The Modifiable Areal Unit Problem
Units make a difference when it comes to geographical analysis: Changing your areal boundaries can dramatically change your results.

---

One More Thing for Planning Researchers to Worry About

The “MODIFIABLE AREAL UNIT” PROBLEM is well known to geographers, but not so much to planning researchers. In An Introduction to Geographical Information Systems, it’s defined as “a problem arising from the imposition of artificial units of spatial reporting on continuous geographical phenomena resulting in the generation of artificial spatial patterns.” Essentially, when you group data for individual points into geographic units of analysis, you lose information and often create bias.

I recently reviewed two papers that illustrate this problem. It’s one researchers and practicing planners alike should be watchful for.

Too small
“Local Variations in the Impacts of Built Environments on Traffic Safety,” published in the Journal of Planning Education and Research, adopts the block group and census tract as the units of analysis for aggregating data on individual crashes in Austin, Texas. Not surprisingly, census block groups and census tracts with more traffic and closer to the central business district experience more crashes.

I recommended acceptance of the paper, as it was well executed, but I also expressed concern due to its analysis of a problem that spills over the census geographies used. I think the census block group, in particular, is too small to study a problem like crash rates. I would imagine a small percentage of the crashes that occur within a census tract, and even more so a block group, involves traffic generated within the block group or tract itself.

Most of the traffic involved in crashes is probably just passing through the block group or tract. The average urbanized census block group is 0.59 square miles, or about 0.77 mile across. The average urbanized census tract is 1.8 square miles, or about 1.34 miles across. At the same time, the average trip length in U.S. metropolitan areas is close to 10 miles. In other words, don’t blame downtown for a crash involving a long-distance suburban commuter who happens to work downtown. A better choice is to use the county as the unit of analysis, since most traffic starts and ends within the boundaries of individual counties.

Turning to the practical import of the JPER paper, traffic volumes are higher where development is concentrated, so the obvious implication of the JPER paper is that development should be dispersed, rather than concentrated. The article concludes: “It is crucial to... design land areas that generate fewer vehicle trips, especially for downtown spaces.” Yet we know from countless studies that dispersed development leads to more vehicle miles traveled, which in turn leads to more crashes. The crashes just occur over many more miles of roadway.

This takes me back to the problem of using census block groups or census tracts as units of analysis, and to the modifiable areal unit problem. These census geographies aren’t large enough to produce meaningful results. In this case, the authors don’t aggregate point data enough.

Too big
I recently reviewed a second paper for Transport Policy: “Means of Transportation to Work and Obesity: A Panel Data Analysis of US States from 2004 to 2013.” It has the opposite problem from the previous paper: The unit of analysis is too large. Using statewide averages, the paper finds that increased automobile usage contributes to rising obesity rates. “...a one percent increase in the share of workers commuting
to work by driving in the previous year will increase the current state obesity rate by 0.15 percent.” Using an area this large creates the real possibility of aggregation bias and ecological fallacy, which occurs when one mistakenly applies the results for a group to its members.

Note that the paper deals with “rates” of obesity, not with individual obesity. If the study had found that, controlling for other variables, individuals driving to work weigh more than those walking to work, that would be notable. Or that those driving farther to work weigh more than those with short drives. Indeed, in studies of individual obesity, starting with Larry Frank’s groundbreaking 2004 study in the American Journal of Preventive Medicine, time spent in cars is implicated in obesity. But a state is too large an areal unit for meaningful causal inference related to individuals.

Just right
That begs the question: What is the right geographic scale for the study of obesity—or anything else? That is, what is the areal unit that might best explain environmental influences?

To answer that question, it is essential to select a geographic scale of analysis that corresponds to your research question. For example, if you’re determining the causes of individual obesity, the right geographic scale is an individual’s activity space.

Since Frank’s study, it has been assumed that the residential neighborhood, defined by a walking distance of about a quarter mile, is the most relevant geography for obesity research. I question this.

According to the 2009 National Household Survey, the average walk trip is 0.69 miles, and the average auto trip is 9.72, both of which take the individual beyond the bounds of a single neighborhood. Most of an adult’s waking hours are spent outside the home: 31 percent of trips don’t start or end at home. It may be that the workplace is as relevant for the study of obesity as is the home location.

Data usually comes in ready-made sets at scales such as traffic analysis zones for travel demand data and Federal Highway Administration urbanized areas for federal highway statistics. If you are unable to choose the scale of your data, test various scales. In travel studies conducted by me and my colleagues, we have been using buffers of various widths, from a quarter-mile to one mile, to describe the relevant environment. Sometimes smaller buffers fit the data best; sometimes larger buffers do.

For planning researchers, it’s vital that you understand and guard against this threat to the validity of your results. For planning practitioners, if research results don’t pass the sniff test, be sure to question the geographic scale of analysis.

—Reid Ewing

Reid Ewing is chair of the Department of City and Metropolitan Planning at the University of Utah, an associate editor of the Journal of the American Planning Association, and an editorial board member of the Journal of Planning Education and Research and Landscape and Urban Planning. More than 50 past Research You Can Use columns are available at mrc.cap.utah.edu/publications/research-you-can-use.

BRT wake-up call
I hope that Jake Blumgart’s article “Nothing Lite About It” in the July issue of Planning serves as a wake-up call to those contemplating bus rapid transit in their communities. BRT has been discussed in the U.S. for close to 20 years, but precious few systems have been implemented that live up to the promise of “light rail at a fraction of the cost.” In fact, a quality BRT system, though less expensive than light rail transit, can still be costly.

That reality, plus the resistance that planners often meet in designing systems that usurp traffic or parking lanes, often results in projects that are “dumbed down.” Instead of landmark waiting areas and signature vehicles, they revert to off-the-shelf bus shelters and conventional buses with fancy paint schemes. These downgrades, along with the lack of exclusive lanes, result in a service that is sometimes called “rapid bus”—somewhat better than normal bus service … but not sufficiently so to attract new riders.

One important objective of implementing BRT, or LRT, for that matter, should be to penetrate the more demanding segments in the local travel market. To do that, BRT systems have to be head and shoulders above the normal expectations for buses. Maybe the article will help bring that principle home.

—William Lieberman, AICP

Transportation Planning Coordinator
CHS Consulting Group
San Francisco, CA