Cites & Insights

Crawford at Large/Online Edition

Libraries • Policy • Technology • Media

Volume 13, Number 5: May 2013 ISSN 1534-0937

Libraries

The Mythical Average Public Library

There is no such thing as the average library.

That's an obvious statement and I'd guess every reader knows it to be true at some level—but it's easy to look at averages *about* libraries and take them to be meaningful, or at least more meaningful than they frequently are.

Caution: This essay is chock-full of numbers and uses a handful of low-grade statistical terms. But I believe this essay will show you some things about the reality of U.S. public libraries (at least in 2010, and as things changed from 2009 to 2010) that you might not have known.

Mythical Numbers

If there was such a thing as an average U.S. public library, it would look like this in 2010 (FY2010). It served 33,017 potential patrons, of whom 18,670 were registered borrowers. It spent $1,201,020 on staff (66% of total expenditures) and $137,170 on the collection (12%), out of total operating spending of $1,174,934. It had 88,395 print books.

Patrons visited the average library 171,438 times. There were 33,668 reference transactions, 268,671 items circulated and PCs with internet access were used 40,119 times. The library had 410 programs with a total attendance of 9,457.

Looked at a little differently, the average public library spent $35.59 per capita (including $4.15 on its collection) and had 2.7 books per patron. The average patron visited 5.2 times, engaged in 1.0 reference transactions, took out 8.1 items, used a PC 1.2 times—and went to 0.3 programs.

But if we assume all library use was by and for registered borrowers (a good assumption for circulation, maybe for PC use, probably not for visits or program attendance), then the average *registered* borrower visited 9.2 times, had 1.8 reference transactions, took out 14.4 items, used a

PC 2.2 times—and went to 0.5 programs. (And there were 4.7 books per registered patrons, along with $62.93 in spending and $7.35 in spending on collections.)

Compared to the previous year, spending was down 1.5%, the bookstock was down 0.5%, visits were down 1% and reference transactions were down 0.2%—but registered borrowers were up 1%, circulation was up 2%, there were 1.4% more programs (but only 0.5% more attendance), and PC use was up 0.6%.

Facts but Not Truth

That barrage of numbers represents facts of a sort. I calculated them directly from the IMLS FY2010 database, stripping out libraries that closed or didn't report key figures, those that weren't active in both FY2009 and FY2010 and territorial libraries (e.g., those in Guam and Puerto Rico).

It's a *larger* set of libraries than IMLS uses in its superb *Public Libraries in the United States Survey*—9,174 libraries and systems rather than the 8,951 included in that report—apparently because there are a couple hundred libraries that don't fully meet the definition of Public Library, or for other reasons. I used the largest plausible set.

Those may be facts, but I'd argue that they're not the *truth* about U.S. libraries, and certainly not relevant for any given library.

Let's look at a few other overall numbers as reported by IMLS, based on a somewhat smaller set of libraries:

- IMLS shows patron visits at 5.28 per (potential) patron, slightly down from 2009 but 33% higher than in 2001.
- It reports spending at $36.18 per capita, a decrease of 2.6% overall (that's adjusted for inflation), staff expenditures as 67% of total spending and collection expenditures as 11.7%. Circulation was 8.3 per capita.

The differences between my figures and IMLS figures are tiny (the percentage change differences are even tinier).

I include the IMLS numbers so you won't be surprised if you read that report—which I strongly recommend, as it is excellent and includes ten-year changes, nearly all of them positive. For the rest of this discussion, I'll use my own numbers—or, for some measures, numbers for a subset that removes difficult cases (e.g., dividing by zero).

I'm going to use "we" in this discussion to mean "most of us," and I could be wrong: Maybe it's just me. I'm certain *some* librarians look at averages with much more jaundiced eyes than I do.

The Average Problem

The problem with averages is that we want them to be more meaningful than they frequently are—and, I think, we're inclined to think of averages in terms of normal distribution. That is, we think of the average figure as being the *most common* figure—and assume that most variations cluster fairly narrowly around that average. I know I have that tendency when it comes to numbers that *could* fit a normal distribution or bell curve, something that could be true for most library measures.

The figure below is one example of a "typical" distribution—and in fact is called a normal distribution. (The figure was created by Mwtoews based on an original graph by Jeremy Kemp; I retrieved it from the Wikimedia Commons.)

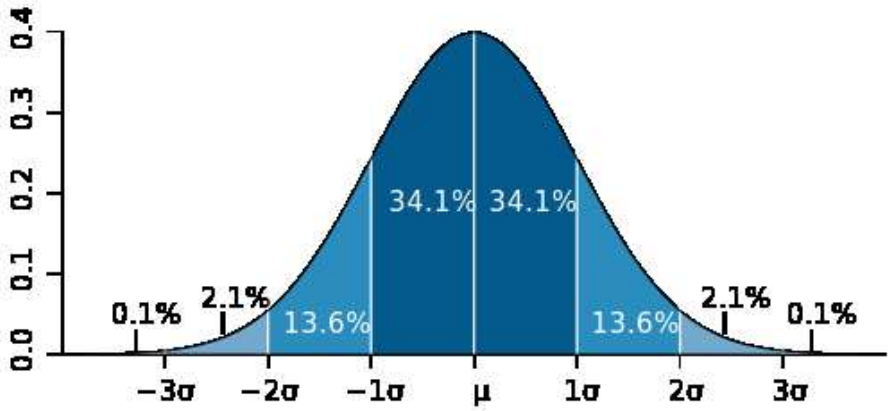

Figure 1. Normal distribution and standard deviation.

The curves in some cases will be broader and shallower or narrower and taller, but that's a detail. The normal distribution—the bell-shaped curve—holds true for *many* phenomena. But not all. Most library statistics don't work this way. Which makes the average considerably less meaningful, as it doesn't say much about the reality of any given library or even *most* libraries: Sometimes most libraries are above average or below average.

One key aspect of normal distribution is that the median (the point at which half the values are lower and half higher) is the same as the mean or average (total values divided by number of values). That's true in Figure 1, where the mode (the single most common value) is also the same. All three *must* be the same for a proper normal distribution.

The further a set of data diverges from this model, the less useful an average will be. That's one reason *Give Us a Dollar and We'll Give You Back Four (2012-13)* almost never offers averages and, where not dividing a universe into levels, offers three figures that *should* be meaningful: The medi-

an, the first quartile and the third quartile. (The first quartile is the point at which 25% of the values are lower and 75% are higher. Need I explain the third quartile? The median is another name for the second quartile.)

I'll talk about "the gap" in some of these discussions—the gap between median and average. The gap is the number of libraries that are either "above average" but below the median or "below average" but above the median. When that's a substantial percentage, it makes the average a misleading figure.

Consider a ludicrous case: How many states have below average population—and how many are in the gap? Dividing total population as estimated for 2012 for the 50 states (excluding DC) by 50 gives 6.266 million as the average state population—but only 17 states have at least 6.266 million people. Thus, 8 states or 16% are in the gap, and 33 (66%) are "below average." Now, if you want *really* ludicrous cases, consider the average number of copies sold for a book—or the average number of potential patrons for an average public library!

Two Statistical Terms

As I go through the rest of this, I'm going to use two other statistical terms—one frequently.

The frequent one (in addition to average or mean, median and quartiles) is *standard deviation*—and Figure 1 also illustrates standard deviation. The darkest section represents values within one standard deviation of the mean (on either side), which should be just over 68% of the population. The slightly lighter area represents values between one and two standard deviations—and, as you can see on the graph, less than 5% of the values should be more extreme than two standard deviations. (Standard deviation is usually represented by σ, the Greek letter sigma. You'll see *s* used as well; that properly refers to *sample* standard deviation, used when looking at a sample rather than the entire population; it's a slightly different formula.)

One nice thing about standard deviation is that Excel or LibreOffice will calculate it for a range of values—and it gives you an immediate sense of whether the data is anywhere *close* to following normal distribution. If the standard deviation is small as compared to the mean, chances are the data follows something like a normal distribution. If it's *enormous*—larger than the mean, for example—that's reason enough to be suspicious of the data's distribution.

In this report, I use the Excel function STDEVPA to calculate the standard deviation because IMLS includes the full population of libraries, not a sample. (For those of a statistical bent, I'm using "n" rather than "n-1" as used by STDEV and STDEVP, both appropriate for samples.)

The other statistic is *correlation*, by which I mean Pearson's correlation (or, to get fancy, Pearson's product-moment coefficient). I use it in

my blog and in *Graphing Public Library Benefits* as a convenient way to show whether two series—e.g., circulation per capita and spending per capita—are "codependent": Whether changes in one are likely to relate directly to changes in the other. Again, it's convenient because Excel will calculate it for a range (the functions CORREL and PEARSON both return Pearson's coefficient), and because it's a reasonable measure of "fit." Based on what I've seen elsewhere, I interpret a correlation of 0.5 or more (or -0.5 or less) as strong correlation and 0.3 to 0.499 (or -0.499 to -0.3) as moderate correlation.

You'll find *lots* more on these terms at *Wikipedia* and the articles appear to be sound, but you'll also encounter lots of equations in those articles, none of which I'm going to subject you to.

After seeing how much library averages and standard deviations were affected by outlying cases, I've gone back and offered two other numbers to show those effects (for most, but not all, measures): the trimmed library average and standard deviation. To arrive at the trimmed figures, I remove 0.5% of the libraries (total) from the top and bottom of the range of numbers; that's normally 24 at each end, except some measures in which some libraries are inherently excluded.

Library spending: I should clarify that, in all cases discussing library spending (or spending per borrower), I'm using library operational expenses, not total library expenses including capital spending and other special cases.

Numbers and Ranges

For most measures or changes, I'll offer some or all of the following, including comments and graphs as appropriate:

- Overall average: What's *usually* called an average, I believe.
- Library average: For any *derived* measure, such as circulation per capita, visits per registered patron, or changes from 2009 to 2010, the average of all the derived measures. The two numbers can be sharply different. Also *trimmed* library average, the average of all but the most extreme half-percent of libraries.
- Median: The point at which half the libraries are lower and half are higher.
- First quartile: The point at which 25% of the libraries are lower and 75% are higher.
- Third quartile: The point at which 75% of the libraries are lower and 25% are higher.
- Standard deviation: The calculated standard deviation, direct from Excel. I may not give this for change percentages, as it seems not to be very meaningful there. Also *trimmed* standard deviation, ignoring the most extreme half-percent of libraries.

- Gap: The number and percentage of libraries between average (either average) and median or between the two averages for derived measures.
- Central cluster and population percentages: For calculated measures, the number and percentage of libraries within 10% in either direction of the median. (For changes, I'll use the actual percentage—that is, if the median change in circulation is +6%, I'll use 96% to 116% as the central cluster.) If the median and average are identical or close to one another, the bigger the central cluster, the more useful the average is as a single figure. (That is: If, say, 90% of libraries circulate 8.4 to 10.2 items per capita and the average is 9.3, then it's fair to say "libraries typically circulate around 8.4 items per capita." If only 30% of libraries fall into the central cluster, I'd argue that it's misleading—even false—to offer such a statement.) In some cases, I'll also offer one or two percentages—namely, the percentage of all registered borrowers represented by libraries in the central cluster and the percentage of all library circulation represented by libraries in the central cluster.
- Extremes and population percentages: Largely anecdotal indications of how many (and what percentage of) libraries fall into high and low extremes for derived measures, and in some cases one or both of the percentages mentioned above.

Most graphs will exclude extremes and I'll always note how many libraries are excluded. Except for one example, there are no graphs for direct measures (they're all power-law curves rather than normal distributions—there are always a few very large cases and many very small cases because of the nature of the American public library "system"), and I don't include central clusters or extremes for direct measures.

Now let's look at the measures already mentioned (and a few others) and see how reality differs from that simple average.

Staff Expense Percentage

Start with staff expenses as percentage of total library spending, because that seems as though it's likely to follow something resembling normal distribution—it *could* be a "well-behaved" measure.

It's also an incomplete measure: only 5,887 libraries and systems (where the same library reported figures in both FY2009 and FY2010) reported staff expenditures for FY2010. That's less than two-thirds of all U.S. libraries. While there are some libraries with no staff expenditures (volunteer-run libraries, which technically don't qualify as public libraries by IMLS standards), most of them—I'm guessing—simply failed to report. (That includes one *very* large California public library system.)

It's probably worth noting that the 5,887 libraries include the vast majority of all libraries' legal service areas and total spending: 294 million of 308 million potential patrons (95%) and $10.6 of $10.8 billion *reported* spending (98%).

Reported staff expenses (salaries, fringe benefits, etc.) were $7.1 billion or 67.2% of total reported spending. So the overall average is 67.2%. The library average is lower: 66.4%. The median is almost exactly the same as the overall average: 67.1%—close enough that this *could* be a normal distribution The first quartile is 61.0%; the third is 72.9%. Standard deviation is 9.6%—which *should* mean that 68% of the libraries spend between 57.5% and 76.7% of their total spending on staff. (The trimmed library average is 66.5%; the trimmed standard deviation is 9.3%.)

Does that work out in reality?

- 916 (15.5%) spend 57.4% or less.
- 747 (12.7%) spend 76.8% or more.
- That leaves 4,224 (71.8%) within the range, so it's a slightly better fit than you might expect.

This *is* a well-behaved measure. Except for relatively few outliers, most public libraries spend somewhere around two-thirds of their budgets on staff.

The gap between library average and median includes 166 libraries (2.8%), a very small gap. The gap between median and overall average is either zero (excluding endpoints) or, by adding another decimal point, 22 libraries.

The central cluster, libraries within 10% of the median in either direction, includes 3,256 libraries (55.3% of libraries reporting staff expenses); in all, those libraries serve 60.1% of the population served by libraries reporting expenditures.

If you're interested in the extremes, 21 libraries spend less than 30%, 49 more less than 40%, and 220 more less than 50%. On the high end, 23 libraries spend at least 90% of their total spending on staff (although that includes one library that reported $0 spending on collections); another 65 spend 85%-89.5%; and other 246 spend 80%-84.5%. The 290 libraries (4.9%) at the extreme low end serve 3.1% of the population; the 334 (5.7%) at the high end serve 5.6%.

Figure 2 shows staff expenses as percentage of all spending, rounded to the nearest whole percentage, in this case *without* removing extremes. (Thanks to Excel's intractability where pivot-table-based graphs are concerned, the percentages appear as decimal numbers.) Although there's a double spike at the top, this is reasonably close to normal distribution.

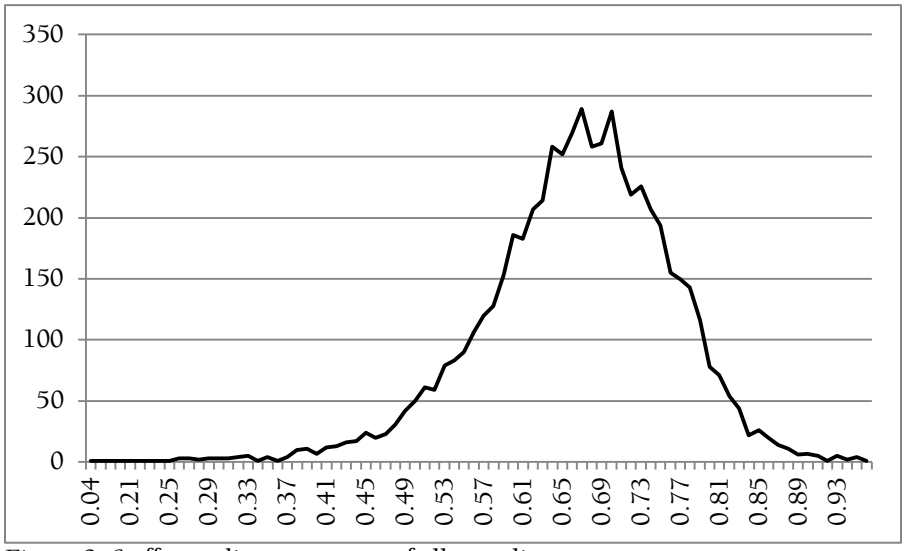

Figure 2. Staff spending as percent of all spending

Collection Spending

How about collection spending as a percentage of total spending for these 5,887 libraries?

I'm going to cheat a little here—removing 14 libraries that reported less than 1.0% spending on collections and three that reported spending at least half of the budget on collections (in one case 85%). That leaves 5,870 libraries.

The overall average (including outliers) is 12.0%. The library average (*excluding* outliers) is 12.2%. The median is 11.8%—not identical, but close. The first quartile is 9.2%; the third, 14.7%

Standard deviation is 4.7%, so two-thirds of the libraries should fall between 7.5% and 16.9%. In fact, 73.3% do fall within that range. (Trimming would not change the library average at all and would reduce the standard deviation to 4.5%.)

The central cluster includes 1,346 libraries (22.9%). Those libraries serve 28.0% of the total population for these 5,870 libraries.

Although Figure 3 technically isn't a true normal distribution neither is Figure 2, since the median and average aren't *quite* identical, it's about as close as you'll see here—and it's certainly close to a classic bell-shaped curve, albeit a bit craggy.

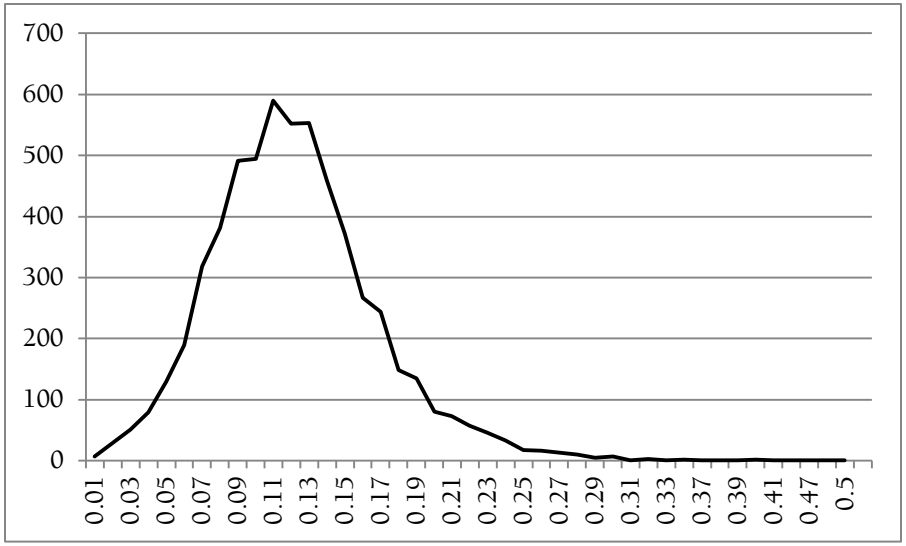

Figure 3. Collection spending as percent of total

Library Service Area

Just for fun, let's go to the other extreme—and it's a case where IMLS does *not* suggest an average library, any more than it does for total budget (probably an even greater extreme). That's LSA, the library service area population.

What's interesting here is just how far the mean or average is from the median. Of 9,211 libraries, the average LSA is 33,066 potential patrons—but the median is 7,061! The first quartile is 2,187. The third quartile is 21,808. When the average is well into the top quartile, you know it's an odd distribution.

The standard deviation is 130,163—a silly number for a measure that bears no relation to normal distribution. The trimmed standard deviation is a little better (78,307), but not much.

The gap is enormous: 2,981 libraries—not quite one-third of them—fall between the median and the average. Since half *necessarily* fall below the median, you can say that more than 80% of America's libraries are smaller than average—because the average is meaningless.

The service population of America's public libraries and library systems follows a power-law curve: A huge number at one extreme and very few anywhere else. Rounding to the nearest thousand (which yields more than 400 libraries that round to zero), it's possible to generate Figure 4—but it's important to note that the vertical scale is logarithmic.

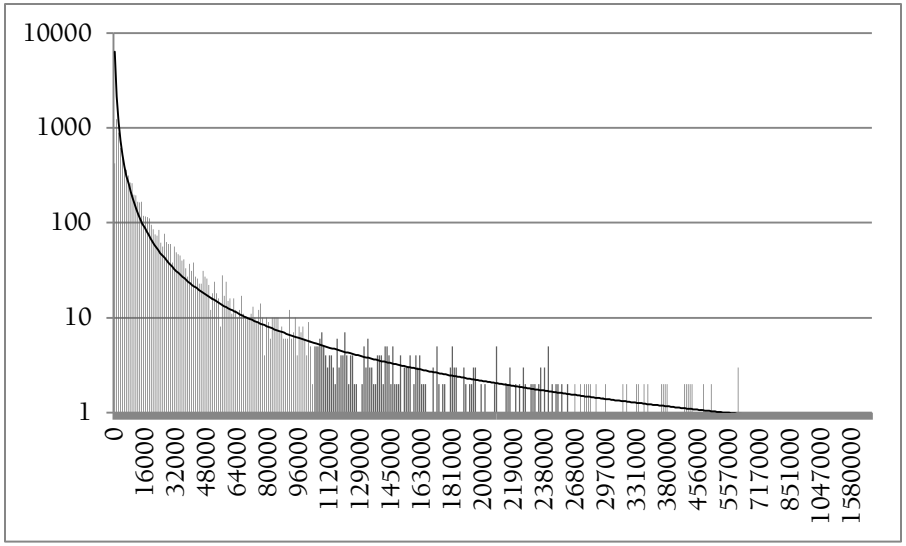

Figure 4. Libraries by LSA, rounded to nearest thousand

I regard Figure 4 as meaningless except to show an extreme case—and for other measures where the distribution is this extreme, including most direct measures, I won't bother with a figure.

Other Direct Measures

Let's run through the others that I mentioned at the start of this article, understanding that they're all wildly diverse, with *thousands* of libraries near the bottom and relatively few near the top. Because I'm lazy, and for better congruence with the rest of this article, I'll do them in the order in which they appear in the IMLS database, not necessarily the order in the introductory paragraphs of this discussion.

Feel free to skip this section—going down to "Derived Measures" roughly three pages away—although you may find some of the numbers mildly interesting—e.g., the (apparent) fact that one-quarter of America's libraries and systems have fewer than 14,641 books each, while only one-quarter have at least 69,559 books.

Spending on Collections

While the *percentage* of spending for collections (both physical and electronic) is a case where the average is meaningful (and for the full set of libraries, it's very close to the subset discussed), the actual *dollars* are wildly skewed.

The average is $137,170. The median is $23,342. The first quartile is $7,599 (one-quarter of libraries spent less than $7,600 on collections) and the third quartile is $81,707—still well below average. The standard

deviation is $627,318. (The trimmed average is $114,159 and the trimmed standard deviation is $347,673.)

A full 3,062 libraries—almost precisely one-third—fall in the gap, spending between $23,343 and $137,169 on collections. More than 83% of public libraries are below average for this figure.

Looking at extremes, one library spent nearly $30 million on collections in 2010. Two dozen more spent $5 million or more, while another 176 spent more than $1 million and less than $5 million. At the other extreme, 38 libraries didn't admit to spending *anything* on collections in 2010, 23 more spent less than $100 and 93 more between $100 and $499.

For what it's worth, the 10% region around the median spending on collection includes 421 libraries (4.5%), but those libraries account for 1.1% of all registered borrowers and 0.9% of all circulation.

Total Spending

As already noted, average spending for FY2010 was $1,174,934. The median was $190,279. The first quartile is $57,695; the third, $709,099. The standard deviation is another ludicrous figure: $5,258,123. (The trimmed average is $980,298; the trimmed standard deviation, $2,904,033.)

How big is the gap between average and median? Big enough for 3,032 libraries—about one-third. More than 86% of libraries spend less than average.

One library spent $231 million in FY2010 and another 15 spent more than $50 million. Another 151 spent $10 million to $49.999 million. Since libraries reporting *no* expenditures were removed, that leaves four libraries spending less than $66, another 11 spending $200 to $994, and another 97 spending $1,017 to $4,988.

Looking at the cluster within 10% of the median in either direction, it includes 371 libraries, 4.0% of the total. Those libraries support 1.0% of all borrowers and handle 0.8% of all circulation.

Print Books

The mythical average library had 88,395 books in 2010—but the median was 29,198. First quartile is 14,641; third is 69,559. The standard deviation is 357,657. (The trimmed average is 75,719; the trimmed standard deviation, 169,661.)

The gap includes 2,776 libraries (30.3%).

One vast library (system) has 20,919,629 books and nine others have at least four million; one library reports having only 48 books, while nine others have fewer than 1,000.

The range within 10% either side of the median includes 595 libraries, 6.5% of the total. These relatively small libraries support 1.6% of all registered borrowers and handle 1.2% of circulation.

Patron Visits

The average for patron visits is 171,438; the median is 34,277. First quartile is 9,869; third quartile is 124,606. Standard deviation is 651,447. (Trimmed average: 147,993; trimmed standard deviation: 418,555.)

The gap includes 2,842 libraries (31.0%).

In this case, the "vast library system" isn't way ahead of others: It had 17.8 million visits, while two others had more than 16 million. Four more had more than ten million visits and 33 more had more than five million. At the low end, five libraries reported fewer than 100 visits each during FY2010 (the lowest was 18) and another 72 reported fewer than 500.

The range within 10% either side of the median includes 364 libraries, 4.0% of the total. Again, these are small libraries: they serve 1.0% of borrowers with 0.7% of circulation.

Reference Transactions

This measure has the most extreme variation between average and median. The average is 33,668; the median, 3,191—less than one-tenth as much. The first quartile is 700 (*really*—one-quarter of the public libraries reported fewer than 700 reference transactions for the year); the third is 13,402. Standard deviation is 214,830. The trimmed average is 25,898 and the trimmed standard deviation is 95,554.

The gap includes 3,351 libraries (36.5%), so more than six out of every seven (86.5%) libraries are below average for this measure.

One library system had more than 10.4 *million* reference transactions, three others had more than five million and 47 more had more than a million. At the low end, 84 libraries reported zero reference transactions and 72 more reported fewer than ten. (The system with the most reference transactions had more than the lowest 5,852 libraries combined!)

The range within 10% either side of the median includes 358 libraries (3.9%) that serve 1.2% of borrowers with 0.9% of circulation.

Registered Borrowers

The average figure for registered borrowers is 18,670; the median is 3,963. The first quartile is 1,200; the third is 12,136. Standard deviation is 76,358. (Trimmed average 16,038; trimmed standard deviation 45,039.)

The gap between median and average includes 2,988 libraries (32.5%).

The largest number of registered borrowers—this time for a different very large library system on the other coast—is just over 3.3 million. Second, at 2.8 million, is the system that leads on several other measures. Four other library systems have more than one million registered borrow-

ers. At the other extreme, three libraries have fewer than 20 registered borrowers and another 101 have fewer than 100.

The range within 10% either side of the median includes 425 libraries (4.6%) that serve 1.0% of borrowers with 1.0% of circulation.

Circulation

Average 268,671; median 47,121. The first quartile is 13,376; third is 158,028. The standard deviation is the second most ludicrous number in this series: 1,106,505. (Trimmed average: 227,383; trimmed standard deviation: 716,915.)

The gap includes 3,057 libraries (almost precisely one-third of the total), so five out of six libraries have below-average circulation.

Five libraries report more than twenty million circulation each (the high is 26.7 million) and another 25 are over ten million. On the low side, five libraries reported less than 100 circulation, with another 55 below 500.

The range within 10% on either side of the median includes 385 libraries (4.2%) serving 1.1% of borrowers with 0.7% of circulation.

Programs

The average number of programs is 410; the median, 131. The first quartile is 39; the third is 354. The standard deviation is 1,450. (The trimmed average is 355; the trimmed standard deviation, 808.)

The gap includes 2,628 libraries (28.6%).

When IMLS says that the average library has more than one program a day, it's worth pointing out that more than three-quarters of US libraries have *less* than one program a day, with one-quarter having just over three per month or fewer.

The most active library (which is *not* one of the very largest library systems) reported more than 53,000 programs; 32 others had at least 10,000 programs each. The low side includes 148 libraries with no programs at all and another 445 with fewer than five programs each.

The range within 10% on either side of the median includes 449 libraries (4.9%) serving 1.7% of borrowers with 1.2% of circulation.

Program Attendance

Average total program attendance was 9,457, while the median was 2,373. The first quartile is 644; the third, 7,465. Standard deviation is 30,739. (Trimmed average is 8,349; trimmed standard deviation is 20,297.) The gap includes 2,660 libraries (29.0%).

The highest reported attendance is just over 853 thousand, with another seven libraries showing more than half a million total attendance each. Naturally, there are 148 libraries with no attendees (since there were no programs), and another 22 show fewer than 10 total attendees.

The range within 10% on either side of the median includes 372 libraries (4.1%) serving 1.2% of borrowers with 0.9% of circulation.

Personal Computer Sessions

Here's the last direct measure before we go on to derived measures where there might plausibly be something like normal distribution. The average is 40,119; the median is 7,733. The first quartile is 2,411; third is 24,013. Standard deviation is 165,468. (Trimmed average is 33,954; trimmed standard deviation, 101,0166.) The gap includes 3,053 libraries (33.3%, but a bit less than one-third).

Two very large library systems reported more than 4.4 million sessions and another dozen reported more than two million. At the other extreme, 85 libraries reported no PC sessions (some don't *have* any public access computers, others don't count sessions) and eleven more reported fewer than ten sessions.

The range within 10% on either side of the median includes 448 libraries (4.9%) serving 1.2% of borrowers with 1.2% of circulation.

Derived Measures

All the numbers for the rest of this discussion are derived numbers that *should* be independent of the size of the library—either percentage changes between 2009 and 2010, registered borrowers as a percentage of legal service area or measures on a per-borrower basis (*not* a per-capita basis, covered in my books). You might expect to see smaller gaps between overall average, library average and median and maybe smaller standard deviations.

Let's use state populations again to illustrate the difference between overall average and library average, but this time I'm going to make up the numbers. Let's suppose that California has 33 million residents, 50% of whom are Hispanic. Let's further suppose that each of the nine other westernmost states has three million residents, 10% of whom are Hispanic. What's the average percentage of Hispanics in the ten western states?

The total population is 60 million. The total of Hispanics is 19.2 million. So the *overall* average is 32%: Just under a third of the western states' population is Hispanic. But only one state has more than a third Hispanics—indeed, only one has more than 10%. The average of Hispanics by state is 14%: The average western state is 14% Hispanic.

Thus the two averages. If you're looking at the nation as a whole, the overall average is the meaningful figure. If you're looking at libraries within the nation, the average library figure *may* be more meaningful.

The median in this imaginary scenario is 10%. So are the first and third quartiles. You could say "almost all of the west is one-tenth Hispanic," but that would also be misleading.

I'm providing trimmed *library* averages (*not* overall averages) and trimmed standard deviations.

Registration Percentage

Across the nation's libraries, registered borrowers represent 56.6% of total population. But the average library registers 68.6% of its potential patrons. In this case, the median rounds to the same as the national average—56.6%. The first quartile is 40.4%; the third, 77.8%.

The standard deviation is 77.5%. The trimmed library average is 66.2%, not much lower than the untrimmed average—and while the trimmed standard deviation is smaller at 45.7%, it's still pretty large. The gap between median and library average includes 1,480 libraries (16%).

The central cluster includes 1,575 libraries (17.2%) that serve 17.8% of borrowers with 18.9% of circulation.

Two other numbers I find interesting:

- 5,524 libraries (60.2%) have at least half as many registered borrowers as potential patrons; those libraries serve 72.2% of all borrowers with 69.4% of all circulation.
- 3,307 libraries (36.0%) have at least two-thirds as many borrowers as patrons; those libraries serve 43.7% of patrons with 39.2% of circulation.

Some libraries (mostly relatively small) register several *times* as many borrowers as their legal service area population, in one case more than forty times as many. Nine more libraries have at least ten times as many registered borrowers as their LSA, and another 31 at least four times as many. In all, 1,209 libraries (13.2%) have more registered borrowers than their legal service area population. Those libraries are, overall, somewhat smaller libraries: They serve 10.4% of all registered borrowers with 7.9% of circulation.

At the other extreme, eight libraries register less than 1% of their potential patrons (the worst case is 0.04%) and another 56 register less than 5%. Combined, those 64 libraries (0.7%) serve 0.03% of borrowers with 0.06% of circulation.

For this graph, rounded to the nearest full percentage point, I'm omitting all values in excess of 100%, but also those below 10%, leaving a total of 7,840 libraries (85% of the total).

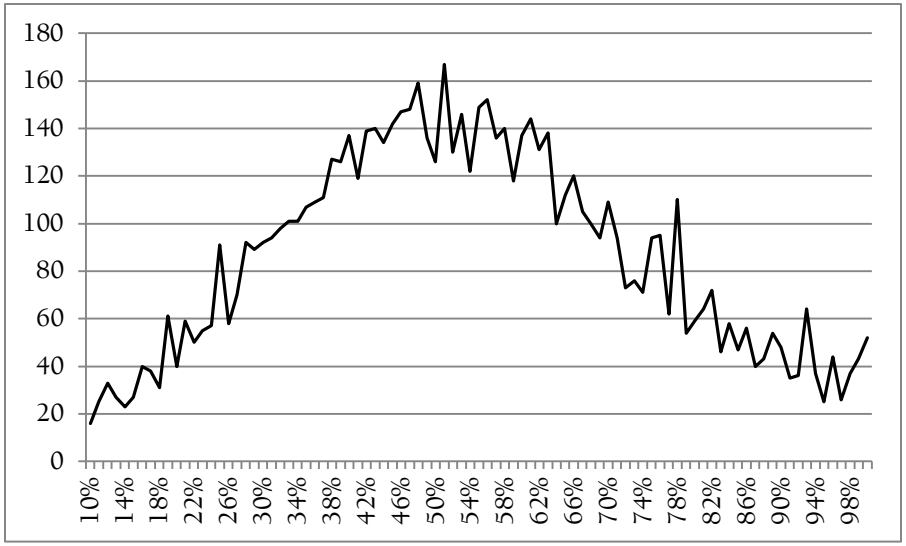

Figure 5. Percentage of registered borrowers (for 85% of libraries)

Figure 5 shows the results—and the shape is close to a normal distribution, a fairly broad one. But note that there are a few libraries not shown off to the left—and *many* (1,209) off to the right.

Changes from 2009 to 2010: Overall

We'll get back to changes in direct measures after this section. Meanwhile, the percentage of registered patrons grew from 2009 to 2010 by 0.2%: Not a huge change, but still a positive one.

That's the overall average. The library average is significantly higher at 6.1% (that is, for the "average library," the percentage of registered borrowers was 106.1% in 2010 of what it was in 2009—if it was 55% in 2009, it was 58.7% in 2010). The median is in the middle—103.1%, that is, an increase of 3.1%.

The first quartile is 96.6%, a drop of 3.4% in *percentage* of patrons registered as borrowers. The third quartile is 108.9%, an increase of 8.9%. Neither of these represents a huge change.

The standard deviation is 109.4%. That seems fairly meaningless. But look what happens when we drop the bottom 24 and top 24 values: the new trimmed library average is 103.6% (up 3.6%), *very* close to the median. The trimmed standard deviation is 27.0%, which is plausible if still a bit high.

The gap between overall average and median includes 1,121 libraries (12.2%); the gap between median and library average, 1,514 libraries (16.5%).

A few hundred libraries had significant LSA changes to deal with—*literally* "a few hundred," with 240 libraries showing at least a 10% drop

in LSA and 263 showing at least a 10% increase. Other dramatic changes can be the result of periodic clearing of outdated registrations or significant drives to increase registration.

The high extreme is indeed extreme, with 26 libraries more than quintupling the *percentage* of potential patrons registered as borrowers. So is the low extreme, with 36 libraries falling to less than one-fifth the *percentage* of potential patrons registered.

If we define the central cluster as 10% on either side of the whole median percentage, that is, 103.1, 6,084 libraries (66.3%) are in the central cluster, supporting 74.8% of borrowers with 73.7% of circulation: Not only most libraries but most larger libraries.

While 3,356 (36.6%) of the libraries saw some reduction in the percentage of patrons registered to borrow, 63.4% either held firm or increased that percentage. Those libraries support 66.0% of borrowers and do 62.3% of circulation.

Figure 6 omits 181 libraries where registered percentage fell by more than 50% (2.0%, supporting 0.4% of borrowers and 0.5% of circulation) and 266 where it grew by more than 50% (2.9%, supporting 1.2% of borrowers and 0.8% of circulation) and rounds the differential percentage (the 2010 percentage divided by the 2009 percentage) to the nearest whole percent.

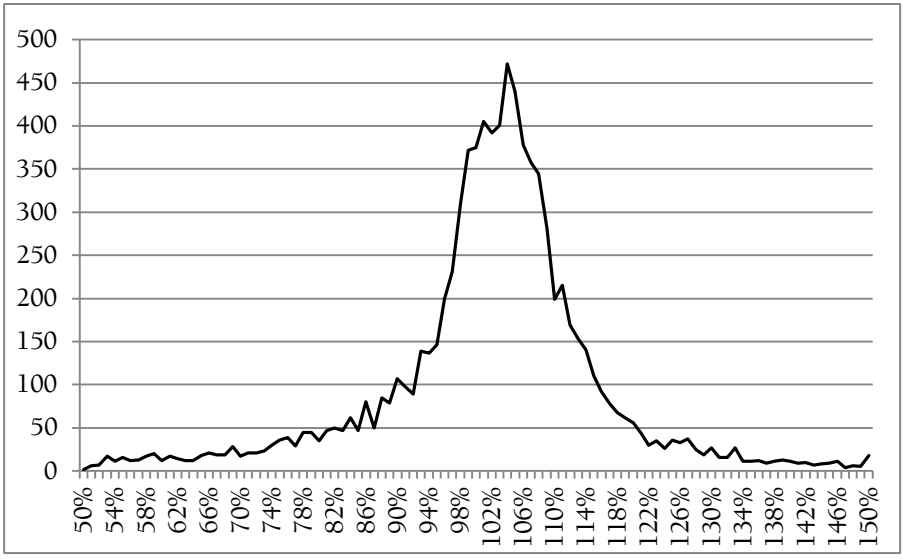

Figure 6. 2010 percentage of registered borrowers as percentage of 2009 (95% of libraries).

The peak is at 104%, As you can see, this is a narrow graph: Not only did most libraries have a higher percentage of registered borrowers, that change is clustered tightly around 103% to 105% and more broadly between 98% and 108%. (If you add up the libraries left off the graph,

while they total almost 5% of libraries they're generally smaller libraries, supporting 1.6% of borrowers and 1.3% of circulation.)

Changes in Whole Numbers

Now we'll go through the changes in whole numbers, perhaps not in as much detail as we did for percentage of patrons registered as borrowers.

Library Service Area

Overall, for the libraries included in this study, LSA rose by 0.7%: 2010 was 100.7% of 2009. The library average is just a bit higher, +0.8%. The median is no change at all. So is the first quartile, while the third quartile is up 0.7%. Standard deviation is 20.5%. The trimmed library average is *very* close to the median, at 100.3% (+0.3%), and the trimmed standard deviation is 6.2%: still high based on the quartiles, but not implausible.

More than four out of ten libraries (3,934) didn't report any change in LSA from 2009 to 2010. 7,384 (80%) saw population change by less than 2.5%; 8,155 (89%) saw it change by less than 5%. In all, 77.5% of libraries either saw no change in LSA or had it grow.

There are always extreme cases, even in one-year (reported) population change. Six libraries saw their service area drop to less than one-fifth its former size, while another 23 saw it cut in half or more. At the other extreme, one library saw its service area population go up more than 12 times and 21 others saw it double or more.

Total Spending

I'm *not* adjusting for inflation, which appears to have been 1.5% for 2010 compared to 2009. Without that adjustment, overall library spending fell by 1.5%—that is, 2010 spending was 98.5% of 2009 spending. But the average library was *up* 5.3%, more than enough to cover inflation. The median was also up, but only 1.1%, not quite enough to cover inflation. The first quartile was down 4.6%; the third, up 7.1%. Standard deviation is 68.9%. Trimming two dozen extreme values at both top and bottom yields a trimmed library average of 103.4% (up 3.4%) and a trimmed standard deviation of 20.7%.

The gap between the overall number and the median is 1,218 (13.2%) while the gap between the median and library average is 1,749 libraries (19.1%). In all, 63% of libraries did better than average on changes in total spending.

Key finding: more than half of America's public libraries spent *more* in 2010 than in 2009 (actually 56%), but slightly less than half kept up with inflation (4,374 or 47.7% had spending grow by at least 1.50%). Libraries keeping up with inflation served 33.3% of borrowers and did 34.8% of circulation.

At the upper extreme, one (very small) library spent 54 times as much in 2010 as in 2009; five more spent at least ten times as much (one of them *not* that small) and 107 more at least doubled spending in 2010. At the lower extreme, three libraries spent less than one-tenth as much in 2010 as in 2009 and 35 others spent less than half as much. (The most extreme loss case, a very small library spending 5.4% as much in 2010 as in 2009, also had only 4.6% as many potential patrons.)

The central cluster includes 6,084 libraries (66.3%) serving 69.4% of borrowers and doing 73.1% of circulation.

To get to Figure 7, I dropped 38 libraries (0.4%) that spent less than half as much in 2010 as in 2009 (serving 0.1% of borrowers with less than 0.1% of circulation) and 233 (2.5%) that spent more than 50% more (serving 0.5% of borrowers with 0.5% of circulation—most "extreme cases" are very small libraries). There's essentially a twin peak—509 libraries that round to no change in spending, 507 that round to a gain of 1%. The encouraging thing here: Not only is the peak at or to the right of staying even, the curve falls off more slowly to the right—that is, more libraries did pretty well than did really badly.

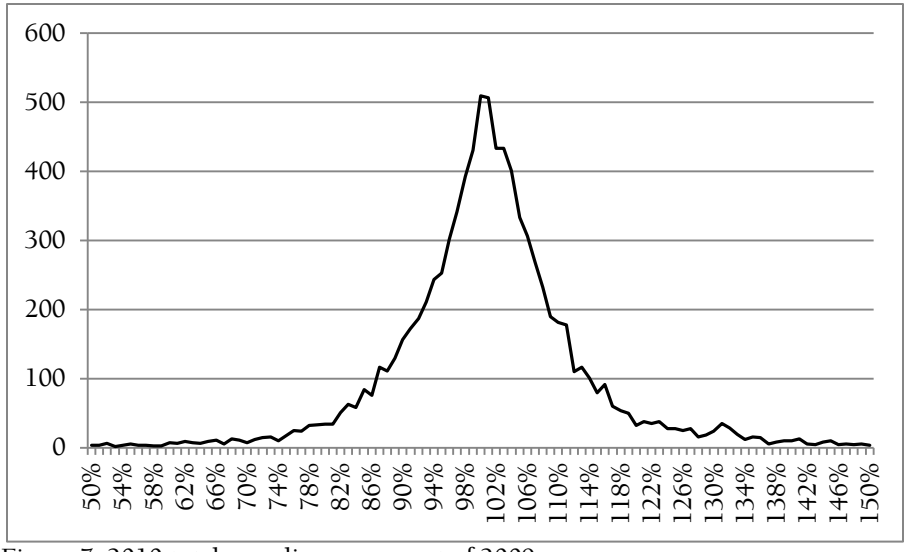

Figure 7. 2010 total spending as percent of 2009

Print Book Collection

For whatever reasons, there were five million fewer print books in the 9,174 libraries considered here in 2010 than in 2009—a drop of 0.5% overall. But the average library had 1.9% *more* books in 2010 than in 2009, suggesting that larger libraries may be weeding more actively. In fact, 62% of libraries had as many or more books in 2010 than in 2009.

The median was an increase of 1.2%, fairly close to the average. Standard deviation was 33.9% (that's related to 101.9%, the average of

2010 figures divided by 2009 figures); there are lots of extreme cases. The trimmed library average is 101.0% (a gain of 1.0%), *very* close to the median—and the trimmed standard deviation is 9.5%.

The first quartile shows a decline of 2.1%—that is, a quarter of libraries lost more than 2.1% of their books—and the third quartile a growth of 4.1%.

The gap between overall change and median change includes 1,478 libraries (16.1%), while the gap between median and average includes 667 libraries (7.3%).

As you'd expect, most libraries stayed about the same. Given that library collections tend to grow, let's consider the range from -2.5% to +5.0%: It includes 5,182 libraries, 56% of the total. 1,858 libraries had more than 5.0% collection growth; 2,134 lost more than 2.5% of their print book collections.

The central cluster includes 7,728 libraries (84.2%) supporting 87.3% of borrowers and 88.2% of circulation.

At the upper extreme, three libraries had book collections grow by more than 16 times and 19 more libraries more than doubled. At the low end, three libraries dropped more than 90% of their book collections and 18 more fell by more than half.

Patron Visits

Reported patron visits were down by 17 million from 2009 to 2010—but that's a decline of only 1.1%. And the average library had 6.0% *more* patron visits. The median was also up, by 1.1%. In all, 5,177 libraries (56%) showed some increase in patron visits. Standard deviation, 116.0%, probably isn't meaningful—but it's clear that this is another case with a significant variance between average and median. The trimmed library average is 103.6% (up 3.6%); the trimmed standard deviation, 25.7%.

The first quartile shows a fairly large drop, down 5.1%; the third, a larger increase, +7.7%.

The gap between overall change and the median includes 764 libraries (8.3%); the gap between the median and the average includes 2,237 libraries (24.4%). Most libraries had more patron visits, and as usual relatively few showed extreme changes, with 1,380 (15.0%) falling more than 10% and 1,297 (14.1%) rising more than 15%. The central cluster includes 5,965 libraries (65.0%) supporting 72.5% of borrowers and doing 74% of circulation.

At the low end, 10 libraries saw patron visits essentially collapse, to less than 10% of 2009 counts, while another 107 dropped by more than half. At the high end, one library went up an improbable 103 times and two others went up more than ten times, with another 127 more than doubling in patron visits.

Reference Transactions

Except for the overall change, I'm ignoring 108 libraries that reported zero reference transactions in either 2009 or 2010, so the studied universe is down to 9,066 libraries. (Those 9,066 libraries account for 99.93% of borrowers and 99.97% of circulation: the missing 1.2% are *very* small.)

The overall change was negative—but tiny. Reference transactions were down about 800,000, or 0.3%. For 9,066 libraries actually reporting reference transactions in both years, the average among library changes was a *gain* of 37.4%. The median is a gain of 0.4%. A majority of public libraries with any reference service *at all* (53.1%) had slightly more reference transactions in 2010 than in 2009.

The standard deviation is 641.8.% The trimmed library average is 115.1%, a gain of 15.1%; the trimmed standard distribution is 114.18%.

The first quartile is 88.7%, that is, a decline of 11.3%. The third quartile is 113.3%, a gain of 13.3%—still way below the average, but this indicates that changes in reference activity vary widely among many libraries.

It may be erroneous to show the gap between overall average and median, since the overall average includes libraries not included here (ones that had reference activity in one of the two years, but not both), but there are 339 libraries (3.7%), inclusive, between 99.73% and 100.43%—and 3,453 (38.1%) between 100.43% and 137.36%. (Here as throughout, I calculate gaps based on results to two decimal places but normally only report one decimal place.) You could also say that 88.1% of the libraries were "below average" in terms of changes in reference service!

Looking at extremes, in addition to libraries where there was some reference activity in 2009 and none in 2010 (no more than 20 such libraries) and those with none in 2009 (88 in all, some of which were probably among the 80 total with none reported in 2010), there are 37 libraries reporting less than one-twentieth as much reference activity in 2010 as in 2009 and another 135 where transactions dropped by at least three-quarters. At the other extreme, six libraries reported reference transactions increasing by at least one hundred times and 51 more where it increased more than tenfold.

The central cluster includes 4,010 libraries (44%), supporting 52.4% of borrowers and 54.2% of circulation.

Eliminating 466 libraries where reference traffic dropped by more than half and a surprising 861 libraries where it *grew* by more than half, we arrive at the rounded graph in Figure 8.

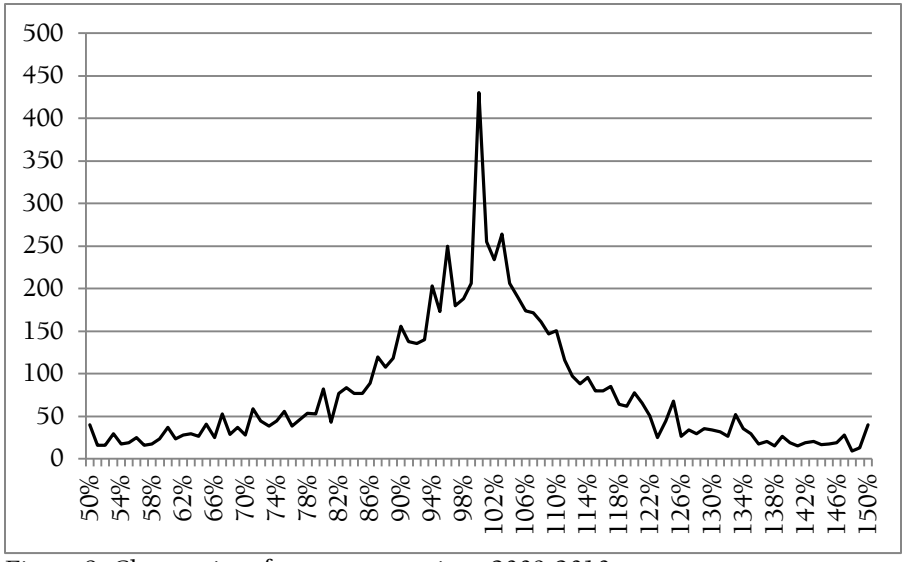

Figure 8. Changes in reference transactions 2009-2010

The peak is at 100% (430 libraries reported changes of less than 0.5% in either direction—considerably higher than the second-highest peak, 264 libraries at 103%). Do note that there are a *lot* of libraries omitted from the graph, almost 15% in all, most of them off to the right of the graph.

Registered Borrowers

Overall, the number of registered borrowers increased by 1.0%—which isn't much, but more than the overall growth in potential patrons (0.7%). But as we've already seen, libraries vary enormously in the percentage of legally-assigned potential patrons who are actually registered borrowers!

The average library saw a 6.2% increase in borrower registration, considerably larger than the overall average. The median was a healthy 3.4% increase: By *any* measure public libraries had more registered borrowers in 2010 than in 2009, even normalizing for population changes. In fact, 6,095 libraries or 66.4% (essentially two-thirds) either stayed the same (about 75 libraries) or showed more registered borrowers. Those libraries serve 70.3% of borrowers and do 67.3% of circulation.

For what it's worth, the standard deviation is 108.7%. The trimmed library average is a gain of 3.1%—so close to the median as to suggest normal distribution. The trimmed standard distribution is 22.29, and that's not unreasonable.

The first quartile is 97.4%, so more than a quarter of libraries dropped at least 2.5% of registered borrowers. The third quartile is 108.8%, a gain of nearly 9% in registered borrowers.

Looking at the gaps, the seemingly small range between overall average and median—from 1.0% gain to 3.4% gain—includes 1,074 libraries (11.7%) while the gap between median and library average includes 1,255 libraries (13.6%).

The central cluster includes 6,347 libraries (69.2%) accounting for 74.5% of borrowers and 73.8% of circulation.

As with most measures, the story behind extremes might be interesting, but all I can do is provide the numbers. One library had 80 times as many registered borrowers in 2010 as in 2009 (and more than 40 times as many registered borrowers as its LSA!); seven others had more than ten times as many registered borrowers in 2010 and 74 more libraries more than doubled registration. At the other extreme, nine libraries had less than one-tenth as many registered borrowers in 2010 as in 2009 and 155 more saw registered borrowers drop by more than half.

Eliminating 164 (1.8%) libraries where the rounded change is more than -50% and 201 (2.2%) where it's more than +50%, we get the graph in Figure 9—a reasonably classic normal distribution, significantly to the right of the no-change point. Note that the 4.0% of libraries not shown on the graph account for 1.4% of borrowers and 1.0% of circulation—they're mostly small libraries. The twin peak is at 104% and 105%.

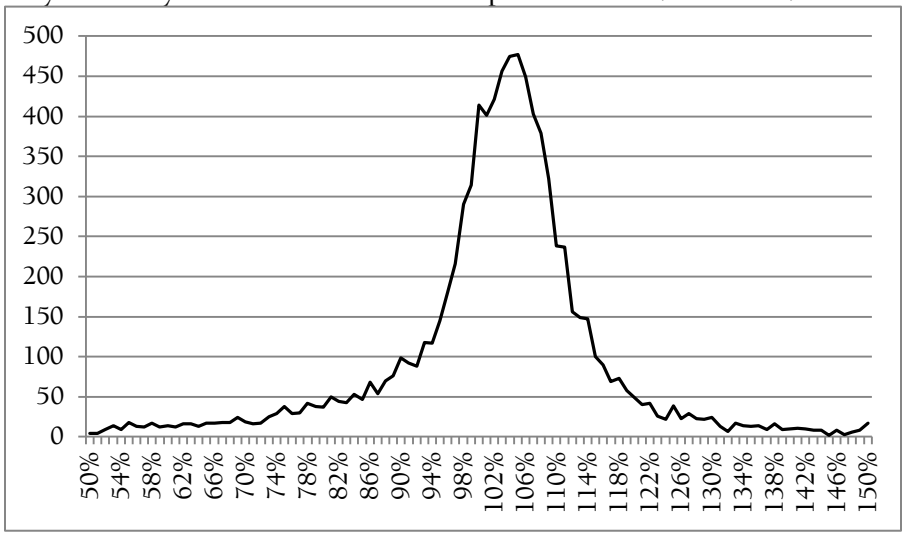

Figure 9. Changes in registered borrowers 2009-2010

Circulation

Overall, circulation grew by 2.1%, a number you may have seen, since it's the Big Number for public libraries. For these 9,174 libraries, that's a growth of some 51.6 million items circulated, up to 2.465 billion in 2010.

The average library increase was 6.3%, an even healthier number. The median? 1.3% growth, a bit less than overall growth but still indicating that, not only did public libraries circulate more items as a whole, so did most individual libraries. The first quartile is 94.6% (a drop of 5.4%); the third quartile, 108.6% (an increase of 8.6%).

The standard deviation is 157.6%. The trimmed library average is 103.2%, a gain of 3.2%; the trimmed standard deviation is 20.5%.

The gap between median and overall average includes 365 libraries (4.0%); the gap between overall and library average, 1,380 libraries (15%).

The central cluster includes 5,836 libraries (63.6%) supporting 78.0% of borrowers and doing 79.8% of circulation.

In all, 5,156 libraries (56.2%) increased circulation (or at least didn't see it fall), supporting 60.4% of borrowers with 63.5% of circulation.

But that *does* mean that 43.8% of libraries saw at least slight declines in circulation. Some 1,325 (14.4%) saw circulation drop by more than 10%; those libraries support 7.2% of borrowers and 5.2% of circulation.

At the extremes, some libraries show phenomenal percentage increases or collapses in circulation—and, as usual, these are mostly (but not entirely) very small libraries. Libraries where circulation fell by more than half number 81 (0.9%) and support 0.2% of borrowers with less than 0.1% of circulation. Those where it grew by more than half number 245 (2.7%), but those are still smallish libraries (by and large), supporting 0.7% of borrowers and 0.5% of circulation.

Excluding those extremes, we get Figure 10.

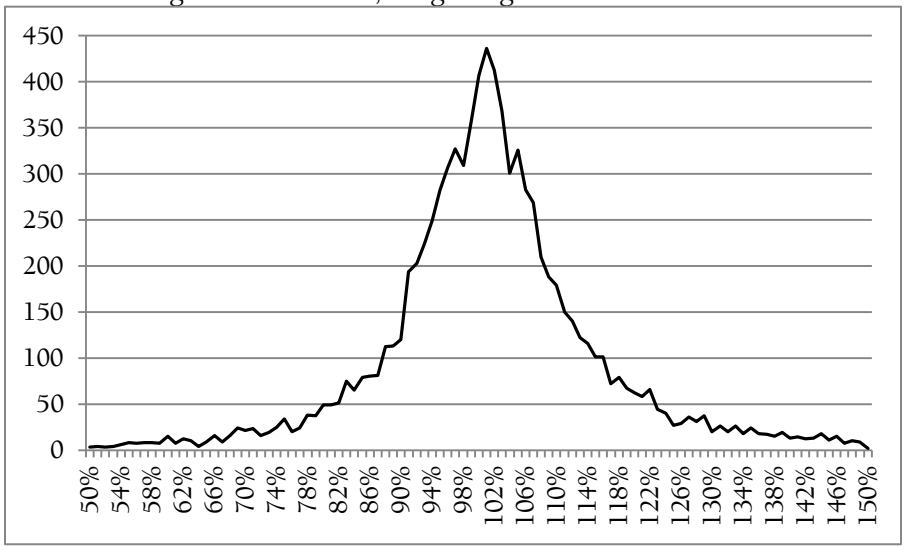

Figure 10. Changes in circulation 2009-2010

The peak in Figure 10 is at 101%.

Programs

This is another measure where some (mostly smaller) libraries either didn't report or reported no programs at all in one of the two years. The remaining 8,973 libraries account for 99.74% of the borrowers and 99.52% of circulation; the other 201 libraries (2.2%) are generally very small.

Overall, programs were up 1.4%. The library average was up 25.7%, but the median was up 1.4%, slightly less than the overall average. (The overall average includes the other 201 libraries, so I won't calculate the tiny gap between that and the median.) First quartile is 88.4% (that is, down 11.6%); third quartile is 120.0% (up 20%--still less than the library average). Standard deviation is meaningless (252.3%). The trimmed library average is 115.9%, a gain of 15.9% and still far from the median; the trimmed standard deviation, 86.7%.

The gap between median and library average includes 2,633 libraries (29.3% of those with at least one program in each year), serving 37.3% of borrowers with 39.7% of circulation.

The central cluster includes 3,438 libraries (38.3%) serving 47.9% of borrowers and doing 50.9% of circulation—generally larger libraries, in other words.

In all, 5,241 libraries (58.4%) having any programs in both years had at least as many in 2010 as in 2009; those libraries served 53.1% of borrowers and did 54.5% of circulation, so they're distributed among all sizes.

At the extremes, including libraries with more than ten times as many programs in 2010, 949 libraries (10.6%) offered more than 50% more programs in 2010; those libraries served 4.5% of borrowers with 3.9% of circulation, so they're generally smaller. On the lower end, 320 libraries (3.6%) offered less than half as many programs in 2010; those are also generally smaller libraries, serving 1.8% of borrowers with 1.4% of circulation. Removing those libraries results in Figure 11.

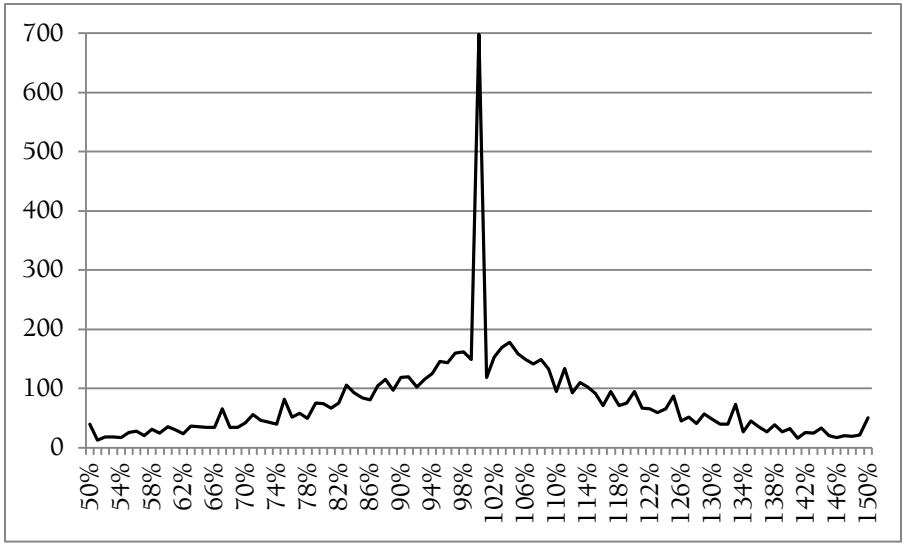

Figure 11. Changes in programs 2009-2010

In this case, Figure 11 points up something that's not quite as clear when scanning the numbers (which are to two more decimal places): that huge spike at 100%, meaning that nearly 9% of the libraries varied number of programs by less than half a percent in either direction.

But it's also an interesting case of a graph possibly overstating the case unless you look at the numbers—because that spike is less than 10% of the total, and the rest fan out fairly widely on either side of the spike rather than having a steep curve on either side: the values for 99% and 101% are 149 and 118 respectively (as you can see if you look closely at the graph, the values for each percentage point from 102% to 109% are actually *higher* than for 101%).

Program Attendance

The next to last of the direct measures: how many people attended the library's programs. Since libraries with no programs are likely not to have any program attendance, this measure also involves 8,973 libraries rather than 9.174.

Overall (including libraries with programs in one year but not the other), program attendance was up 0.5%—a bit less than programs. As with programs, the library average is *considerably* higher at 120.2% (a 20.2% gain). The median is 102.2%, up 2.2%.

The first quartile is 87.2%—for a quarter of libraries, attendance at programs was down at least 12.8%. The third quartile is 120.3% (up 20.3%), essentially the same as the library average.

The standard deviation is meaningless at 215.2%. The trimmed library average is 113.0%, a gain of 13.0%; the trimmed standard deviation,

70.7%. The gap between the median and library average includes 2,236 libraries (24.9%), not surprising given that the library average is just barely under the third quartile.

The central cluster—in this case ranging from libraries with an 8.06% drop in program attendance to those with a 12.38% increase—includes 3,143 libraries (35%) that account for 43.7% of borrowers and 47.8% of circulation: They're typically somewhat larger or busier than average.

4,864 libraries (54.2%) saw program attendance hold steady or increase; those libraries account for 48.0% of borrowers and 49.2% of circulation.

375 libraries (4.1%) saw program attendance fall by more than 50%; those libraries are generally small, accounting for 1.7% of borrowers and 1.1% of circulation. At the high end, a large number of libraries—1,002 (11.2%)—had an increase of more than 50% in program attendance. Those libraries are also, by and large, small, serving 5.1% of borrowers with 4.1% of circulation.

Removing those libraries leaves 7,596 libraries (84.7%) that account for 93.2% of borrowers and 94.8% of circulation; for those libraries, changes in program attendance appear in Figure 12.

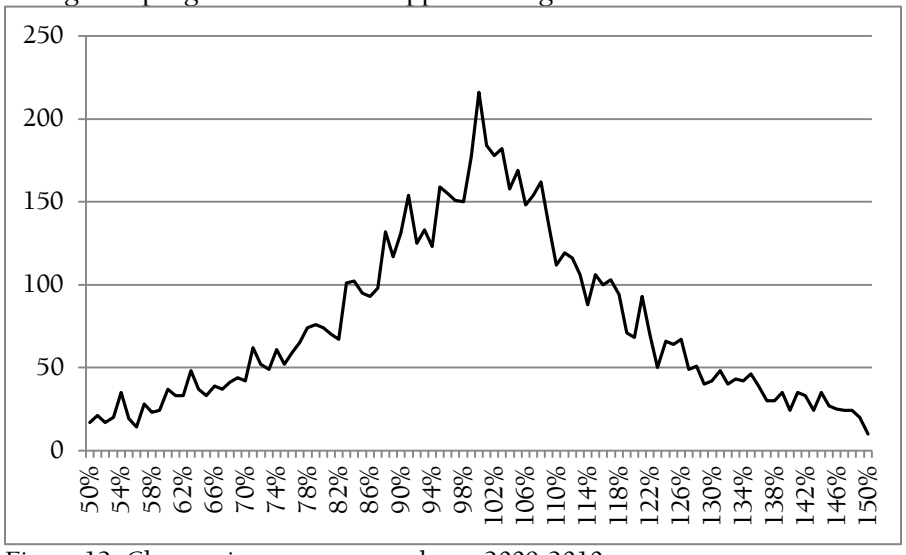

Figure 12. Changes in program attendance 2009-2010

What I find most interesting here is the difference between Figure 11—changes in actual program count—and Figure 12—changes in program attendance. While the peak is once again at 100%, that peak (216 libraries) is only slightly higher than the levels on either side (less than 20% higher), not the *enormous* difference for programs. Once again, the curve below the peak is tilted to the right.

Personal Computer Sessions

The last of the direct measures, again one where some libraries reported no sessions during 2009, 2010 or both (some very small libraries don't *have* any public access computers with internet access). That includes 101 (1.1%) libraries that serve 0.1% of borrowers with less than 0.1% of circulation—in other words, mostly *very* small libraries. Thus, the universe for this discussion is 9,073 libraries serving 171.09 million borrowers with 2.463 billion circulation. The overall average *including* libraries with no reported sessions in one year, but some in the other year, is 100.6% (that is, up 0.6%).

The library average among the 9,073 libraries is a fairly astonishing 277.1%—an increase of 177.1%, no doubt affected by 135 libraries where PC use more than tripled from 2009 to 2010 (and seven cases where it increased by ten times or more). The median is 101.5%, an increase of 1.5%. The first quartile is 90.3%; the third is 113.8%. Which is to say that roughly a quarter of libraries saw a drop of 10% in counted computer sessions and another quarter saw an increase of roughly 14% or more. (The standard deviation is even more absurd than the library average, at 12,598.6%!)

The trimmed library average is 110.5%, a gain of 10.5%—a case where dropping a few outlying (but presumably real) reports makes the library average a whole lot more plausible. The trimmed standard deviation, 72.6%, may still be extreme, but it's a *lot* better than the untrimmed version.

It's useful to note here that PC session counts presumably only include sessions on library-provided public access computers. They won't normally include all the computing done by patrons using their own devices and the library's WiFi. That's likely to be a growing factor in library service, especially in more affluent neighborhoods—and with WiFi-capable tablets costing $150 or less, maybe even in less affluent communities.

The gap between the median and library average includes 4,377 libraries (48.4%) accounting for 42.5% of borrowers and 42.0% of circulation.

The central cluster—ranging from a drop of 8.69% to an increase of 11.6%—includes 4,065 libraries (44.8%) supporting 50.0% of borrowers with 50.7% of circulation.

In all, 5,075 libraries (55.9%) had as many or more PC sessions in 2010 as in 2009 (which also means that 44.1% had fewer); those libraries serve 51.0% of borrowers with 49.5% of circulation.

As usual, there are extremes at both ends—e.g., five libraries with less than one session in 2010 for every *hundred* in 2009. 237 libraries (2.6%), serving 1.6% of borrowers with 1.5% of circulation, saw a drop of more than 50% in PC sessions—and 743 more (8.2%), representing

5.4% of both borrowers and circulation, had more than 50% increase. Removing those outliers yields the graph in Figure 13.

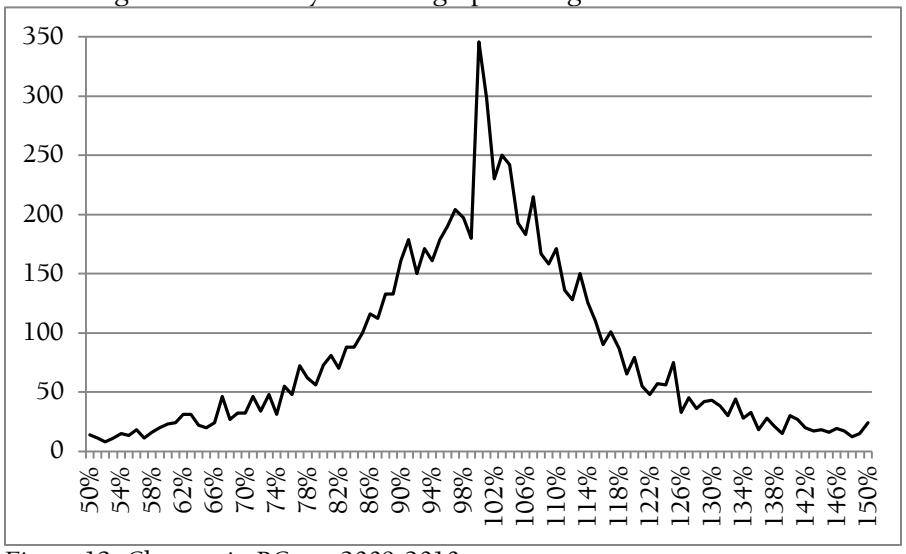

Figure 13. Changes in PC use 2009-2010

Once again, the peak is at 100%; this time, it's 16% higher than the next highest value, with the curve higher to the right of 100% than to the left.

That completes the discussion of direct measures. I was wrong in assuming that very few public library distributions would be something like normal; in fact, most distributions of *change percentages* are some variation of normal, although distributions of actual numbers are, pretty much automatically, not.

Per Capita

In the interest of making this essay less than book length, I'm going to skip per capita figures (that is, whole numbers divided by LSA). Most (not all) of them are addressed in *Give Us a Dollar and We'll Give You Back Four* and its companion *Graphing Public Library Benefits*, although not in the same manner as I'm doing here. You *will* find the median, first quartile and third quartile for most of these metrics (and a couple of others) for most libraries in those books.

Instead, I'd like to focus on a somewhat different set of derived measures, ones I haven't seen noticed much: measures on a *per borrower* basis, dividing whole numbers by the number of registered borrowers. That yields higher numbers than per capita figures for most libraries—but sharply lower ones for those libraries with many more borrowers than their LSA would suggest. These discussions will combine per borrower figures for 2010 and changes from 2009 to 2010.

Measures appear in the same order as in previous sections, with four changes:

- Library Service Area doesn't enter into these discussions.
- I'm adding notes on collection spending per borrower.
- Registered borrowers per borrower are always 1, so that measure is ignored.
- While I do discuss program attendance, I omit number of programs.

Spending Per Borrower

Three libraries report spending less than $1.00 per borrower in FY2010 (two of them tiny, one—spending *a dime* per borrower—not so small). Ten libraries report spending more than $1,000 per borrower, including one at more than $10,000—but all of them have very small numbers of borrowers (six of the ten have fewer than 100).

Everybody else is in between, but that's a *big* in between, albeit not quite as big as the range of actual spending.

The overall average in FY2010 was $62.93. The library average is higher, but not *much* higher: $70.32. The median is lower: $52.91. It's interesting that the three figures are roughly equally spaced, but perhaps not significant. (If you remember that only about one-tenth of American public libraries spend at least $73 per capita, remember that these are *per borrower* figures.)

The first quartile is $33.96—one-quarter of U.S. public libraries spend less than $34 per capita. The third quartile is $81.87. The standard deviation—which should be more meaningful, as this isn't a percentage—is $155.30. The trimmed library average is $68.62, not a big change; the trimmed standard deviation is $111.19, not a big enough change to make it a useful figure.

The gap between median and overall average includes 991 libraries (10.8%) that account for 11.4% of borrowers and 12.2% of circulation. The gap between overall average and library average includes 584 libraries (6.4%) that account for 6.6% of borrowers and 7.7% of circulation.

The central cluster around the median includes 1,139 libraries (12.4%) accounting for 12.2% of borrowers and 12.0% of circulation. Note that this 20% range—10% on either side of the median—includes only 12% of libraries.

Rather than look at the true extremes, I'll divide things into five areas:

- Libraries spending less than $10 per borrower: 172 (1.9%), accounting for 1.2% of borrowers but only 0.2% of circulation. Given that spending and circulation *per capita* typically correlate fairly well, this should be no surprise.

- Libraries spending from $10 to half the median ($26.46): 1,236 (13.5%), accounting for 10.5% of borrowers and 4.4% of circulation.
- Libraries spending from half to twice the median: In other words, the *very* broad middle, from $26.47 to $105.81. This group includes 6,406 libraries (69.8%) that account for 77.4% of borrowers and 76.2% of circulation.
- Libraries spending more than twice the median but less than $200 per borrower: 1,121 libraries (12.2%), accounting for 9.6% of borrowers—and 16.4% of circulation.
- Libraries spending $200 or more per borrower: 238 libraries (2.6%), accounting for 1.3% of borrowers and 2.8% of circulation.

Note that libraries at both extremes have somewhat fewer borrowers than those in the middle: As in most other measures, extreme cases are typically smaller libraries. (Of the bottom ten, for example, only one has more than 620 registered borrowers and three have fewer than 75.)

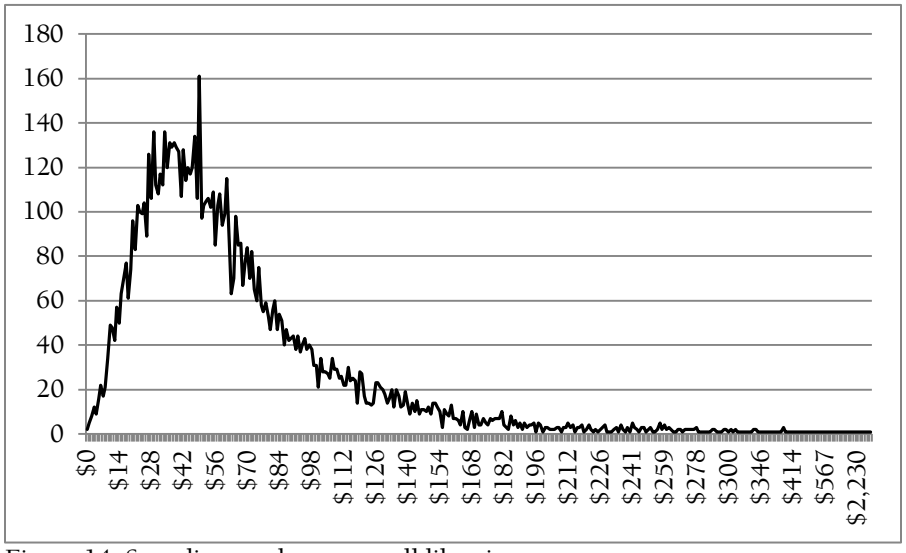

Figure 14. Spending per borrower, all libraries

Figure 14 includes all libraries. The horizontal axis is only linear up to $200; after that, because it includes only actual values, it's extremely and increasingly non-linear (e.g., the range from $538 to $4,318 gets the same space as the range from $20 to $40). The peak on the graph is at $49, somewhat to the left of the median and well to the left of the library average.

Let's normalize this somewhat: Drop off roughly 15% at each end (libraries spending either less than half as much as the median or more than twice as much) and round spending to the nearest $2 rather than the nearest $1 (as in Figure 14). That yields Figure 15, showing spending per borrower for 69.8% of libraries.

Figure 15. Spending per borrower for 70% of libraries

I find Figure 15—which I prepared with no idea what would actually happen—interesting in a couple of ways. First, the "peak" seems to be even lower than before. Second, this doesn't look like a normal distribution, and that's right. The actual median here (for the raw data, without the highs and lows but before rounding to the nearest $2) is a bit higher—$53.10—while the library average is lower ($56.61). Neither would suggest much about the actual distribution, which I'd call "roughly equal from $30 to $56, then declining fairly slowly from there." Remember: Even in this group, 50% of the libraries spend at least $53.10; if it feels as though most of the space under the curve is to the left of $50—and that's how it looks to me at first glance—then the graph is correct but misleading.

Changes from 2009 to 2010

A library can see extreme *change* in spending per borrower on a year-to-year basis either because of unusual funding changes—or because of unusual changes in number of registered borrowers. That said, the changes cover quite a range: eight libraries where 2010 spending per borrower was less than one-tenth of 2009—and 13 others where it was at least ten times as high. All of the huge increases are in libraries with fewer than 500 borrowers; two of the huge drops are in libraries with more than 50,000 borrowers. (Both libraries have at least 15 times as many borrow-

ers as legal potential patrons; both saw registration jump by more than 25 times—both odd cases.)

The overall average change in spending per borrower is -2.4%: 2010 spending per borrower is 97.6% of 2009 spending per borrower. But the library average is 109.7, a gain of 9.7%. The median is close to the overall average: 98.2%, a loss of 1.8%. The first quartile is 89.7% (a drop of 10.3%); the third quartile is 109.3% (a gain of 9.3%).

The standard deviation is 108.7%. The trimmed library average is 106.2%, that is, a gain of 6.2%; the trimmed standard deviation, 43.3%.

The gap between overall average and median includes 189 libraries (2.1%) that account for 2.1% of borrowers and 2.7% of circulation. The gap between median and library average includes 2,350 libraries (25.6%) that account for 26.8% of borrowers and 29.8% of circulation. The central cluster includes 4,726 libraries (51.5%) that account for 58.0% of borrowers and 60.3% of circulation.

- 4,094 libraries (44.6%) spent at least as much per borrower in 2010 as in 2009; those libraries account for 33.9% of borrowers and 36.6% of circulation. They're *somewhat* smaller libraries, by and large.
- 121 libraries (1.3%) spent less than half as much per borrower in 2010 as in 2009; those libraries account for 0.5% of borrowers and 0.3% of circulation. They're mostly considerably smaller libraries.
- At the other end, 587 libraries (6.4%) spent more than half again as much per borrower (that is, >150.0%) in 2010 as in 2009. Those libraries account for 2.2% of borrowers and 2.8% of circulation—still somewhat smaller libraries, but active ones.

Eliminating the 7.7% of libraries at either end yields Figure 16. The highest peaks are at 97% and 99%.

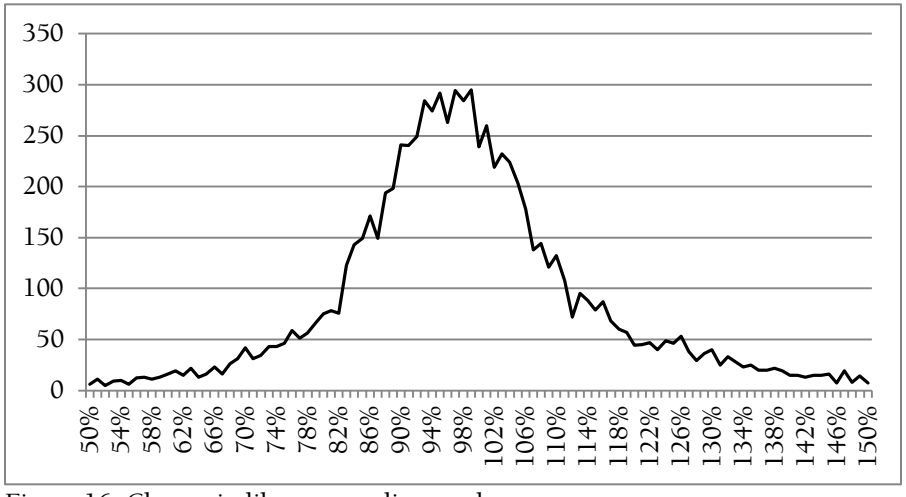

Figure 16. Change in library spending per borrower

How does this compare with Figure 7, change in library spending (*not* on a per-borrower basis)? Not all that well: Figure 7 has a much narrower and taller central area.

Collection Spending Per Borrower

Thirty-eight libraries report spending nothing at all on collections in FY2010, while three libraries spent more than $500 per borrower (all with fewer than 100 borrowers).

Overall, including those non-spending libraries, libraries spent $7.35 per borrower. The library average is higher but not absurdly so, at $9.05. The median is $6.76 (again including libraries that spent nothing at all), just a bit lower than the overall average. The first quartile is $3.91; third is $10.87. Standard deviation is $16.62. The trimmed library average is $8.53, while the trimmed standard deviation is $7.11.

The gap between median and overall average includes 402 libraries (4.4%) serving 4.6% of borrowers with 4.9% of circulation. The gap between overall average and library average includes 1,008 libraries (11.0%) serving 10.6% of borrowers with 11.9% of circulation.

The central cluster—which in this case means libraries spending $6.08 to $7.44 per borrower on collections—includes 955 libraries (10.4%) serving 8.7% of borrowers with 8.9% of circulation.

On the low side—spending less than $3.38 per capita, or half of the median—are 1,838 libraries (20.0%) accounting for 22.5% of borrowers but only 10.5% of circulation: That's not terribly surprising.

On the high side—spending at least $13.52 per capita, or *twice* the median—are 1,518 libraries (16.5%) serving 9.7% of borrowers, but with 18.8% of circulation—also not terribly surprising. Libraries that spend more on materials tend to circulate more.

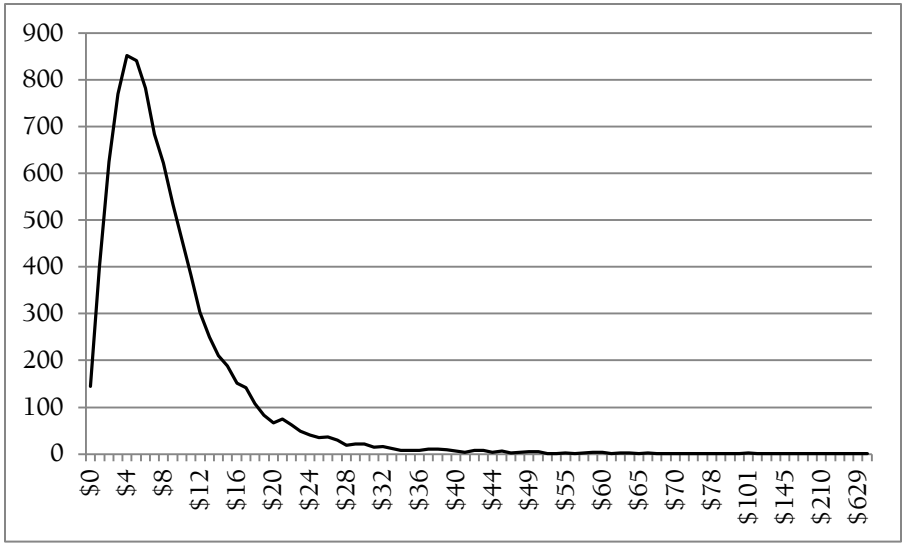

Figure 17. Collection spending per borrower

Figure 17 is interesting and, as far as it goes, reasonably meaningful—but note the shape of the curve: very steep on the left side of the central area (the peak figures are at $4 and $5, to the left of the median) and much shallower on the right, going off into a long tail of relatively few libraries spending much more money. (The horizontal axis is only linear up to $42.)

What if we drop some cases (let's say the bottom 10% and top 10%) and redraw the chart, this time to $0.50 increments? The remaining libraries—80% of the whole, serving 85.1% of the borrowers with 86.2% of the circulation—all round to somewhere between $2 and $17, yielding the graph in Figure 18.

Figure 18. Collection spending per borrower (for 80% of libraries), rounded to nearest fifty cents

Changes from 2009 to 2010

As a whole, collection spending per borrower declined by a significant percentage, with 2010 spending being 95.2% of 2009 spending—a decline of 4.8%. That includes 14 libraries with no reported spending in 2010 and 40 others with no reported spending in 2009 (those 54 libraries account for only 0.03% of borrowers—about 59,000 out of 171 million—and roughly 0.01% of circulation, about 250,000 out of 2.465 *billion*).

For the remaining 9,120 libraries, those reporting *some* spending on collections in both years, the library average is a *gain* of 9.6% (2010 spending per borrower was 109.6% of 2009 spending per borrower). The median, the most meaningful figure, is 96.5% (a loss of 3.5%), making the library average mostly useless.

The standard deviation is 115.4%. The trimmed library average is 105.6% (a gain of 5.6%, still a *long* way from the median) and the trimmed standard deviation is a still-useless 58.3%.

The first quartile is 82.7% (a drop of 17.3%, so a lot of libraries cut collection spending per borrower substantially). The third quartile is 112.6%--that is, a quarter of libraries *raised* collection spending per borrower by more than 12.5%, more than one-eighth. "All over the place" may be a good summary for changes in library spending per borrower.

The gap between overall average and median includes 277 libraries (3.0%) that serve 3.5% of borrowers with 3.4% of circulation. The gap between median and library average includes 1,977 libraries (21.7%) serving 21.0% of borrowers with 23.7% of circulation.

The central cluster includes 3,324 libraries (36.2%) that serve 37.2% of borrowers with 40.0% of circulation.

Overall, 3,892 libraries (42.7%) saw no drop or some gain in collection spending per borrower; those libraries serve 34.5% of borrowers with 37.6% of circulation, so they are *on the whole* a little smaller than libraries that did see reductions.

The extremes are, as usual, *very* extreme (20 libraries spent less than 10% as much per borrower on collections in 2010 as in 2009, including one down to 1%; 20 others spent more than nine times as much, including one at 50 times as much for its, well, *25* registered borrowers).

Let's drop a relatively small group at each end in order to get a possibly workable graph. At the low end, 367 libraries (4.0%) dropped by more than half; those libraries served 3.4% of borrowers with 2.2% of circulation. At the high end, 331 libraries (3.6%) more than doubled collection spending per borrower; those libraries served 1.5% of borrowers with 1.9% of circulation. Dropping those yields the graph in Figure 19 (noting that percentages are expressed as decimal numbers, because this is a direct PivotGraph, and you can't change the way numbers are displayed). The highest peak is at 0.97 (97%), a 3% loss.

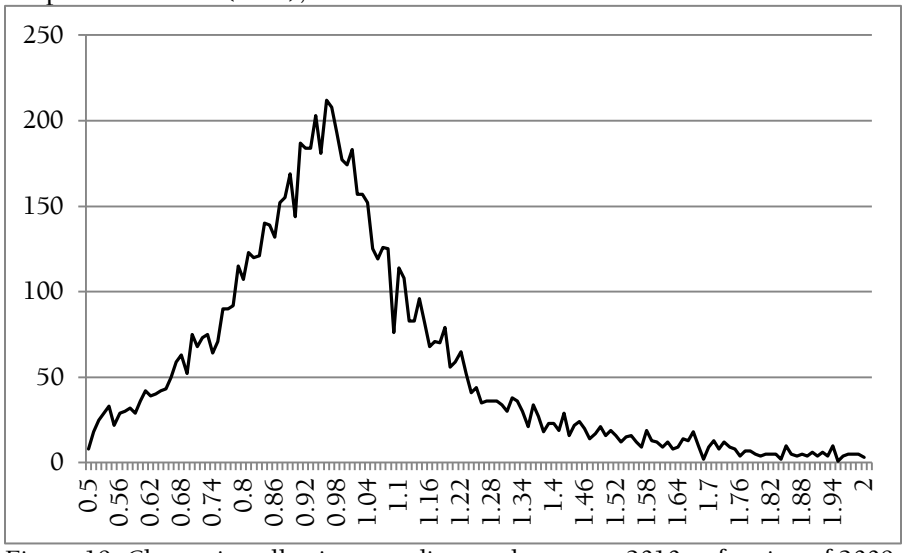

Figure 19. Change in collection spending per borrower, 2010 as fraction of 2009, excluding 7.6% of libraries

Print Books per Borrower

Here's a somewhat unusual measure, at least in my experience: The extent of a collection measured on a per-borrower basis. The overall average is 4.7—that is, 4.7 books per borrower across all libraries.

The library average is *much* higher, 12.4. The median is also much higher than the overall average: 8.0 books per borrower.

The first quartile is 5.0, so more than three-quarters of libraries have "more books per borrower than average," if the average is the national overall average. The third quartile is 13.3, so you could also say that something shy of three-quarters are *below* average in books per borrower, if you're using the other average. (The actual figure: 6,633 libraries or 72.3% have 12.4 books per borrower or less.) Since standard deviations should be somewhat meaningful for non-percentage numbers, here it is for this measure: 21.1, or more than 2.6 times the median. The trimmed library average is 11.7, still considerably higher than the median; the trimmed standard deviation is 13.0.

The gap between overall average and median (includes 2,512 libraries (27.4%) that serve 24.3% of borrowers with 27.1% of circulation. The gap between median and library average includes 2,040 libraries (22.2%) that serve 7.8% of borrowers with 10.8% of circulation. Libraries in the first gap are roughly typical; libraries in the second gap are typically smaller and relatively well used.

The central cluster isn't very large in absolute or relative terms: ranging from 7.2 to 8.8 books per borrower, it includes 1,015 libraries (11.1%) serving 7.3% of borrowers with 8.2% of circulation.

As with almost every measure, there are some *true* extremes. At the low end, 21 libraries (0.2%) with more borrowers than books serve 0.9% of borrowers, but only 0.09% of circulation (that is, they're used about one-tenth as much as is typical). At the high end, 57 libraries (0.6%) have more than 100 books per borrower; those libraries serve 0.005% of borrowers (only one has more than 1,000 borrowers and most have fewer than 100) with 0.033% of circulation: both tiny numbers, but the relative use is about six times average. (Those percentages are so small that I should include actual numbers: The libraries with *lots* of books per borrower serve 8,904 borrowers with 809,472 circulation—including one library with 291 borrowers and nearly 85,000 circulation!)

While I was tempted to cut off the bottom 5% (libraries with fewer than 2.68 books per borrower) and the top 5% (libraries with at least 33 books per borrower) to make a more readable chart, the fact is that the libraries with relatively few books per borrower include quite a few large and very large libraries. That may make sense: A system with 3.3 million borrowers may serve them well with 6.3 million books. (At the top end, on the other hand, the 456 libraries—which are 5% of the libraries—with at least 33 books per borrower serve a mere 0.13% of borrowers and 0.26% of circulation: They're heavily used but almost entirely very small libraries.)

So we'll leave in all but the 78 at the extreme top and bottom, 0.8% of libraries in all, and round to the nearest whole number to produce Figure 20.

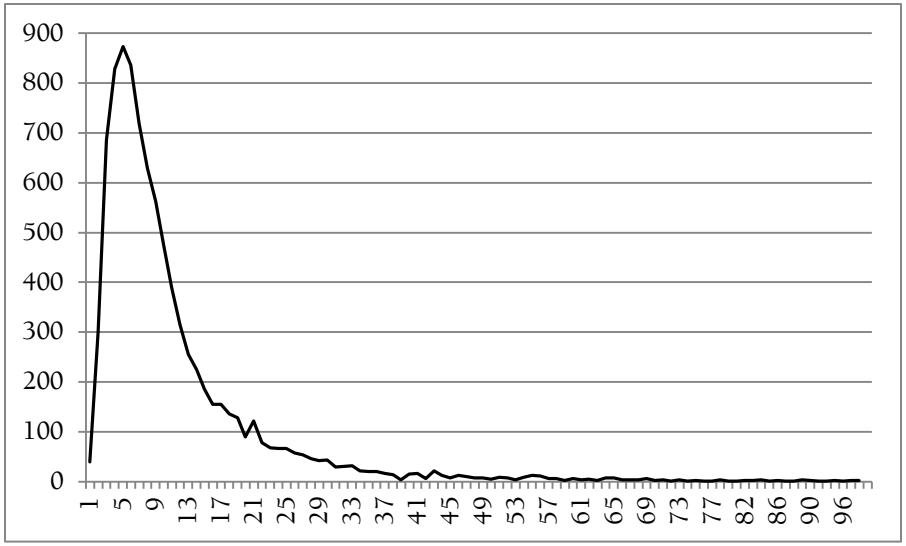

Figure 20. Books per borrower for 99.2% of libraries

The peak (the mode) is at five books per borrower, and as you can see the curve is sharply skewed to the right. Since we already knew that 75% of the libraries have more than five books per borrower, this is no surprise.

Changes from 2009 to 2010

You wouldn't expect extreme figures here—unless, as with percentage of patrons that are registered, there's been some sudden change or, possibly, massive weeding. Since the total size of collections shrank a little and borrowers grew a little, the overall average is necessarily down—but not a lot: 98.5%, down 1.5%. The library average is up a fair amount, at 106.4% (up 6.4%). The median is down a little more than the overall average, at 98.1% (down 8.9%). The first quartile is 91.6% (one out of every four libraries had a drop of more than 8% in books per borrower), while the third quartile is 105.2% (a gain of 5.2%). This is another case where more than three-quarters of libraries (actually 79.5%) are below average—based on the library average.

The standard deviation is 98.5%. The trimmed library average is 103.2% (up 3.2%); the trimmed standard deviation is 35.1%.

The gap between median and overall average includes a mere 154 libraries (1.7%), accounting for 2.4% of borrowers and 1.8% of circulation. The gulf between overall average and library average includes 2,339 libraries (25.5%) serving 24.6% of borrowers with 25.4% of circulation.

The central cluster includes 5,686 libraries (62.0%). Those libraries serve 62.3% of borrowers with 64.9% of circulation. In this case, nearly two-thirds of libraries *are* "somewhere around the median." In all, 3,852

libraries (42.0%) held steady or gained on this measure; those libraries serve 35.6% of borrowers and do 37.5% of circulation, so they're (as a whole) *slightly* smaller libraries.

At the low end, 98 libraries (1.1%) saw a drop of more than half in books per borrower; those libraries serve 0.5% of borrowers with 0.2% of circulation. At the high end, 406 libraries (4.4%) saw an increase of more than half in books per borrower; those libraries serve 1.9% of borrowers with 2.4% of circulation. As is frequently the case, extremes tend to be smaller libraries.

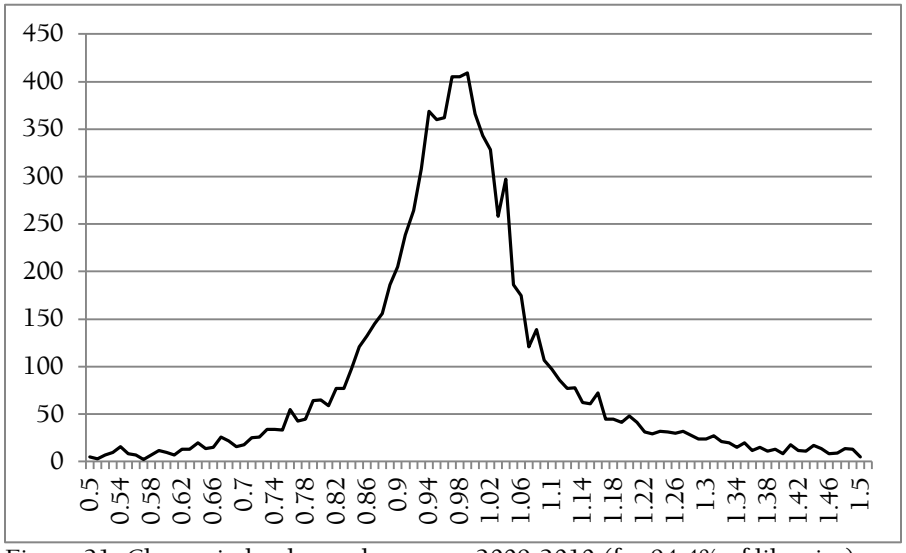

Figure 21. Change in books per borrower, 2009-2010 (for 94.4% of libraries)

Figure 21 excludes the low and high extremes just noted. The peaks are at 0.98 and 0.99 (98% and 99%).

Visits per Borrower

The previous three measures are *input* measures, showing what a library provides for its patrons. I've omitted a couple of others that only scale on a per-1000-patron or per-1000-borrower basis: available personal computers and total programs. The remaining five measures are, at least to some extent, *output* measures, although this first one is somewhat indirect.

The overall average number of visits per registered borrower is 9.2. The library average isn't all that far away, at 12.0. The median is 9.1—*almost* identical to the average, which means this measure could follow normal distribution.

The first quartile is 5.9. The third quartile is 13.5. The standard deviation is an astonishingly high 98.0. But look what happens when we drop 24 libraries at the low end and 24 at the high end. The trimmed li-

brary average drops to 10.7, even closer to the median; the standard deviation drops to a plausible 7.4, less than one-twelfth of the value including the most extreme cases.

The gap between median and overall average includes only 64 libraries (0.7%), serving 0.9% of borrowers with 0.8% of circulation. The gap between overall average and library average is much larger, 1,594 libraries (17.4%) serving 20.4% of borrowers with 25.7% of circulation. The central cluster includes 1,239 libraries (13.5%) serving 14.9% of borrowers with 15.6% of circulation.

Libraries with extremely few visits per patron include 68 (0.7%) with less than one (serving 0.8% of borrowers—but with only 0.06% of circulation) and, *including those 68*, a total of 259 (2.8%) with less than two visits per borrower. Those 259 libraries are also lightly used—serving 1.4% of borrowers, they only circulate 0.3% of circulation.

At the high end, 2,911 libraries (31.7%) average at least one visit per month per borrower; those libraries serve 19.7% of borrowers but provide 32.6% of all circulation. Narrowing that down further, 127 libraries (1.4%) average at least *three* visits per month (36 per year). Those libraries serve 0.3% of borrowers but do 0.8% of all circulation.

Removing the 2.8% at the bottom and 1.4% at the top, and rounding to whole visits, yields Figure 22 (noting that both endpoints are too low, since, for example, "2" doesn't include libraries with 1.51 to 1.99 visits per borrower).

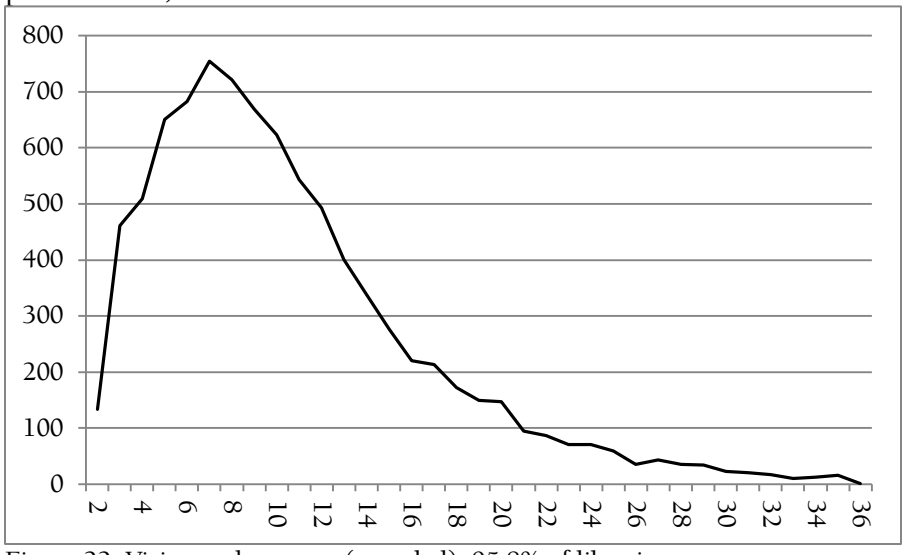

Figure 22. Visits per borrower (rounded), 95.8% of libraries

The peak here is 7—but most libraries are to the right of the peak.

Change from 2009 to 2010

Since visits were down (slightly) and registered borrowers were up (slightly), it follows that visits per registered borrower *have* to be down—at least at some libraries.

The overall average is down 2.1%. The library average is up 9.1%. The median is down 1.7%, so most libraries did see a drop in visits per borrower—but not a big drop. The first quartile is 89.7% (down 10.3%); the third is 109.5% (up 9.5%), very close to the library average).

The standard deviation is 93.4%. The trimmed library average is up 6.0%, while the trimmed standard deviation is 46.4%.

The gap between overall average and median includes 119 libraries (1.3%) serving 2.2% of borrowers with 1.9% of circulation. The gap between median and library average includes 2,246 libraries (24.5%) serving 25.6% of borrowers with 27.7% of circulation. The central cluster includes 4,660 libraries (50.8%) serving 58.5% of borrowers with 59.4% of circulation.

In all, 4,119 libraries (44.9%) had flat or increased visits per borrower. Those libraries account for 38.2% of borrowers and 42.3% of circulation.

At the *true* extremes, 49 libraries (0.5%) had less than one-quarter as many visits per borrower in 2010 as in 2009; those libraries served 0.3% of borrowers with 0.1% of circulation—and 70 libraries (0.8%) had at least *four times* as many visits per borrower; those libraries served less than 0.1% of borrowers with 0.1% of circulation.

For the purposes of Figure 23, we eliminate libraries falling below 50% or above 150%. At the low end, that means 187 libraries (2.0%) serving 1.0% of borrowers with 0.5% of circulation. At the high end are 625 libraries (6.8%) with more than a 50% increase in visits per capita; those libraries serve 2.8% of borrowers with 3.3% of circulation.

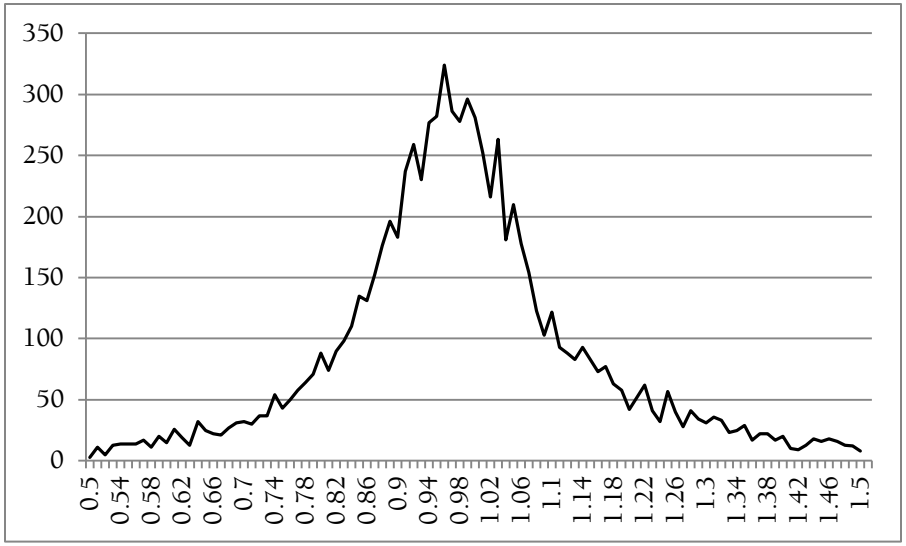

Figure 23. Change in visits per borrower, 2009 to 2010

The peak in Figure 23 is at 96% (0.96).

Reference Transactions per Borrower

For U.S. public libraries as a whole, there were an average of 1.8 reference transactions per borrower in 2010. In this case, the library average isn't much different: 1.7. The median, however, is *far* lower: 0.9.

The first quartile is 0.4. The third is 1.7: roughly halfway between the overall and library averages. This is a case where 73.5% of libraries are below average (using the lower of the two averages; 75.8% using the higher). Standard deviation is 17.0. Trimmed library average is 1.41, a little closer to the median; trimmed standard deviation is 1.8, a plausible value.

The gulf between median and library average includes 2,124 libraries (23.2%) that serve 29.4% of borrowers with 30.6% of circulation. The much smaller gap between library average and overall average includes 200 libraries (2.2%) serving 3.1% of borrowers with 4.1% of circulation. The central cluster, within 10% in either direction of the median, is quite small: 731 libraries (8.0%) serving 10.8% of borrowers with 9.0% of circulation.

At the low end, 173 libraries (1.9%) reported less than one reference transaction for every 50 borrowers—but those libraries serve only 0.3% of borrowers and do 0.1% of circulation. Perhaps more significantly, 712 libraries (7.8%) reported less than 0.10 transactions per borrower; those libraries serve 2.3% of borrowers with 0.9% of circulation.

At the highest end, 90 libraries (1.0%) reported more than 10.00 reference transactions per borrower; those libraries served 0.5% of borrowers with 0.9% of circulation. More broadly, 616 libraries (6.7%) reported more than 4.00 reference transactions per borrower; those libraries

served 8.8% of borrowers with 11.8% of circulation, so they're slightly larger than average.

Removing 14.5% of libraries at the top and bottom and rounding to the nearest 0.1 transactions yields Figure 24.

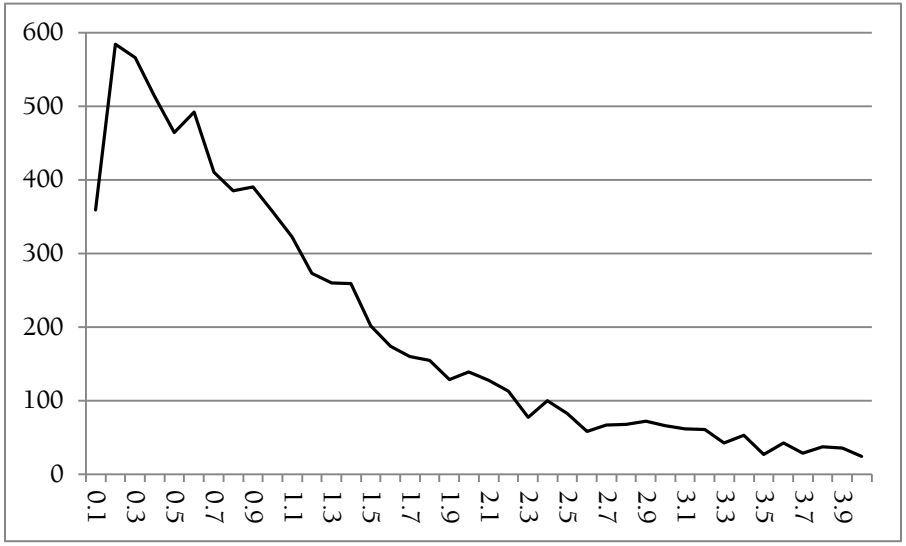

Figure 24. Reference transactions per borrower

That's a fairly startling curve, perhaps more so when you realize how many libraries are missing *on either side*—more on the left than on the right. The peak is 0.2, but almost 90% of the libraries represented in the curve are to the right of that peak.

Changes from 2009 to 2010

Including all libraries, reference transactions per borrower declined a small 1.24% from 2009 to 2010 (from 1.83 per borrower to 1.80 per borrower)—but, as with several other changes, it's not feasible to include libraries reporting no transactions in one or both years. That reduces the universe to 9,066 libraries. Among those libraries, the library average is a *gain* of 50.5%—but the median is a loss, albeit a small one (2.1%: the 2010 figure is 97.9% of the 2009 figure).

The first quartile is 83.8% (a loss of 16.2%); the third is 116.2% (a gain of 16.2%). Standard deviation for the 9,066 libraries is an absurd 1293.3%. Trimmed library average (removing 22 libraries at the top and at the bottom) is a gain of 18.6%; trimmed standard deviation, a still high but less absurd 139.2%. The gap between median and overall average is quite small, including 188 libraries (2.1%) that support 2.4% of borrowers with 2.5% of circulation. The gulf between overall average and library average includes 3,284 libraries (36.2%) that support 40.7% of borrowers and circulation. The central cluster includes quite a few libraries: 3,238

(35.7%). Those libraries support 44.2% of borrowers with 44.1% of circulation.

At the extremes, 75 libraries (0.8%) that had *some* reference transactions in both years had a drop of more than 90% in transactions per borrower; those libraries support 0.3% of borrowers with 0.1% of circulation. Another 275 libraries (3.0%) reported more than three times as many reference transactions per borrower; those libraries support 1.3% of borrowers with 1.1% of circulation.

In all, 4,131 (45.6%) libraries saw no loss or some gain in reference transactions per borrower. Those libraries served 44.8% of borrowers with 45.2% of circulation.

Truncating at 49.99% (a drop of more than half) and 200.01% (more than doubling) removes 579 libraries (6.3%) on the low side, serving 2.9% of borrowers with 2.2% of circulation, and 571 libraries (6.3%) on the high side, serving 3.1% of borrowers with 2.9% of circulation. The remaining libraries yield Figure 25. The peak in Figure 25 is at 0.97 (97%, a three percent drop). Note that this graph isn't directly comparable to some of the other change graphs, since cutting it off at 1.5 (+50%) would have removed too many of the values.

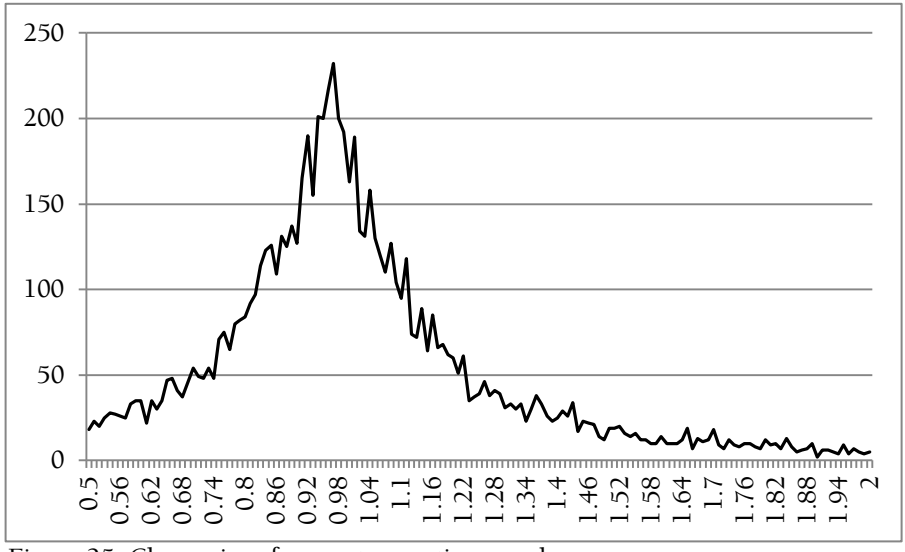

Figure 25. Change in reference transactions per borrower

Circulation per Borrower

The overall average for this key figure was 14.4 for 2010. That's surprisingly close to the library average at 15.0. The median is lower, at 12.5, but that's still an average of more than one circulation per month per borrower.

The first quartile is 7.7; the third quartile, 19.3. Standard deviation is 12.9. The trimmed library average is 14.7; the trimmed standard deviation, 10.0.

The gap between median and overall average includes 787 libraries (8.6%) that support 9.5% of borrowers with 9.0% of circulation. The gap between overall average and library average includes 223 libraries (2.4%) that support 2.7% of borrowers with 2.7% of circulation. The central cluster includes 1,147 libraries (12.5%) supporting 12.1% of borrowers with 10.5% of circulation. That's not a big chunk; in fact, as the quartiles indicate, circulation per borrower covers a broad range.

At the low end, 71 (0.8%) libraries circulate less than one item per borrower—and while those libraries support 0.7% of borrowers, they account for only 0.03% of circulation. Moving up one notch (and *including these libraries*), 235 libraries (2.6%) circulate fewer than two items per borrower; those libraries support 1.6% of borrowers but only 0.1% of circulation.

The true extreme for high activity is 15 libraries (0.2%) with more than one hundred circulation per borrower. Those 15 are (with one exception) *very* small, supporting 0.002% of borrowers with 0.03% of circulation. (Let's spell those out: The 15 libraries have a total of 3,984 borrowers and circulated 792,888 items.) Moving down one notch, 115 libraries (1.3%) circulated more than 50 items per borrower. Those libraries support 0.6% of borrowers and 2.5% of circulation. (One of them is a large library with astonishingly high circulation; another is a medium-sized library with similarly high circulation.)

Leaving out the libraries at either extreme yields Figure 26.

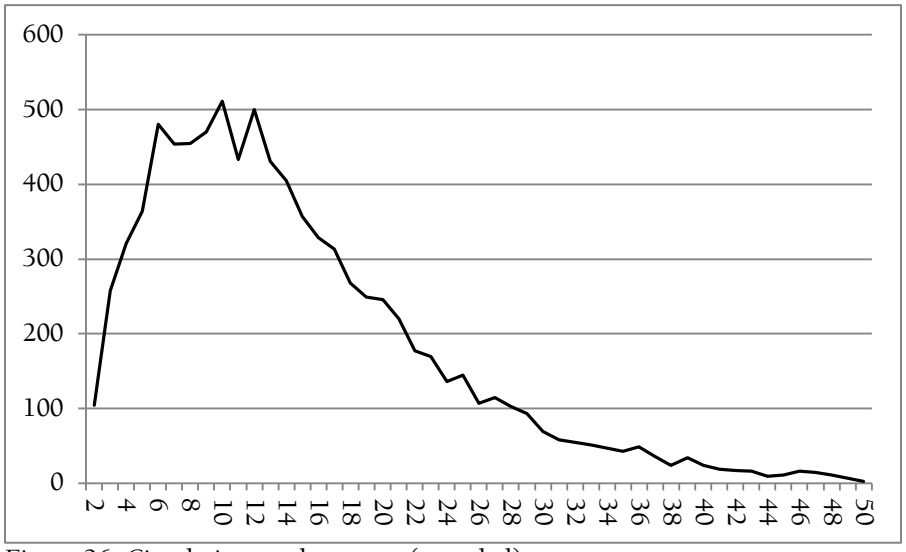

Figure 26. Circulation per borrower (rounded)

The two tallest peaks are at 10 and 12—and this graph slopes broadly to the right for quite a ways.

Changes from 2009 to 2010

We know that circulation rose overall—and so did circulation per borrower, although not by much. Overall, circulation per borrower was 101.1%, up 1.1% from 2009. The library average is *much* higher, 108.3% (up 8.3%).

Unfortunately, the median is 98.8%: down 1.2%. In other words, more than half the libraries in the U.S. had slightly *lower* circulation per borrower in 2010 than in 2009, although they had higher overall circulation.

The first quartile is 89.9% (down 10.1%); the third is 110.4% (up 10.4%). Pretty clearly, the combination of registering more borrowers (or clearing out old registrations) and circulating more items (although, remember, 44% of libraries saw an overall drop) yields a wide range of results—more than half of libraries either dropped at least 10% or gained at least 10%.

The standard deviation is 94.2%. The trimmed library average is 105.3%, that is, a gain of 5.3%; the trimmed standard deviation is 40.8%.

The gap between median and overall average includes 644 libraries (7.0%) supporting 8.2% of borrowers with 9.1% of circulation. The larger gap between overall average and library average includes 1,358 libraries (14.8%) supporting 20.6% of borrowers with 22.9% of circulation—somewhat larger libraries. The central cluster includes 4,557 libraries (49.7%) supporting 62.0% of borrowers with 63.7% of circulation: Almost half the libraries and not much less than two-third of borrowers and circulation. In all, 4,232 libraries (46.1%) had at least as much circulation per borrower in 2010 as in 2009.

As always, there are truly extreme cases: ten libraries (0.1%) dropping to less than one-tenth as much circulation per borrower (those libraries serving 0.1% of borrowers but with only 0.06% of circulation) and nine libraries (0.1%) with at least ten times as much circulation per borrower in 2010 as in 2009 (those libraries are *tiny* but very active: taken together they serve 0.001% of borrowers—2,038 in all—with 0.007% of circulation—164,280).

Less extreme outlying cases at the low end include 149 libraries (1.6%) where circulation per borrower dropped by more than half; those libraries are still mostly small, serving 0.5% of borrowers with 0.2% of circulation. At the high end, 223 libraries (2.4%) more than doubled circulation per borrower; those libraries are also small, serving 0.5% of borrowers with 0.6% of circulation.

Removing those larger groups of outlying libraries yields Figure 27.

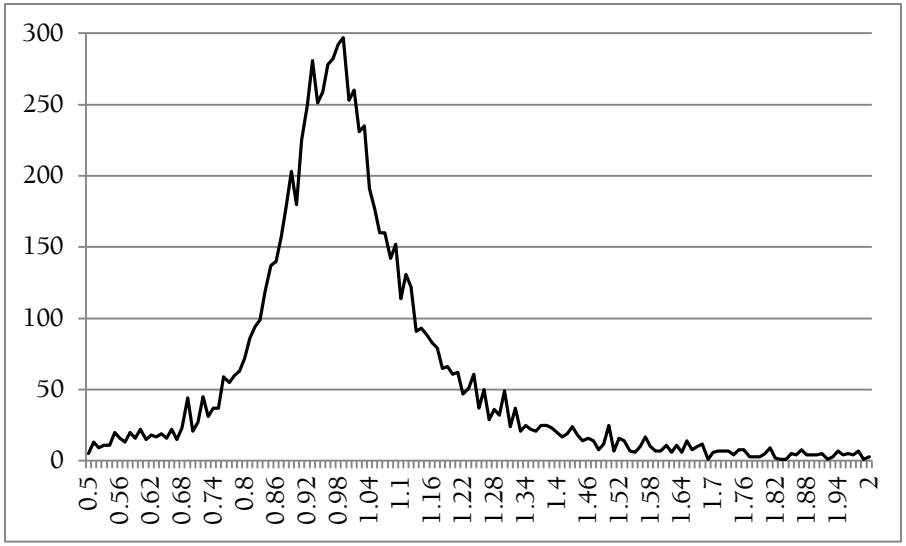

Figure 27. Change in circulation per borrower for 96% of libraries

The tallest peak in Figure 27 is at 0.99 (99%).

Program Attendance per Borrower

This measure is a *little* more encouraging than attendance per patron: At least it's more than 0.5 (although barely—the overall average is 0.51, or 51 total attendance per hundred borrowers). The library average is *much* higher at 0.86—but the median, although higher than the overall average at 0.57, is still pretty low.

The first quartile is 0.29; the third is 1.01, which is to say that more than one-quarter of U.S. libraries managed to average at least one attendance per borrower.

Standard deviation is 1.24. The trimmed library average is 0.82; the trimmed standard deviation is 0.91, an entirely plausible figure.

The gap between overall average and median (this is one of the rare cases where the overall average is *lower* than the median: most libraries are above average) includes 371 libraries (4.0%) supporting 4.1% of borrowers with 4.4% of circulation. The gap between median and library average includes 1,789 libraries (19.5%) supporting 16.2% of borrowers with 21.0% of circulation. The central cluster, within 10% on either side of the median, is relatively small: 948 libraries (10.3%) supporting 11.7% of borrowers with 13.9% of circulation.

Some 2,375 libraries (25.9%) managed at least one attendance per borrower. Those libraries support 8.9% of borrowers with 13.3% of circulation—they're generally smaller and more heavily used.

Since I've previously noted the number of libraries with *no* programs (and, thus, no attendance per borrower), here's a broader low extreme:

155 libraries (1.7%) with less than one attendance for every hundred borrowers—actually less than one for every 199, since it includes those that *round* to less than 0.01. Those libraries support 0.1% of borrowers with 0.07% of circulation—they're mostly very small libraries.

At the positive end, there are a few libraries (including one with more than 25,000 borrowers) that manage more than *ten* attendances per borrower—but the broader outlying group is those with at least five: 102 libraries (1.1%) with 0.1% of borrowers and circulation. Removing outliers on both ends and rounding to the nearest tenth of an attendance yields Figure 28.

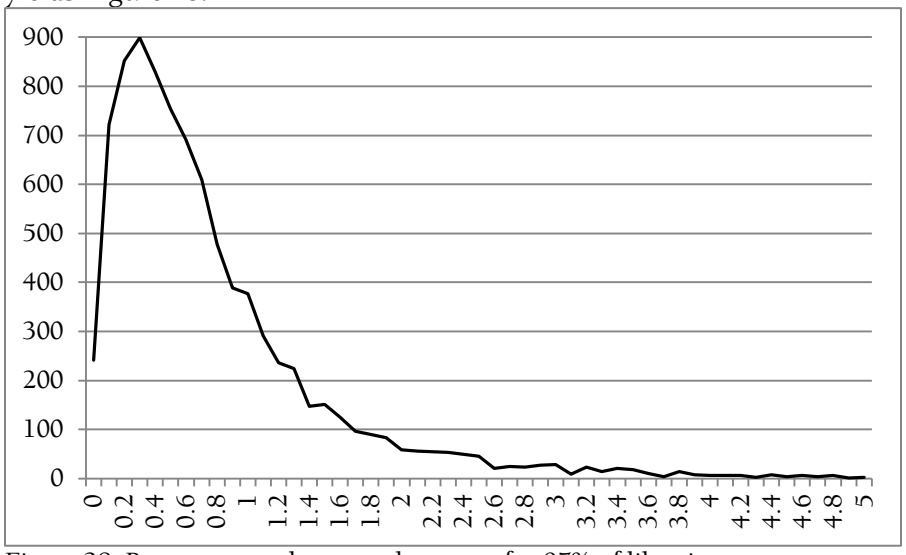

Figure 28. Program attendance per borrower for 97% of libraries

The peak is at 0.3. Note that including the low-lying extreme cases would raise the "0" mark to 396 instead of the current 242.

Changes from 2009 to 2010

Overall, program attendance per borrower is down just slightly from 2009 to 2010: 99.5%, down 0.5%. For all other figures, we ignore libraries reporting no attendance at all in either year (which would raise the overall average slightly, since it reduces the number of borrowers). The library average for the remaining 8,973 libraries is remarkably *high*, up 23.6% at 123.6%. The median is also up, albeit just barely: 100.1%, up one-tenth of one percent. Still, that's encouraging: program attendance per borrower *was* up for more than half of libraries with programs.

The first quartile is 82.9% (down 17.1%); the third, 122.3% (up 22.3%). Since that's lower than the library average, that figure is clearly thrown off by a number of very high percentage increases. Standard deviation is 220.5%. Trimmed library average (ignoring the 22 lowest and 22

highest percentages from the remaining 8,973 libraries) is 116.2%, up 16.2%; trimmed standard deviation is 91.7%.

The gap from the overall average to the median—a tricky one, as the former includes no-program libraries—includes 111 libraries (1.2%) serving 1.2% of borrowers with 1.2% of circulation. The larger gap between the median and the library average includes 2,328 libraries (25.9%) supporting 28.5% of borrowers with 30.7% of circulation. (In all, 4,471 libraries—49.9%—had the same or increased attendance per borrower. Those libraries account for 43.8% of borrowers and 45.9% of circulation, so on the whole they're *slightly* smaller.)

The central cluster ranges from 90.1% (a drop of 9.98%) to 110.1% (an increase of 10.1%) and includes 2,681 libraries (29.9%) supporting 35.1% of borrowers and 39.0% of circulation: Slightly *larger* on the whole.

On the low side, *not* including libraries with no attendance at all in one of the years, 452 libraries (5.0%) dropped by more than half in program attendance per borrower. Those libraries support 2.2% of borrowers with 1.3% of circulation—they're smaller and more lightly used.

On the high side, 552 libraries (6.2%) more than doubled program attendance per borrower. Those libraries support 1.8% of borrowers with 1.6% of circulation: They're generally small and *slightly* less actively used.

Removing those libraries and rounding to the nearest percentage yields Figure 29, as usual using decimal numbers rather than percentage points. Figure 29 is unusual among these graphs because the highest peak is "right in the middle"—at 100%, neither growth nor shrinkage.

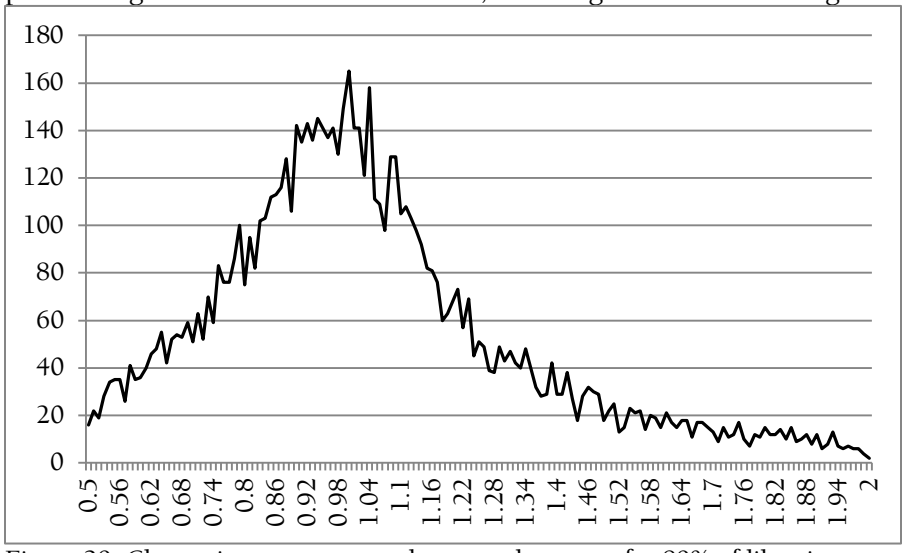

Figure 29. Change in program attendance per borrower for 89% of libraries

Personal Computer Sessions per Borrower

Overall, there were on average 2.2 logged PC sessions per borrower (where "PC" means library-provided computer with internet access available for personal use). The library average was 2.7. The median was 2.0, not much lower than the overall average.

The first quartile was 1.2; the third, 3.2. Standard deviation is 3.5. The trimmed library average is 2.6 and the trimmed standard deviation is 2.4, a realistic figure.

The gap between median and overall average includes 448 libraries (4.9%) supporting 8.4% of borrowers and 7.9% of circulation. The gap between overall average and library average includes 1,064 libraries (11.6%) supporting 11.6% of borrowers with 14.0% of circulation. The central cluster includes 1,070 libraries (11.7%) supporting 16.5% of borrowers with 17.5% of circulation.

At the bottom end, 158 libraries (1.7%)—some of them without public computers—report less than 0.1 sessions per borrower. Those libraries support 0.7% of borrowers with 0.1% of circulation: They're small and poorly used. At the top end, 189 libraries (2.1%) had more than ten sessions per borrower. Those libraries are also small (0.8% of borrowers) but better used (1.5% of circulation).

Removing the bottom and top ends and rounding to the nearest quarter-use yields Figure 30 (the peak is at 1).

This may be a good place to repeat that sessions per borrower includes only logged sessions on library-provided personal computers with internet access. It ignores wifi use and uses that aren't logged sessions.

Figure 30. PC sessions per borrower for 96% of libraries

Changes from 2009 to 2010

PC sessions per borrower declined ever so slightly for all libraries, including 101 with no sessions reported in one of the two years. The overall average is down from 2.16 sessions per borrower to 2.15, a decline of 0.4%. The rest of these figures will deal with the subset of 9,073 libraries that have at least one reported session in each year. (The missing libraries account for some 200,000 borrowers and less than two million circulation.)

The library average is an astonishing 281.5%— almost *tripling* the number of sessions per borrower. The median is starkly different: 99.1%, a decline of 0.9%—very small, but even a little worse than the overall average. The standard deviation is more ludicrous than the library average: 13,036.1%. Both library average and standard deviation are severely affected by a handful of extreme cases: The trimmed library average is 114.54, up 14.5%; the trimmed standard deviation is 104.0%.

The first quartile is 85.6%, a drop of 14.4% (roughly one-seventh); the third is 115.8%, an increase of 15.8%.

The gap between median and overall average is tiny, containing only 94 libraries (1.0%) that support 0.9% of borrowers and 0.7% of circulation. The gulf between overall average and library average is absurdly large, containing 4,212 libraries (46.4%) that support 43.0% of borrowers with 45.5% of circulation. The central cluster includes 3,330 libraries (36.7%) that support 41.6% of borrowers with 41.3% of circulation.

Some 4,367 libraries (48.1%) gained in PC sessions per borrower; those libraries support 42.8% of borrowers with 45.5% of circulation.

At the low end, 323 (3.6%) libraries saw more than a 50% drop in PC sessions per borrower; those libraries support 2.7% of borrowers and 2.6% of circulation. At the high end, 470 (5.2%) libraries more than doubled sessions per borrower; those libraries support 3.0% of borrowers with 3.5% of circulation. (The *extreme* high end is absurd: one library shows 2010 PC sessions per borrower at 12,369 *times* as high as in 2009.) Removing those libraries and rounding to whole percentages yields Figure 31 (again showing decimals rather than percentages).

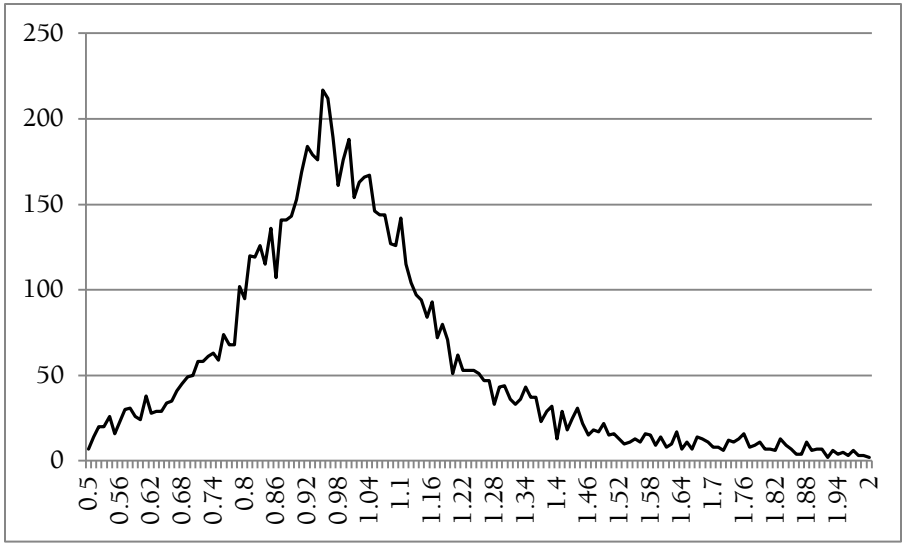

Figure 31. Change in PC sessions per borrower, 2009 to 2010, 91% of libraries

A Note on the Charts

That's it for the distribution charts. Although quite a few of them vaguely resemble bell-shaped curves, almost none are true normal distributions, since the median is almost never the same as—and frequently not even close to—the average (either average!).

In many cases, the average is wildly misleading, and even the median doesn't tell you that much about where libraries can be expected to lie.

I believe that the trio of quartile figures (first quartile, median, third quartile) offers the most realistic quick information on the *majority* of libraries—but it's possible for both the top quarter of libraries (those higher than the third quartile) and the bottom quarter to include a fair number of extreme cases, sometimes too many or too large to simply write off as outliers or reporting errors.

Now let's look at two other issues: Select possible correlations (omitting those discussed in *Graphic Public Library Benefits*) and whether changes and per-borrower metrics vary widely depending on the general size of libraries.

Correlations

Graphing Public Library Benefits shows correlations between spending per capita and other per-capita measures for (almost) all public libraries and for libraries within size groups. That PDF-only ebook costs $4; I refer you to it for those figures.

This brief runthrough looks at some other correlations that might be meaningful, ignoring a fair number that fall into the "yeah, so what?" cate-

gory (e.g., there's likely to be a strong correlation between patron visits and circulation: so what?). I'll only list cases where correlation or Pearson's coefficient is at least 0.30 (the low end of moderate correlation).

Changes in Whole-Number Measures

Most changes in whole-number measures (e.g., change in reference transactions) don't correlate well with changes in spending or, where appropriate, changes in collection spending. But there is one exception: Changes in spending on collections correlate *very* strongly (0.80) with changes in circulation.

Per-Borrower Measures

Here are cases where I found a significant correlation between measures that aren't obviously related.

- Visits per borrower correlate *very strongly* (0.71) with spending per borrower.
- Excluding libraries that report no reference transactions at all (102 of them), reference transactions per borrower correlate *very strongly* (0.71) with spending per borrower.
- Circulation per borrower correlates moderately (0.43) with spending per borrower—as does (0.46) circulation per borrower with collection spending per borrower.
- For those libraries that reported program attendance, program attendance per borrower correlates moderately (0.35) with spending per borrower.
- For those libraries reporting PC sessions, PC sessions per borrower correlate strongly (0.53) with spending per borrower.
- Number of books per borrower correlates moderately (0.37) with circulation per borrower.

Changes in per-borrower figures

There are a few significant correlations:

- Changes in spending per borrower correlate *very strongly* (0.72) with changes in visits per borrower.
- Changes in spending per borrower also correlate *very strongly* (0.72) with changes in circulation per borrower.
- For libraries that spent money on collections in both years, changes in spending on collections also correlate *very strongly* (0.76) with changes in circulation per borrower.
- For libraries reporting programs in both years, changes in spending per borrower correlate moderately (0.34) with changes in program attendance per borrower.

- Finally, changes in books per borrower correlate *very strongly* with changes in circulation per borrower.

Small, Medium, Large

Do distributions of derived library measures vary widely depending on the size of libraries? It might be interesting to look at a four-way split, one I'll call very small, small, medium and large, based on registered borrowers in 2010 (not LSA).

- Very small includes 2,637 libraries with 14 to 1,499 registered borrowers.
- Small includes 2,461 libraries with 1,500 to 4,999 registered borrowers.
- Medium includes 2,565 libraries with 5,000 to 19,999 registered borrowers.
- Large includes 1,511 libraries with at 20,000 to 3,305,875 registered borrowers.

For each size category, I show three tables with the first quartile (Q1), median (Med) and third quartile (Q3) for the measure. Tables showing year-to-year changes also show the percentage of libraries that didn't shrink for that particular measure (1+). Where the figure is at least 20% higher than the corresponding figure for all libraries, it appears in bold. Where it's at least 20% lower, it appears in *italics*. That's 20% *of the measure*—so, for example, if a measure is 12.0%, 20% less is 9.6% and 20% more is 14.4%. To make the figures readable, measure names are abbreviated somewhat.

Very small libraries (14-1,499 borrowers)

Changes	Q1	Med	Q3	1+%
LSA	*-0.5%*	0.0%	*0.5%*	69.7%
Spending	-5.3%	2.1%	11.1%	59.2%
Print Books	-1.9%	1.3%	4.7%	65.0%
Visits	*-6.4%*	1.1%	9.9%	56.8%
Reference	-13.3%	*0.0%*	17.7%	53.4%
Borrowers	*-4.7%*	2.8%	10.0%	63.0%
Circulation	*-8.3%*	*1.0%*	11.7%	54.6%
Programs	-4.6%	*0.0%*	24.7%	61.0%
Prog. Att.	*-18.7%*	*1.4%*	26.9%	52.9%
PC Sess.	*-11.8%*	2.0%	18.0%	56.8%

Table1. Changes in measures, very small libraries

Per Borrower	Q1	Med	Q3
Spending	$37.56	$57.42	$88.18
Coll. Spend	$5.04	$8.71	$13.82
Print Books	12.42	18.17	27.21
Visits	6.05	9.54	14.57
Reference	*0.27*	0.80	1.77
Circulation	7.53	12.58	20.21
Prog. Attend	0.32	0.73	1.44
PC Sessions	1.29	2.48	4.40

Table 2. Per-borrower measures, very small libraries

Change PB	Q1	Med	Q3	1+%
Spending	-11.6%	0.1%	18.2%	50.2%
Coll. Spend	-19.3%	-1.9%	22.7%	47.4%
Print Books	-9.0%	-1.1%	8.0%	46.1%
Visits	-11.7%	0.7%	14.1%	48.2%
Reference	-18.2%	-1.8%	26.4%	47.1%
Circulation	-11.8%	-0.6%	15.8%	48.7%
Prog. Att.	*-23.4%*	0.4%	35.1%	50.6%
PC Sess.	-16.6%	0.1%	25.2%	50.4%

Table 3. Changes in per-borrower measures, very small libraries

Small libraries (1,500-4,999 borrowers)

Changes	Q1	Med	Q3	1+%
LSA	0.0%	0.0%	*0.5%*	77.0%
Spending	-3.7%	1.7%	7.1%	59.5%
Print Books	-2.1%	1.4%	4.4%	63.3%
Visits	-4.4%	1.1%	8.2%	59.3%
Reference	-11.8%	0.7%	13.5%	54.4%
Borrowers	-2.3%	3.8%	8.8%	68.1%
Circulation	-5.2%	1.4%	8.9%	56.4%
Programs	-11.2%	2.0%	21.4%	59.5%
Prog. Att.	-12.2%	3.1%	22.4%	56.9%
PC Sessions	-9.3%	2.0%	14.0%	57.0%

Table 4. Changes in measures, small libraries

Per Borrower	Q1	Med	Q3
Spending	$32.35	$50.11	$74.85
Coll. Spend	$3.94	$6.43	$10.17
Print Books	6.32	8.59	11.14
Visits	5.63	8.90	13.44
Reference	0.34	0.81	1.65
Circulation	7.83	12.44	18.77
Prog. Attend	0.30	0.61	1.08
PC Sessions	1.17	1.99	3.18

Table 5. Per-borrower measures, small libraries

Change PB	Q1	Med	Q3	1+%
Spending	-10.0%	-1.4%	9.7%	45.3%
Coll. Spend	-16.8%	-2.5%	13.5%	44.8%
Print Books	-8.4%	-2.0%	5.1%	41.6%
Visits	-10.0%	-1.7%	10.1%	44.5%
Reference	-16.6%	-2.0%	15.7%	46.1%
Circulation	-10.3%	-1.3%	10.8%	45.8%
Prog. Att,	-16.6%	0.5%	24.1%	50.8%
PC Sess.	-4.3%	-0.4%	15.9%	49.2%

Table 6. Changes in per-borrower measures, small libraries

Medium libraries (5,000-19,999 borrowers)

Changes	Q1	Med	Q3	1+%
LSA	0.0%	0.0%	0.6%	80.7%
Spending	-3.9%	0.9%	6.2%	56.1%
Print Books	-1.9%	1.2%	3.9%	61.9%
Visits	-4.6%	1.1%	6.8%	57.6%
Reference	-10.5%	0.6%	11.4%	53.1%
Borrowers	-2.2%	3.3%	8.1%	65.4%
Circulation	-4.6%	1.1%	7.3%	56.2%
Programs	-9.6%	2.9%	19.5%	57.1%
Prog. Att.	-10.4%	2.5%	17.7%	55.3%
PC Sess.	-8.2%	1.7%	12.2%	56.2%

Table 7. Changes in measures, medium libraries

Per Borrower	Q1	Med	Q3
Spending	$32.56	$52.77	$83.19
Coll. Spend	$3.60	$6.13	$9.92
Print Books	4.20	*5.08*	7.77
Visits	6.04	8.86	13.92
Reference	0.40	0.86	1.68
Circulation	7.79	12.58	19.25
Prog. Att,	0.32	0.57	0.95
PC Sess,	1.19	1.84	2.75

Table 8. Per-borrower measures, medium libraries

Change PB	Q1	Med	Q3	1+%
Spending	-8.9%	-1.6%	*7.0%*	44.3%
Coll. Spend	-15.7%	-3.6%	*9.3%*	41.2%
Print Books	-7.6%	-1.5%	4.7%	43.2%
Visits	-9.2%	-1.6%	8.1%	44.8%
Reference	-14.8%	-2.0%	14.0%	45.3%
Circulation	-9.3%	1.0%	9.1%	46.2%
Prog. Att.	-14.3%	0.6%	19.8%	51.5%
PC Sess.	-12.3%	-0.6%	14.1%	48.6%

Table 9. Changes in per-borrower measures, medium libs

Large libraries (20,000-3.3 million borrowers)

Changes	Q1	Med	Q3	1+%
LSA	0.0%	0.0%	1.3%	86.8%
Spending	*-6.0%*	*-0.6%*	*3.8%*	45.8%
Print Books	*-2.8%*	0.5%	*3.2%*	55.1%
Visits	-5.4%	*0.6%*	*5.7%*	53.3%
Reference	-9.9%	*0.0%*	11.7%	50.5%
Borrowers	-1.1%	4.2%	8.4%	71.5%
Circulation	-3.8%	1.5%	*6.7%*	58.7%
Programs	-11.0%	1.8%	*15.3%*	54.6%
Prog. Att.	-10.7%	*1.4%*	*15.2%*	53.3%
PC Sess.	-9.3%	*0.6%*	*10.5%*	52.2%

Table 10. Changes in measures, large libraries

Per Borrower	Q1	Med	Q3
Spending	$33.95	$51.18	$78.23
Coll. Spend	$3.52	$5.90	$9.50
Print Books	*2.91*	*3.96*	*5.46*
Visits	5.84	8.12	11.36
Reference	0.63	1.12	2.07
Circulation	7.62	12.23	18.63
Prog. Att.	0.26	*0.42*	*0.67*
PC Sess.	1.18	1.74	*2.54*

Table 11. Per-borrower measures, large libraries

Change PB	Q1	Med	Q3	1+%
Spending	-11.1%	*-4.3%*	*3.4%*	*34.3%*
Coll. Spend	-17.8%	*-6.4%*	*5.4%*	34.5%
Print Books	-8.6%	*-3.2%*	*2.5%*	*33.5%*
Visits	-10.7%	*-2.9%*	*5.4%*	40.0%
Reference	-15.2%	*-3.1%*	*11.5%*	42.6%
Circulation	-8.7%	*-1.9%*	*5.9%*	42.0%
Prog. Att.	-5.2%	*-1.7%*	14.1%	47.1%
PC Sess.	-14.1%	*-3.4%*	*10.5%*	41.6%

Table 12. Changes in per-borrower measures, large libs

Conclusions

If nothing else, this study should make you wary of statistical results from small or even fairly large samples if those results are presented as being true for libraries as a whole.

Pay What You Wish

Cites & Insights carries no advertising and has no sponsorship. It does have costs, both direct and indirect. If you find it valuable or interesting, you are invited to contribute toward its ongoing operation. The Paypal donation button (for which you can use Paypal or a credit card) is on the *Cites & Insights* home page. Thanks.

Masthead

Cites & Insights: Crawford at Large, Volume 13, Number 5, Whole # 161, ISSN 1534-0937, a journal of libraries, policy, technology and media, is written and produced by Walt Crawford.

URL: citesandinsights.info/civ13i5on.pdf