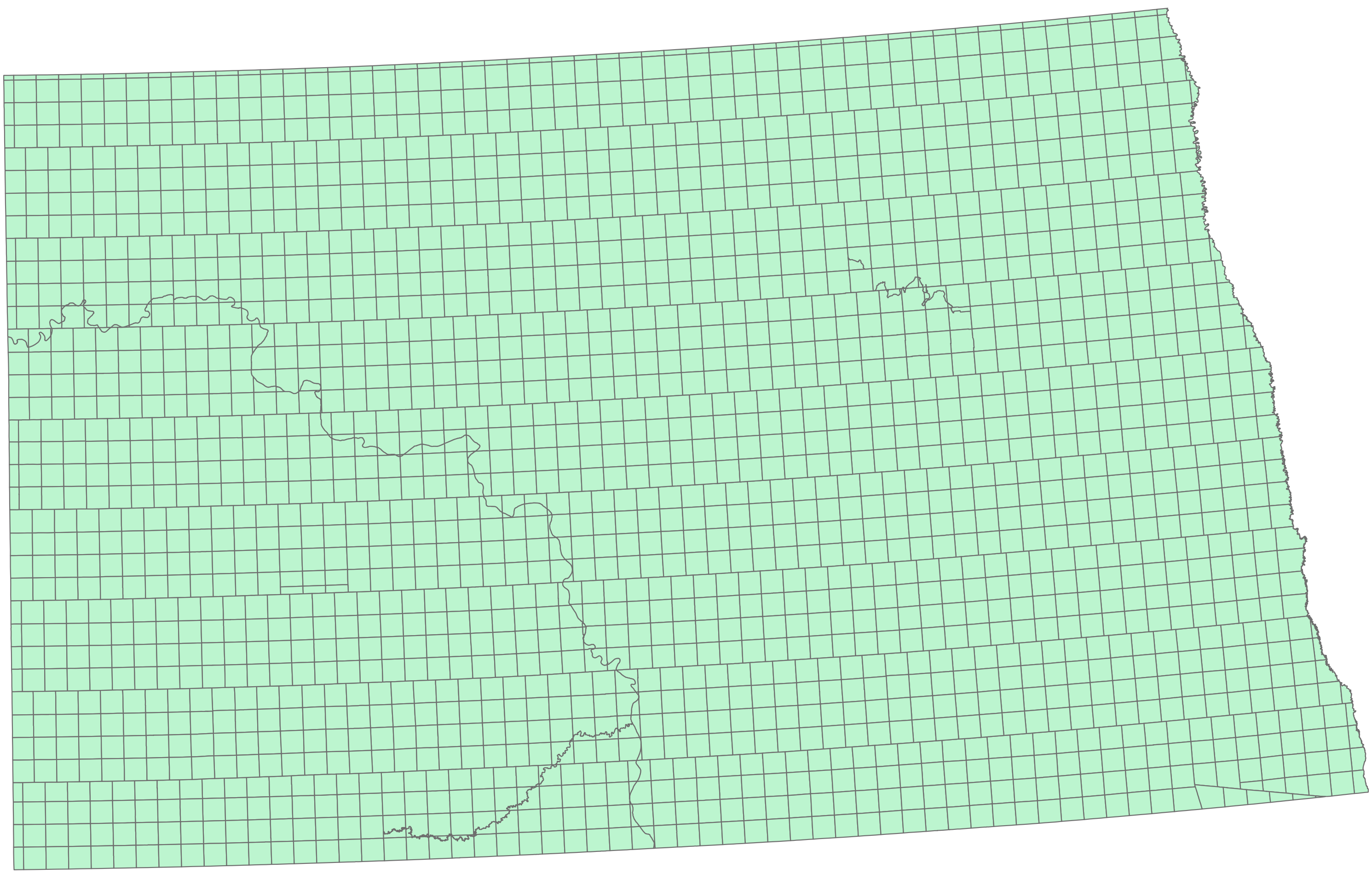


Figure 3: Zipf's Law (Population) for All 15 States

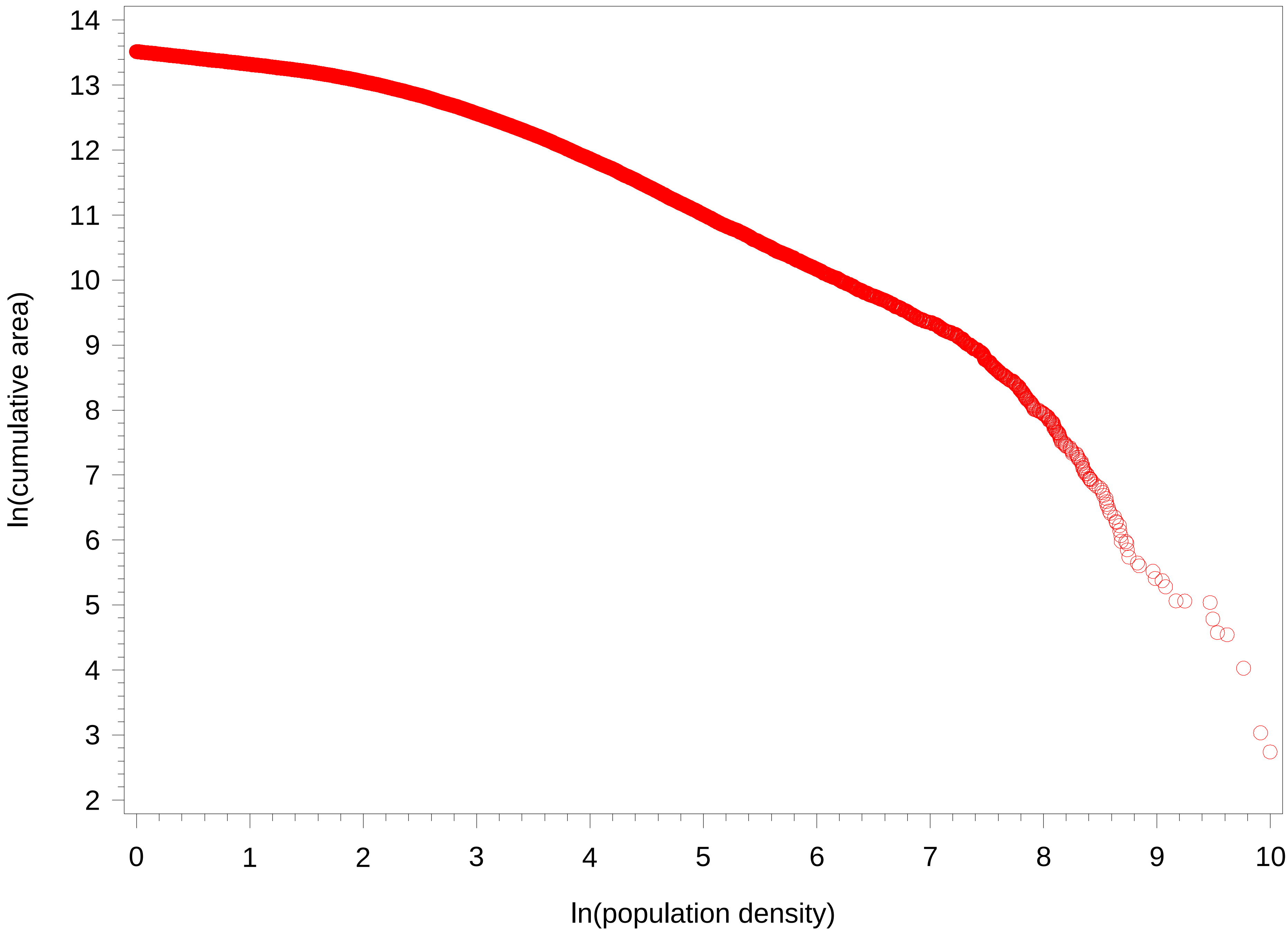


Figure4:Zipf's Law (Employment) for All 15 States

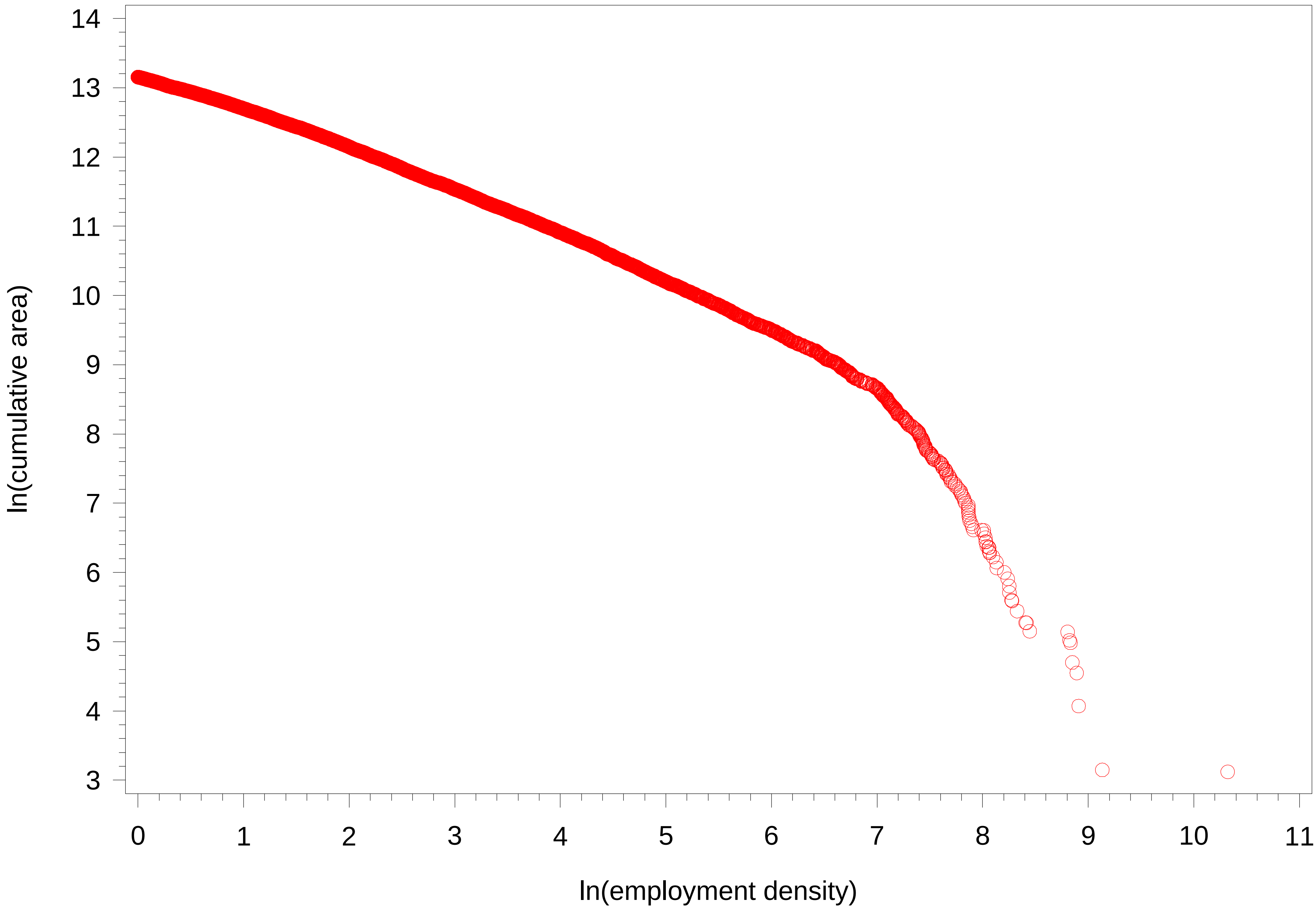
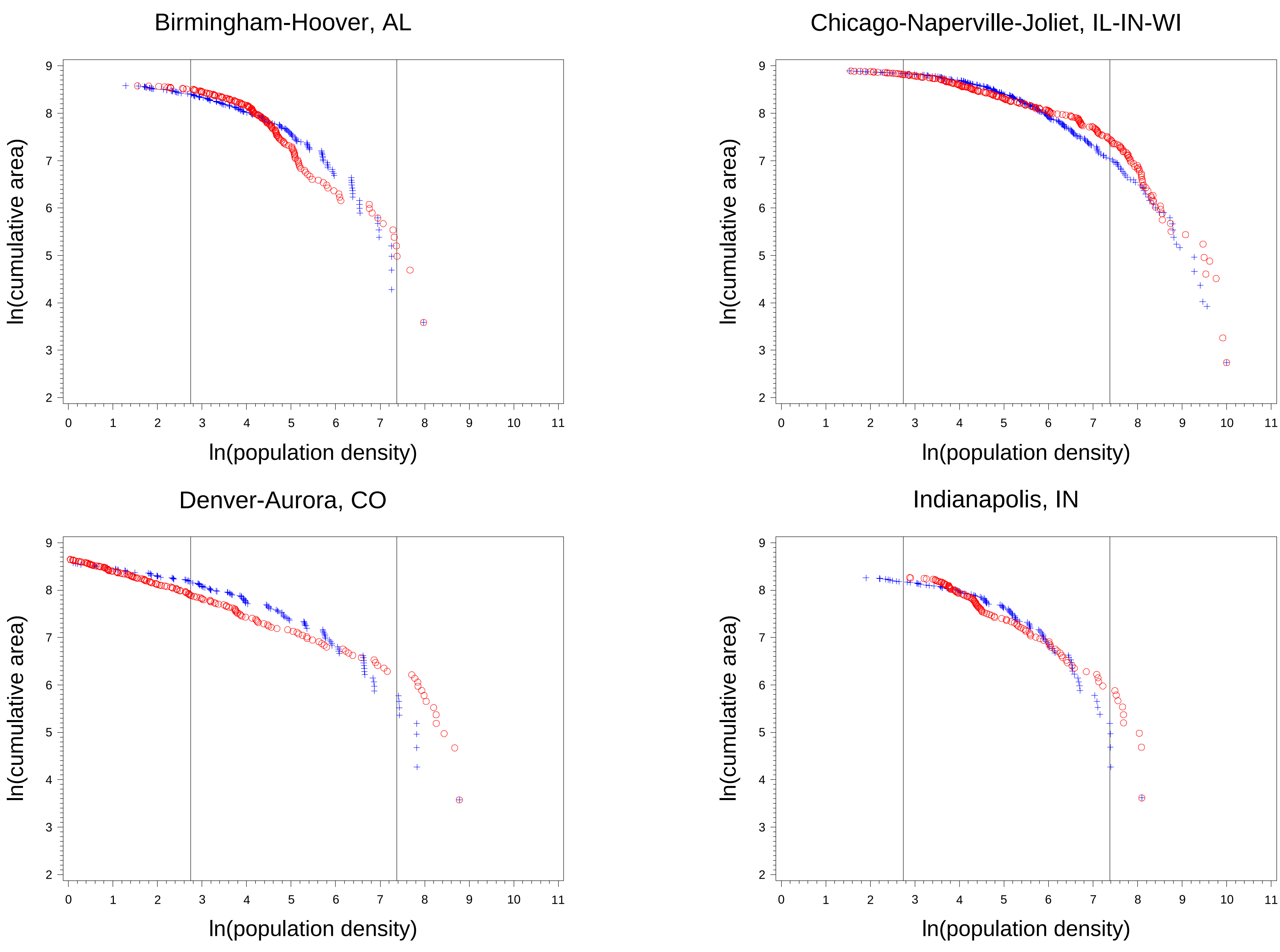
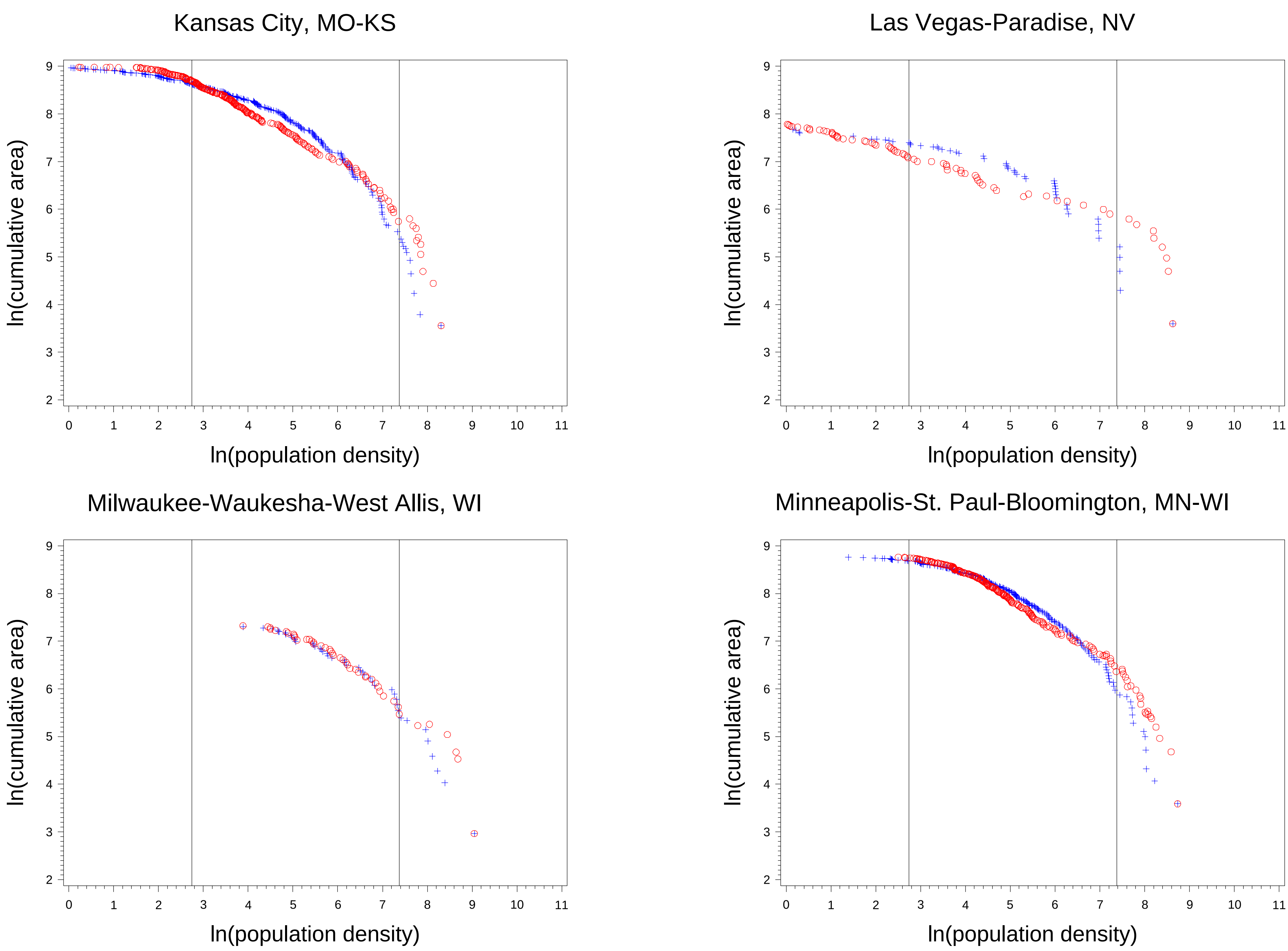


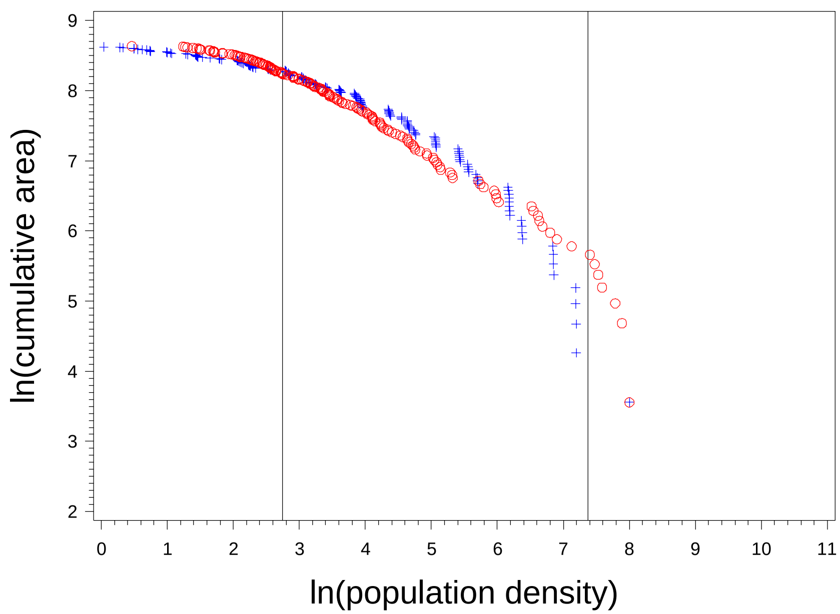
Figure 5: Individual Metropolitan Areas



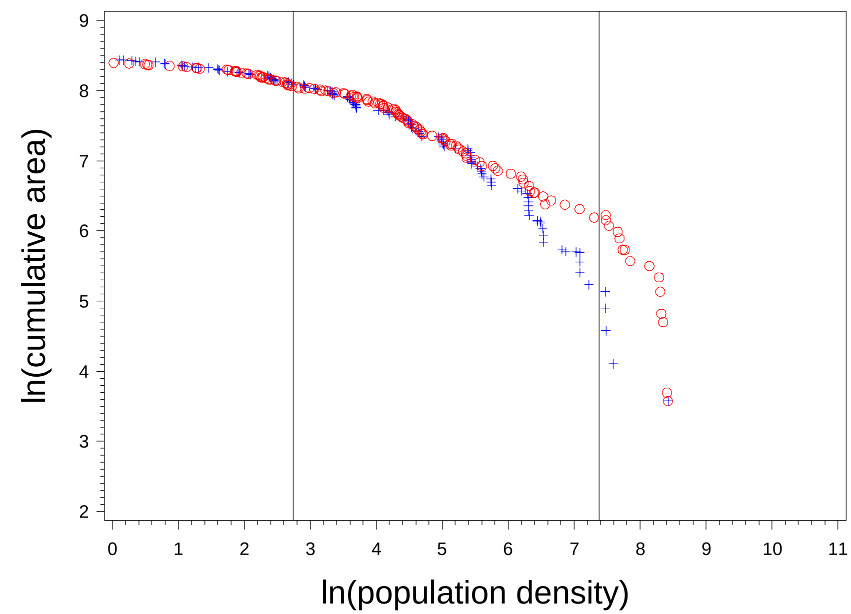


Kansas City, MO-KS
Las Vegas-Paradise, NV
Milwaukee-Waukesha-West Allis, WI
Minneapolis-St. Paul-Bloomington, MN-WI
ln(cumulative area)
ln(population density)

Oklahoma City, OK



Portland-Vancouver-Beaverton, OR-WA



Seattle-Tacoma-Bellevue, WA

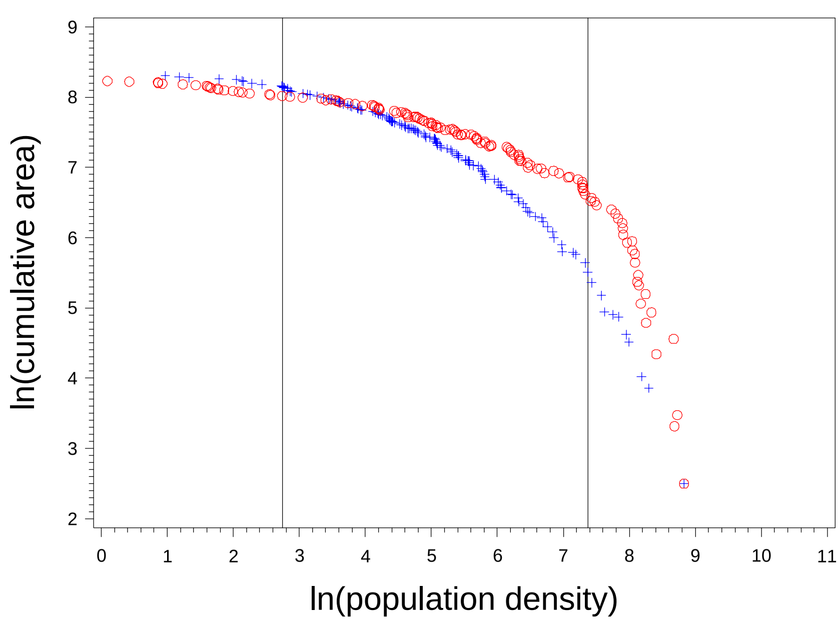


Figure 6: The NAS Rule in North Dakota

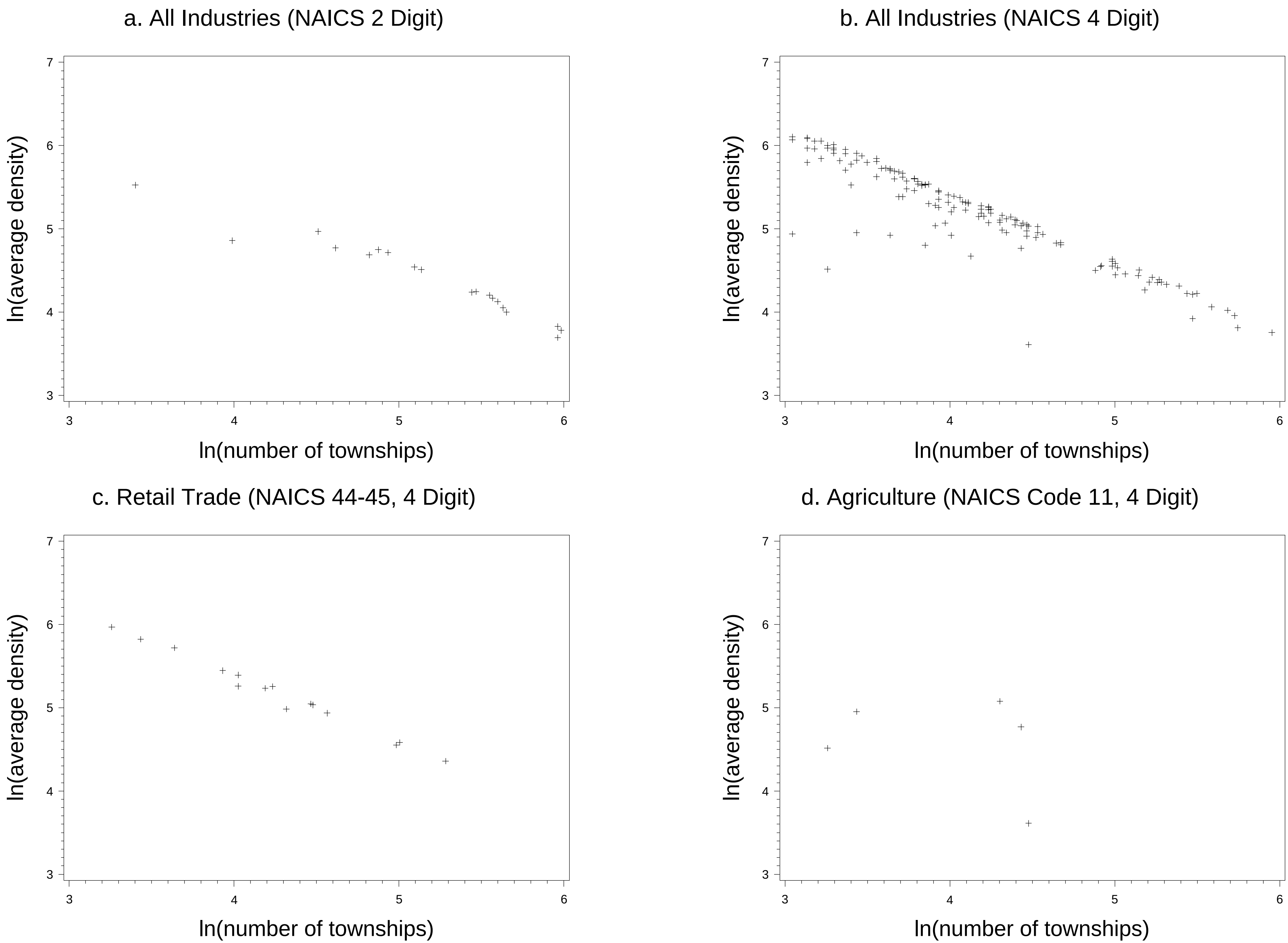


Figure 7: The NAS Rule in Chicago

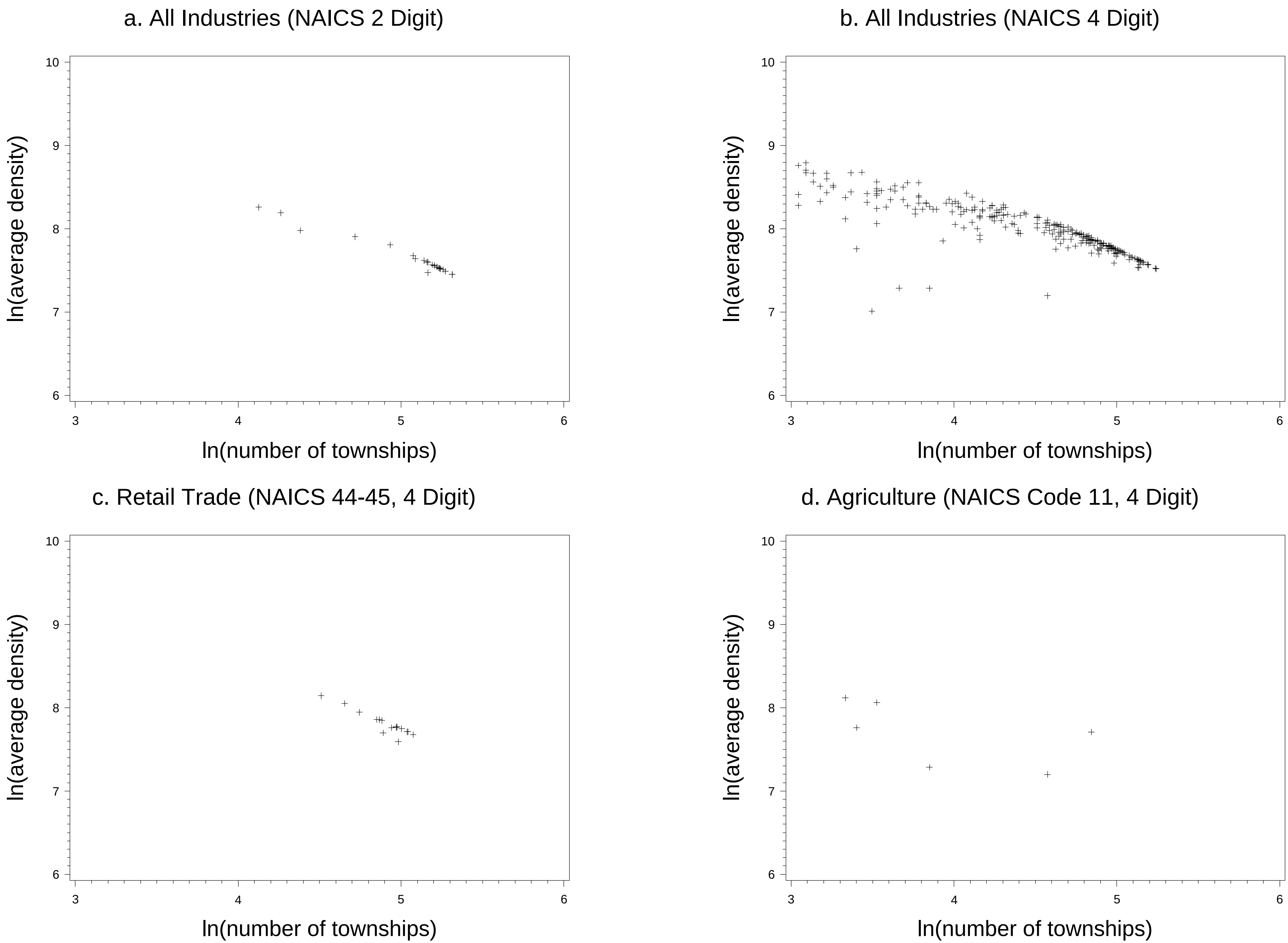


Figure 8: The NAS Rule in Chicago (Dartboard)