PERSPECTIVES

PHOENIX CENTER FOR ADVANCED LEGAL & ECONOMIC PUBLIC POLICY STUDIES

Measuring Broadband at the FCC: A Peculiar Approach

George S. Ford, PhD*

February 19, 2026

Introduction

The Federal Communications Commission's *Section 706 Reports* are a useful source for information on broadband deployment, even if largely used by Democratic administrations to justify regulation. Armed now with the new broadband fabric data, which attempt to measure broadband availability down to the location level, the FCC's staff is faced with some new complexities. Summary statistics provided in the *Report* require linking the finely-tuned broadband data to population statistics. The problem is that the fabric collects data for Broadband Serviceable Locations while the Census Bureau only collects data down to the census block level, and it does so only every ten years. Population, therefore, must be allocated to the Broadband Service Location by some mechanism to produce population-averaged statistics.

This PERSPECTIVE discusses the statistical procedure used by the FCC—summarized in its *2024 Section 706 Report* (the most recent *Section 706 Report* available)—to do so.[1] While the FCC's approach seemed interesting at first glance, it turns out to be a bad idea as it is used for the wrong purpose. The permutation approach outlined in the *Report* may be suitable for constructing confidence intervals, but not for allocating population (except in repeated draws). Here, I demonstrate the problems with the Commission's approach.

The problem is that the fabric collects data for Broadband Serviceable Locations while the Census Bureau only collects data down to the census block level, and it does so only every ten years. Population, therefore, must be allocated to the Broadband Service Location by some mechanism to produce population-averaged statistics.

Statistical Procedure in the Section 706 Report

Section 706(c) requires the Commission to:

> compile a list of geographical areas that are not served by any provider of advanced telecommunications capability (as defined by subsection (d)(1)) and to the extent that data from the Census Bureau is available, determine, for each such unserved area—
>
> (1) the population;
>
> (2) the population density; and
>
> (3) the average per capita income.[2]

I have looked at many *Section 706 Reports* and have never seen such a list, though the *Reports* always contain statistics related to the number of people without access to broadband, so somewhere the list exists.

Computing such statistics requires the FCC to assign population to geographic areas. Prior to the broadband fabric, this task was somewhat straightforward as the Form 477 broadband data were collected at the census block level. The Census Bureau reports census block level population estimates (from the decennial census), requiring only a summing up of the population in unserved blocks as determined by FCC deployment data.

As is well documented, the Form 477 data overstate availability given the assumption that all locations within a block are served if any single location within the block was served. My work, based on the fabric data, suggests the overstatement was about 3.3 percentage points on average, though the bias can be sizable in rural areas where blocks are geographically large.[3]

The new fabric data provide location data below the block level. A Broadband Serviceable Location ("BSL") is "a business or residential location in the United States at which mass-market fixed broadband Internet access service is, or can be, installed."[4] A BSL may have one or multiple units—a single-family home versus an apartment complex.

Using the fabric data, it is easy enough to count the share of *locations* served and unserved, and that should be the target statistic reported by the FCC—it is the most precise available in that data and broadband gaps are best measured by location. Yet the statute requires an assessment of "population," at least to the extent Census Bureau data permits it. If the FCC wants to exploit the fabric's detail, then it must impute population data down to the BSL, and that requires addressing the variation in the number of housing/business units contained within a BSL. A single-family home in a BSL might contain, on average, 3 persons, but a twenty-unit apartment building, also a single BSL, would service many more people, and there is no reason to suspect the number of persons per unit equals the mean household size. Mean household size in the U.S. is about 2.54 persons.[5]

Allocating population from a higher to a lower level of aggregation is tricky, especially without any external information providing some guidance. We must live with estimates. One approach is to allocate a census block's population, which is observed, across BSLs based on the number of units in the BSL, or equivalently the share of units across BSLs in the block. This is the maximum likelihood estimate of population count per BSL. It doesn't get much better than that given what we know.

In the last *Section 706 Report*, however, the FCC describes an alternative procedure. An example is provided with a census block with a population of 20 people and 3 BSLs; the first BSL has one unit, the second 2 units, and the third 3 units. The maximum likelihood estimate (unit share) of the population count is 3.33 in BSL1 [= (1/6)20], 6.67 in BSL2, and 10 in BSL3. Or, if integers are your thing, then we have 3 in BSL1, 7 in BSL2, and 10 in BSL3.

This approach is not what the FCC is doing. Rather, the FCC's procedure is as follows:

> … each person is assigned to a BSL unit with probability 1/(# of units in block). For example, if a block has a population of 20 persons [] and six units contained within three BSLs [], each person is essentially assigned to a unit in turn by rolling a six-sided die (because there are six units). [] We then estimate the number of households by counting the number of units within populated BSLs.[6]

The first part sounds like the unit-share approach, but the "six-sided die" component points in a different direction. Specifically, it appears that FCC is making a single random draw from the binomial distribution based on total population and unit shares. A binomial distribution is a discrete probability distribution that models the number of successes in a fixed number of independent trials, each with the same probability of success.

To grasp this procedure, consider a bucket that contains six balls, one with a "1" on it, two with a "2" on them, and three with a "3" on them. The FCC reaches into the bucket and draws one ball and records its number, returns the ball, and then repeats the draw 19 more times. There are 20 draws in total, and the population in each BSL is based on the tally of what number appears on the ball across these draws. Notably, this approach will not provide the unit-share (maximum likelihood) allocation of population (except by coincidence in integer form). In fact, if the FCC repeated this process a second time (making another 20 draws), then the allocation populations would not be the same as the first set of draws (except by coincidence). There are 231 unique draws available in this very simple case.

Table 1 summarizes the results of three such simulations for this hypothetical census block. For each random draw, the population of 20 is allocated differently across BSLs—very differently in some cases.[7] We get 7, 6, or 15 people allocated to BSL3, and 4, 7 or 1 people allocated to BSL1. Which do we use? The FCC, it appears, uses the numbers from the first draw. This single draw might allocate the population across BSLs using any of the three outcomes in Table 1, or one of 228 other possible allocations for this simple case. (If there were 2,500 persons across 100 BSLs, the number of unique allocations is near infinite.)

Table 1. Random Draws

				Draw	
BSL	**Units**	p	**1**	**2**	**3**
1	1	0.167	4	7	1
2	2	0.333	9	7	4
3	3	0.500	7	6	15

This "single draw" approach might be justified if it was more accurate than the simpler unit share allocations? It is not. In fact, for any true distribution of the population across BSLs (which is unobserved, but can be simulated by Monte Carlo analysis), the unit-share allocation is a better estimate *on average* of the truth (any truth) than the single draw approach used by the FCC, and more accurate in two-thirds of cases.

Here's a thought experiment: Recognizing that the allocation will vary in each random draw, why not take many random draws and take the average? Doing so results in the unit-share allocation (3.33, 6.67, and 10, or the integers 3, 7, and 10). Yet, using only a single draw the allocation could be (0, 0, 20), (6, 7, 7), (12, 6, 2), (4, 7, 1), (9, 7, 4), (7, 6, 15), or any one of an additional 225 unique alternatives in this simple example. The simplest approach of all, and the most accurate absent additional information, is simply to multiply the block's population by the share of units with broadband service (avoiding any assignment of population to BSLs or units).

The FCC's random draw approach requires a random seed for reproduction, which permits random draws to be reproduced exactly.[8] In setting the random seed, the first draw and subsequent draws may be accurately reproduced. But, even with the same seed, a second draw will produce a different population allocation than the first draw. Why is the first, or any particular draw, so special? It's not. Or, if the seed is changed from 12345 to 12346, then the first draw will be different between the two seeds. Why is the population allocation changing simply because of a change in the seed? It should not.

Does the FCC's approach improve the estimates of relevant statistics summarized in the *Section 706 Report*? You probably guessed by now the answer is "no." Say, for example, we wanted to compute share of population with two or more broadband providers (a popular statistic from the reports). For any true competition level, the unit-share allocation is a better fit to the truth *on average* than a single draw. The root mean square error of the FCC's approach is 40% larger than the unit-share approach.

To the point, the complex (and computationally expensive) procedure to allocate population across BSLs described in the last *Section 706*

Report is not an improvement over a simple unit-share allocation. It performs worse, on average, in all respects. Certainly, that one draw may hit the nail on the head (though we can't know for sure), but that would be coincidental and a near zero probability event, especially across millions of census blocks.

And, it is entirely feasible for the FCC's approach to assign zero population to a BSL, another problem observed in the data for which the FCC proposes a fix. There is no reason to employ a complicated procedure when a simple procedure performs better; and one that can be easily reproduced across analysts and software platforms. Integers are not required and only complicate matters when BSL probabilities are small (leading to zeros).

The simplest procedure is the compute the share of units within a blocks that meet the required threshold (say, two or more providers). Then, multiply that proportion by block population. Absent more information (though some is provided below), this calculation is the maximum likelihood estimate.

There are uses for permutations of the sort proposed by the FCC. For instance, permutations may be used to construct a sample distribution for a statistic of interest with an unknown distribution. Reporting the mean of calculated statistics (say, persons with two or more providers) and its margin of error ("MOE") might be useful. The MOE may be available analytically (though perhaps requiring adjustments for "survey" design), but the permutation or bootstrap approach would provide an exact measure. It's worth investigating.

Allocation Refinements

There are many complications in this endeavor, some of which may be informed from the analysis of data. For example, multi-dwelling units may have fewer people on average than single-family homes, and population-per-unit may decline the larger is the number of units.

> ***There is no reason to employ a complicated procedure when a simple procedure performs better; and one that can be easily reproduced across analysts and software platforms.***

Such differences may be estimable. Let's take the latest fabric data (v6) and the FCC's 2023 population estimates. First, business locations are excluded and the fabric data is collapsed to the block level for a sum of total units (U_i) and the number of unique BSLs (B_i). The block level population estimates are P_i. I then run the regression,

$$\ln\left(\frac{P_i}{U_i}\right) = \beta_0 + \beta_1 \ln\left(\frac{U_i}{B_i}\right) + \mu_i + \varepsilon_i \tag{3}$$

where μ_i is a county-level fixed effect and ε_i is the disturbance term. The left-hand side is the natural log of "household size" and the continuous variable on the left-hand size is the log of units per BSL. The U_i/B_i is 1.0 when all the homes are single units and gets larger as the number of units per BSL gets larger. The log-log specification is used as the relationship is non-linear in levels. The regression is weighted by block population. Blocks with zero population or unit counts are excluded. Both coefficients (1.03 and -0.198) are statistically different from zero at the 1% level (using county clustered standard errors). The within-R^2 is 0.071.

Table 3. Population by Unit Size

Units	Pop/Unit	Lower Bound (95% CI)	Upper Bound (95% CI)	Weight
1	3.347	3.332	3.361	1.000
2	2.917	2.911	2.922	0.872
3	2.691	2.677	2.706	0.804
4	2.542	2.522	2.562	0.760
5	2.432	2.408	2.456	0.727
6	2.345	2.318	2.373	0.701
7	2.275	2.245	2.304	0.680
8	2.215	2.184	2.247	0.662
9	2.164	2.131	2.197	0.647
10	2.119	2.085	2.153	0.633
20	1.847	1.806	1.888	0.552
30	1.704	1.660	1.748	0.509
40	1.610	1.564	1.655	0.481
50	1.540	1.493	1.587	0.460
100	1.342	1.293	1.391	0.401

From the estimates of Equation (3), I calculate the predictions of household size at varying unit sizes. Results are summarized in Table 3. All these values are statistically different from one another (but $N \approx 5.7$ million). As expected, we see that the average household size (population per unit) declines as the number of housing units per BSL rises. The population counts in one-unit dwellings is a little rich, however. ACS data indicates the average household size for one-unit dwellings in 2.74 people, for two-to-four unit dwellings in 2.29 people, and for five-or-more dwellings in 1.90 people.[9] The pattern is similar, but not exact.

Of course, this sort of regression approach could be expanded or restricted as desired. Household sizes differ, for example, by race. If the parameters estimates are sufficiently similar, then the same adjustment could be applied irrespective of the statistical tests (the large sample size tends to make difference statistically significant). In any case, there is the possibility of using statistical models to improve population allocations. The benefits of such refinements are certainly subject to diminishing marginal returns, but some may prove useful, while some may not be worth the hassle.

Naturally, there will be errors in allocating population to BSLs—we simply have no idea what the true populations are. Some sort of MOE calculation may be prudent.

Conclusion

This PERSPECTIVE discusses the statistical method used to assign population to Broadband Serviceable Locations outlined in the *2024 Section 706 Report*. The FCC's current approach of using a single random draw to allocate population across Broadband Serviceable Locations, while perhaps well-intentioned, introduces unnecessary complexity and potential error compared to simpler unit-share allocation methods. The most obvious approach—a unit-share allocation—is the simplest and the best absent more information. A margin of error may also be computed to account for the estimation. Adjusting population allocations based on regression analysis may or may not prove useful.

NOTES:

[*] **Dr. George S. Ford** ***is the Chief Economist of the Phoenix Center for Advanced Legal and Economic Public Policy Studies. The views expressed in this* PERSPECTIVE *do not represent the views of the Phoenix Center or its staff. Dr. Ford may be contacted at*** ford@phoenix-center.org.

[1] *2024 Section 706 Report*, Federal Communications Commission, GN Docket No. 22-270 (rel. March 18, 2024) at Appendix A (available at: https://docs.fcc.gov/public/attachments/FCC-24-27A1.pdf).

[2] 47 U.S.C. § 1302(c).

[3] G.S. Ford, *Overstating Broadband Availability: An Assessment of the "All-In" Assumption for FCC 477 Data*, PHOENIX CENTER POLICY PERSPECTIVE No. 22-04 (September 14, 2022) (available at: https://www.phoenix-center.org/perspectives/Perspective22-04Final.pdf).

[4] *About the Fabric: What a Broadband Serviceable Location (BSL) Is and Is Not,* Federal Communications Commission (April 15, 2024) (available at: https://help.bdc.fcc.gov/hc/en-us/articles/16842264428059-About-the-Fabric-What-a-Broadband-Serviceable-Location-BSL-Is-and-Is-Not).

[5] Data available at: https://www.census.gov/quickfacts/fact/table/US/HCN010222.

[6] *2024 Section 706 Report*, *supra* n. 1 at Appendix A.

[7] Population counts are produced using a random draw from the binomial distribution with 20 trials at the BSL's unit share. A leave-one-out procedure is used where BSL3 receives the remainder. Which BSL receives the remainder affects the variance but the not the mean of the allocations. Another approach is to expand the sample to the size of the population by random draws based on the unit share and then count the occurrences of BSLs.

[8] A random seed is a number used to initialize a pseudorandom number generator. When you set a specific random seed, it ensures that the same sequence of random numbers will be generated each time you run your code.

[9] Data from the 2023 ACS Five-year data using B25124 and B25033.