Bad Data:
A Giant Challenge to Management, Budgeting and Policy

Katherine Barrett & Richard Greene

It's no surprise that all local governments confront a number of challenges in developing strong management systems and good policies. Often, this is because of budgetary shortages. But in fact, hidden in the weeds of state and local government efforts is another obstacle: the remarkable amount of local and state data used for decision making that is flawed, out-of-date, biased, inaccurate, and on and on.

This issue has come to the forefront in recent months, as it became increasingly clear that the fight against coronavirus was stymied by a hodge-podge of incomparable, insufficient, and slow-moving data for use in tracking the progress of the disease and determining policies to help advance the fight against the greatest threat to American health safety in a century.

As Janet Hamilton, director of the Science and Policy Council of State and Territorial Epidemiologists, testified to the House Subcommittee on the Departments of Labor, Health and Human Services, Education, and Related Agencies, about a year ago, "The nation's public health data systems are antiquated, rely on obsolete information sharing methods, and are in dire need of security upgrades."

Some of the issues she sited were lack of interoperability, reporting consistency, and data standards, all of which can lead to errors in quality, completeness, timeliness, and communication. One of her major concerns? Many places still rely on paper records and error-prone manual data entry, which are, she says, "still in widespread use (and) have important consequences, most notably delayed detection and response to public health threats of all types: chronic, emerging, and urgent."

But though absence of strong information in the face of COVID-19 may be the most obvious instance of the problem with bad data, it is only one of countless examples from coast to coast.

Though a number of cities are making efforts to control overtime, many are still hobbled by inadequate data systems. "In Atlanta, six departments with the highest overtime use did not have documentation for half of overtime hours recorded, according to a February 2019 city audit, which stated that "Poor record keeping makes overtime use vulnerable to abuse."

In Atlanta (population 500,000), the Department of Watershed Management could support only 6 percent of overtime sampled with documentation. Meanwhile, the Department of Aviation could support only 11 percent of sampled overtime. With those kinds of lapses, it becomes difficult to diminish the unnecessary use of overtime where it exists.

Then there's Shreveport Louisiana (population 190,000). There, according to an April 10, 2019, city audit, when a citizen requests service—most notably for pothole and sidewalk repair—the request is recorded manually on forms that have not had all the necessary information completed, including the total costs for labor, equipment materials, and priority. In fact, the only data that are reliably recorded are the date a repair project begins and the date it's completed. That's a pretty minimalistic approach.

The ultimate problem emanating from these lapses in data-gathering lapse is in prioritizing the projects to be overtaken, according to Leanis Steward, internal auditor

for Shreveport. As he told us, it's important to know what jobs are most important for avoiding physical risks to citizens and tackling those sooner rather than later.

Though absence of strong information in the face of COVID-19 may be the most obvious instance of the problem with bad data, it is only one of countless examples from coast to coast.

Without that kind of data, Shreveport undertakes public works efforts on a first come, first served basis. This means that if a small pothole on a side street is reported at 8:00 on a Monday morning, and the city becomes aware of a pothole that could eat a Volkswagen on Tuesday, the relatively unproblematic hole is filled first, leaving the Volkswagens to watch out for themselves.

Meanwhile, the City of Beverly Hills, California (population 35,000), has been risking property loss or even lives every day because the city doesn't have good data about the functionality of its fire hydrants. "We don't think of them much in municipalities, though they're ubiquitous" says Eduardo Luna, city auditor there for about two years. "But in an emergency, we expect them to be fully functional," he says, referring to December 30, 2019, audit. "What could

have been enhanced is the collection and maintenance of the data on them, and a solid sense of whether the inspections were being done."

In fairness, the city has maintained maps showing the geographic location of all the hydrants. Luna's fear, though, is that a fire truck could easily pull up to a hydrant to put out a significant blaze only to discover that a flawed hydrant may not allow them to hook up their hoses.

Over the last number of years, as we have become increasingly concerned about the shortage of sufficient accurate data in the nation's cities, counties, and states, we've developed a list of guidelines for appropriate data collection and dissemination. We wrote about five of our most important notions in a blog item for our website at greenebarrett.com about eight months ago, and we have revisited and revised that list:

▪ **Absolute figures that come from different size cities or states may not be comparable and can make them unusable for benchmarking.** If, for example, hundreds of people in Los Angeles are affected by a crime, that may not be nearly as alarming a situation as when the same thing happens to dozens of residents in Helena, Montana.

▪ **Over-precision often means lesser precision.** Reports and articles frequently use large numbers, articulating them to the penny. But when project spending is reported as $3,205,432.15, the one thing that's certain is that there was no legitimate way to be sure the expenses were accumulated and recorded credibly. A tight range is often more useful and believable.

Data can't have much meaning if you don't know the **logic behind their creation.**

- **Along the same lines, it's problematic to report an expense as a multiple of the smaller end of a range.** How many times have you seen estimates of expenditures or revenues that fall in ranges like $100 million to $500 million? Clearly, whoever came up with that range didn't have enough information available to come to a more precise number. The problem comes when the city council is trying to make a decision based on that data—individual council members can pick the number they want to vote for, often depending on political preferences.

- **If you're relying on data for which no assumptions are provided, dig deeper.** Data can't have much meaning if you don't know the logic behind their creation. This has long been a problem with pension plans, for example, where there's a range in cities' inflation assumptions, yet their liabilities are treated as though they are strictly comparable. Cost benefit analyses, often by competing interests, arrive at vastly different conclusions, depending on the assumptions they've used.

- **Common sense should prevail.** Some years ago, there was a lot of talk about 1 million children being abducted each year. Yet in New York City (population 8.4 million), news reports were dominated by the story of just one little boy who had gone mission in lower Manhattan. How could it be that if such huge numbers of children were disappearing, one child was getting so very much attention? According to the Denver Post, the "national paranoia" raised by the one million figure wasn't the result of scary people luring children into their cars with candy, but rather children taken in custody battles. And even that often-repeated one million figure was an exaggeration. In 2017, the Justice Department reported that the number of serious parental abductions is closer to 156,000 a year, of which about 30,500 reach the level of a police report. ▣

Katherine Barrett and *Richard Greene* *are principals of Barrett and Greene, Inc. and are co-authors of the recently released* Making Government Work: The Promises and Pitfalls of Performance-Informed Management. *They are: columnists for the Government Finance Officers Association; columnists and senior advisors at Route Fifty; senior advisors at the Government Finance Research Center at the University of Illinois in Chicago; consultants to the National Association of State Personnel Executives; special projects consultants for the Volcker Alliance; columnists for IPMA-HR; and fellows in the National Academy of Public Administration. Greene has been named chair of The Center for Accountability and Performance at the American Society for Public Administration. Their website is greenebarrett.com*