PERSPECTIVES

PHOENIX CENTER FOR ADVANCED LEGAL & ECONOMIC PUBLIC POLICY STUDIES

Overstating Broadband Availability: An Assessment of the "All-In" Assumption for FCC 477 Data

Dr. George S. Ford*

September 14, 2022

Broadband availability is an important part of policymaking in the United States. Until recently, the only source for broadband availability data was the Form 477 data collected (since 2014) by the Federal Communications Commission ("FCC").[1] These data are collected by the Commission from broadband providers at the census block level (averaging about 20 homes). Under an "all-in" assumption, a census block is deemed to have broadband (at a specified speed threshold) if a provider serves (or could serve in a few months) a single location within the block. This assumption tends, of course, to overstate broadband availability. As a result, the Form 477 data have been heavily criticized for doing so and likened to "fake news."[2]

Importantly, there are at least two sorts of errors in the Form 477 approach of policy relevance. First, there is the issue of over-stating the total availability of broadband service at an aggregate level (e.g., state or nation). Second, there is the question of the quality of data to identify the exact location of unserved locations. The latter is more significant since it may affect where subsidy dollars are spent, while the former speaks mostly to a general distrust of the data in general policy debates.

In prior works, I have attempted to quantify the overstatement. Using a purely statistical approach, I found that the "all-in" assumption resulted in about a 3.3 percentage point overstatement of availability.[3] In a second study, I compared newly-released data from Georgia, collected at the location level (and summarized at the block level), to the Form 477 data and found that the FCC data overstated broadband by 3.5 percentage points in the state, which was well below that estimated by other researchers (7.3% for Georgia).[4] In my latter work, however, the comparison was between two data sets, and thus the overstatement reflected both disparities in the data (in part, a short temporal difference) and the effect of the "all-in" assumption.

When the Form 477 data are analyzed at the state level ... the overstatement is small—less than four percentage points. Criticisms of these high-level statistics, therefore, are somewhat overblown.

More states have now begun to collect and report detailed surveys of broadband availability, relying on location-specific surveys. Georgia was the first to release this new, more detailed data and Iowa now reports availability data down to the address in an easy-to-use format. Using these data, in this PERSPECTIVE I estimate the overstatement of the FCC 477's "all-in" assumption by comparing availability data at the block level using data from Georgia and Iowa. This analysis permits a direct assessment of the consequences of the "all-in" assumption, ignoring differences in the data between these

more detailed surveys and the Form 477 data. My approach is straightforward: actual availability rates in a census block are compared to the "all-in" availability rate constructed under the assumption that all locations have broadband in the census block if the availability rate exceeds zero.

[I]f one wants to know the exact locations without broadband service availability, then the consequences of the Form 477's "all-in" assumption can be severe. In blocks that cover many square miles, the overstatement of availability can be very large.

The effect of the Form 477's "all-in" assumption is found to be small—about 3.6 percentage points in Georgia and 2.5 percentage points in Iowa. The results are consistent with my prior work. High-level estimates of broadband availability are not much affected by the "all-in" assumption, but the assumption has more severe consequences, of course, if one wants to know exactly which locations do not have broadband service available. The Form 477's overstatement affects large census blocks (in square mileage) much more than small census blocks, and the overstatements are comparable between the two.

Data

Broadband availability data is obtained for Georgia and Iowa from their respective state broadband maps.[5] The Georgia data are for June 2022, and the Iowa data are for August 2022. The data are linked to Census Tiger Shapefiles for population and density data (2020 blocks for the Georgia data, 2010 blocks for the Iowa data).[6]

Analysis

For each of the Georgia and Iowa Broadband Maps, the number of locations with broadband at the at the 25/3 Mbps level, labeled *A*, are calculated. The "all-in" availability rate, Z, is calculated by assuming all locations within a census block have broadband if any location within the block has broadband. The "all-in" assumption's overstatement, *E*, is computed as *Z* – *A*, which is measured in percentage points ("pp"). Since all the data are from the same source, the effect of the "all-in" assumption is directly estimated, rather than comparing the state data to the FCC's Form 477 data.

Table 1. Overstatement from the "All-In" Assumption

State	*Z*	*A*	*E*
Georgia	94.9%	91.3%	3.6 pp
Iowa	92.8%	90.4%	2.4 pp
Both	94.5%	91.1%	3.4 pp

The results are summarized in Table 1. For Georgia, the measured availability rate, *A*, of broadband service is 91.3% of locations. Under the "all-in" assumption, the availability rate is calculated to be 94.9%. The error, therefore, is 3.6 percentage points. This difference is nearly identical to my earlier analysis comparing the Georgia data to the Form 477 data. For Iowa, the difference is smaller. The availability rate, *A*, of broadband service is 90.4% of locations, while the "all-in" assumption implies an availability rate of 92.8%, for an overstatement of 2.5 percentage points. Across both states, the overstatement is 3.4 percentage points, reflecting the larger population of Georgia.

Overstatement and Block Size

The overstatement of availability from the "all-in" assumption is presumably related to block size. Rural blocks can be quite large, so the "all-in" assumption is prone to overstate availability more in these areas. To quantify the relationship, the spread between the actual and "all-in" availability rates is calculated for each census

block. This variable is regressed on the natural log of square miles, but only when the "all-in" availability rate is 1.0. Since the dependent variable is restricted to the unit interval, the model is estimated using a Generalized Linear Model ("GLM") of the binomial family with a logit link, which ensures the predictions are on the unit interval.

Figure 1. Overstatement and Density

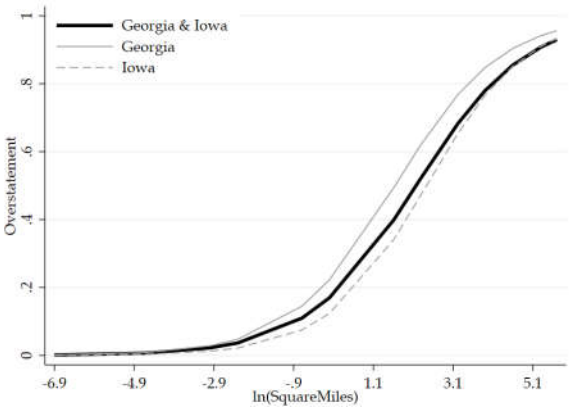

Figure 1 illustrates the mean overstatement across block size. Three relationships are presented: (1) both states; (2) Georgia only; and (3) Iowa only. The overstatement rates are comparable between the two states. The similarity between the two suggests adjusting the available rate by block size may offer a plausible estimate of the true availability rate from "all-in" data.

Table 2. Overstatement and Block Size

Miles²	Over-Stated	Share of Locations	Overstated Georgia	Overstated Iowa
0.001	0.001	0.100%	0.002	0.001
0.010	0.007	16.30%	0.009	0.003
0.050	0.022	39.60%	0.028	0.012
0.100	0.037	51.20%	0.048	0.022
0.500	0.110	80.40%	0.145	0.075
1.000	0.170	90.20%	0.223	0.124
5.000	0.399	99.40%	0.494	0.341
10.00	0.523	99.80%	0.622	0.474
25.00	0.683	100.0%	0.768	0.654
50.00	0.780	100.0%	0.847	0.766
100.0	0.855	100.0%	0.904	0.852
200.0	0.908	100.0%	0.942	0.910
300.0	0.930	100.0%	0.956	0.933

Table 2 summarizes the overstatement at various block sizes. For smaller blocks, the overstatement rate is near zero, but begins rising sharply when square mileage exceeds 0.05. Nearly 40% of the locations are in these relatively small blocks. The overstatement is less than 10 percentage points for 80% of locations in the sample (square mileage less than 0.50). The overstatement is sizable for very large blocks: a block of 100 square miles has an overstatement rate, on average, of about 86 percentage points, though a trivial number of locations are in such large blocks. Yet, few locations are found in the large blocks.

Conclusions

Collection of the FCC's Form 477 data assumes that if a single location in a census block has broadband, then *all* locations in that block have broadband (the "all-in" assumption). As a consequence, the Form 477 data presumably overstate broadband availability, and the data have been heavily criticized for doing so. When the Form 477 data are analyzed at the state level, the analysis in this PERSPECTIVE (consistent with the analysis in two of my prior works) finds that the overstatement is small—less than four

percentage points. Criticisms of these high-level statistics, therefore, are somewhat overblown.

Fortunately, several states and the federal government have begun collecting much more detailed data, and new broadband maps from the FCC are expected soon—both of which should improve subsidy allocation.

Still, if one wants to know the exact locations without broadband service availability, then the consequences of the Form 477's "all-in" assumption can be severe. In blocks that cover many square miles, the overstatement of availability can be very large. Thus, for subsidy allocation purposes, the Form 477 data are unreliable, at least for partially-served census blocks. Fortunately, several states and the federal government have begun collecting much more detailed data, and new broadband maps from the FCC are expected soon—both of which should improve subsidy allocation.

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 Data available at: https://www.fcc.gov/economics-analytics/industry-analysis-division/form-477-resources.

2 B. Nuelle, *Senators Push FCC to Improve Broadband Data Maps*, AGRI-PULSE (March 20, 2019) (available at: https://www.agri-pulse.com/articles/12010-senators-push-fcc-to-improve-broadband-data-maps); *see also* D. Clark, *Broadband Mapping Is a Mess. No One Knows What To Do About It*, BROADBAND COMMUNITIES MAGAZINE (July 2019) (available at: https://www.bbcmag.com/law-and-policy/broadband-mapping-is-a-mess-no-one-knows-what-to-do-about-it), noting that Senator Jerry Moran (R-Kansas) claimed the accuracy and value of the map "is nearly nil"; Representative Peter Welch (D-Vermont) said the maps are "bogus" and "phony maps"; and Senator John Tester (D-Montana) colorfully noted the "maps stink" and "we need to kick somebody's ass, truthfully."

3 G.S. Ford, *Quantifying the Overstatement in Broadband Availability from the Form 477 Data: An Econometric Approach*, PHOENIX CENTER POLICY PERSPECTIVE No. 19-03 (July 11, 2019) (available at: (available at: https://www.phoenix-center.org/perspectives/Perspective19-03Final.pdf).

4 G.S. Ford, *A Quality Check on Form 477 Data: Errors, Subsidies, and Econometrics,* PHOENIX CENTER POLICY PERSPECTIVE NO. 21-05 (October 27, 2021) (available at: https://www.phoenix-center.org/perspectives/Perspective21-05Final.pdf). The 7.3% figure is calculated as the implied availability in the Form 477 data from the estimated availability rate reported by BROADBANDNOW. *See* J. Busby, J. Tanberk, and T. Cooper, *BroadbandNow Estimates Availability for all 50 States; Confirms that More than 42 Million Americans Do Not Have Access to Broadband*, BROADBANDNOW.COM (August 21, 2021) (https://broadbandnow.com/research/fcc-broadband-overreporting-by-state).

5 Data available at: https://broadband.georgia.gov/maps; https://ocio.iowa.gov/broadband-availability-map-version-5.

6 Data available at: https://www2.census.gov/geo/tiger/TIGER2020/TABBLOCK20.