2024 Urban Rate Survey – Fixed Broadband Service

Introduction

Every year, the Wireline Competition Bureau (Bureau) and the Office of Economics and Analytics (OEA) (together, Bureau/OEA) conduct the fixed broadband Urban Rate Survey (broadband URS). The broadband URS collects data on rates for standalone Internet access service charged by a representative sample of fixed broadband providers in urban census tracts[1] in the United States.

The main purpose of the broadband URS is to produce reasonable broadband comparability benchmarks for every possible service tier (i.e., a service plan with specified minimum download speed, minimum upload speed, and monthly capacity allowance). These benchmarks serve as rate caps to "help ensure that universal service support recipients offering [fixed voice and] broadband services do so at reasonably comparable rates to those in urban areas."[2]

To calculate these benchmarks, the Bureau/OEA have, over the years, used essentially the same sample design, data collection mode, edit checks, and estimation methodology as what it had originally adopted in 2013. In particular, the same fixed sample size of approximately 500 sampling units had been selected and the same benchmark definition had been used every year.

To account for the fact that the Bureau/OEA must calculate benchmarks for a much wider range of speeds than when the broadband URS was first initiated, since the 2022 broadband URS, the Bureau/OEA have increased the sample size to about 2,000 sampling units. Correspondingly, we have slightly modified the sample design and estimation methodology to better capture the variation in broadband service rates across the United States, and thereby improve the quality of the benchmark estimates.

This methodology report follows the format of previous years' reports and describes how the Bureau/OEA calculated the fixed broadband reasonable comparability benchmarks for 2024.

Sample Design

Primary sampling unit and sampling frame

The 2024 broadband URS retains the same definition of primary sampling unit (PSU) as had been used in past survey cycles. That is, a PSU is a pair consisting of a broadband service provider and an urban census tract where the provider offers at least one terrestrial fixed broadband service tier to residential customers therein. In rare cases where this pair is distinguishable based on the provider's designation as both an incumbent local exchange carrier (ILEC) and a non-ILEC in the census tract, or the availability of both fixed wired and wireless service options, the PSU definition accommodates these distinctions.

As in previous years, the Bureau/OEA developed the sampling frame for the 2024 broadband URS based on data from the Broadband Data Collection tool (previously, this information was collected via FCC Form 477 data collection tool). This sampling frame uses data as of December of the year prior to data

[1] Prior to the 2023 URS, urban census tracts were defined as tracts with at least one populated block located within an urban area or urban cluster that is also located within a county designated as a metropolitan statistical area. Because the Census Bureau has updated the definition of urban areas using the results of the 2020 Census, the Bureau/OEA adopted a new definition of urban tracts: a 2020 tract is urban if at least 80 percent of its housing units are within a 2010 tract Urban Area that has a population of at least 50,000.

[2] *Connect America Fund*, WC Docket No. 10-90, Order, 28 FCC Rcd 4242 (WCB/WTB 2013).

collection. The 2024 broadband URS frame consists of 253,802 PSUs, encompassing 1,065 service providers and 58,568 census tracts.

Stratification

The broadband URS uses a stratified sample design. Stratification is the division of a heterogenous population (represented by the sampling frame) into subpopulations called strata (singular: stratum), each of which is internally homogenous with respect to the population characteristic(s) of interest. When properly implemented, this commonly used sample design element can produce gains in precision in the estimates of characteristics of the whole population.[3]

For the past several years, the Bureau/OEA has stratified the broadband URS sampling frame based on combinations of the following five major factors:

1. Continental United States[4] versus Alaska;
2. Provider's affiliated holding company;
3. Low- versus high bandwidth, where the distinction lies in whether the provider's service tier(s) in the census tract has minimum download speed lower than, or at least equal to, 500 Megabits per second (Mbps);
4. Providers of terrestrial fixed wireless (TFW) versus non-TFW (i.e., wired) service; and
5. "Major" versus "Minor" providers, where the distinction is algorithmically determined by a clustering method that captures dissimilarities in the number of occupied housing units to which the providers offer service.

Because of the significant increase in sample size, the Bureau/OEA stratified the frame slightly differently for the 2022 broadband URS. Specifically, the Bureau/OEA implemented the following modifications:

- More granular division of the range of download speeds based on the following cut points, in Mbps: 2, 36, 75, 115, 155, 250, 300, 500, 750, and 1,000

  The goal of this design modification is "to reduce the number of survey responses requested … while still capturing the [variation in] offered rates in accordance with their estimated effect on the reasonable comparability benchmark."[5] This granular division does not apply to holding company-based strata, which retained the low- versus high bandwidth distinction.

  If not enough sampling units in the frame are available to construct separate strata based on this more granular scheme, the Bureau/OEA collapsed the strata appropriately.

- Removal of the "Major" versus "Minor" strata

  Bureau/OEA staff's analysis of historical broadband URS sampling frame data suggests that carving out separate strata based on the number of occupied housing units introduces design redundancy while contributing little, if any, to achieve the goals of stratification. This is because

[3] William G. Cochran, Sampling Techniques ch. 5 (3rd ed. 1977).

[4] All 50 U.S. states except Alaska, together with the District of Columbia and Puerto Rico.

[5] Supporting Statement, 2020 Urban Rate Study Statistical Methodology Part B, OMB Control Number 3060-1192 (Dec. 2019) *available at* https://omb.report/icr/201911-3060-013/doc/97437001.

the sample design uses this number to assign a measure of size to each sampling unit before selecting the sample (as described in more detail below). As a result of this finding, the Bureau/OEA decided not to define strata based on the number of occupied housing units.

Similar modifications were applied to 2023 broadband URS. Starting in 2024, TFW and non-TFW providers were not distinguished as separate strata; instead, census division was used as a stratification variable to reflect variation in price data across geographic areas, in addition to the predefined strata for top holding companies.

These modifications resulted in the formation of a total of 23 strata for the 2024 broadband URS, containing 12 holding company-based strata, nine census division-based strata (not including Alaska), Puerto Rico, and Alaska. By comparison, the 2023 broadband URS had 33 strata. The table in the section "Measure of Size and Sample Selection" shows summary information on these strata.

Sample Allocation

For the 2024 broadband URS, the Bureau/OEA first assigned 176 sampling units to Alaska to make sure all PSUs of Alaska are surveyed and then allocated the remaining 1,824 sampling units to the 22 strata of Continental United States using proportional allocation.[6] Proportional allocation produces sample counts that vary according to the number of potential subscribers: the higher the number of potential subscribers in a stratum, the more sampling units are allocated for selection from that stratum. The Bureau/OEA switched to a proportional method last year because it reflected the distribution of potential subscribers. Also, the year-to-year strata mapping became impractical due to constant changes in the landscape of what companies could offer within a year. In addition, the Bureau/OEA implemented a minimum sample size of 20 for a stratum to avoid unreliable estimate due to small sample size.

The table below includes information on the sample allocation for the 2024 broadband URS.

Measure of Size and Sample Selection

The broadband URS implements probability sampling, i.e. every sampling unit has *some* chance of being selected in the sample, but not equal probability sampling, where every sampling unit has an *equal* chance of selection. Instead, the broadband URS sample design calculates a measure of size (MOS) for every sampling unit in the frame, and selects the sample independently within each stratum based on this MOS. Thus, for example, if sampling unit A has a MOS that is twice that of sampling unit B, then A is twice as likely to be selected in the sample compared to B. This type of unequal probability selection is called probability proportional to size (PPS) sampling.[7]

In the previous broadband surveys, the MOS was calculated differently for TFW and non-TFW providers based on the concept of a provider's presence ratio in an urban tract it served. Because TFW and non-TFW are no longer differentiated in stratification in the 2024 broadband URS, the Bureau/OEA calculates the MOS for the 2024 broadband URS sampling units based on data directly from Broadband Data Collection filings. Specifically, tract-level residential service subscription counts as reported by providers in their 2022 December Broadband Data Collection filings were used as MOS.

After completing the stratification, sample allocation, and measure of size calculation steps, the Bureau/OEA calculated selection probability for all PSUs to ensure that the units with selection probability greater or equal to one are separated and treated as certainty units (multiple rounds may be

[6] Cochran, *supra* note 6 at 96-99.

[7] *Id*. at 251.

needed) before using SAS proc surveyselect with systematic PPS selection option[8] to select the final sample of noncertainty units. The procedure selected a total of 2,000 sampling units.

The below table summarizes the sample design for the 2024 broadband URS.

	Frame			Sample		
Strata	Sampling Units	Providers	Census Tracts	Sampling Units	Providers	Census Tracts
AT&T	28,598	9	28,598	80	8	80
Altice	3,912	2	3,912	80	2	80
Charter	21,007	1	21,007	80	1	80
Comcast	26,353	1	26,353	80	1	80
Cox	5,439	1	5439	80	1	80
Frontier	5,642	1	5,642	80	1	80
Google	1,997	11	1,971	80	10	80
Lumen	9,479	2	7,945	80	2	80
Radiate	3,189	10	3,189	80	9	80
T-Mobile	56,948	1	56948	80	1	80
Verizon	53,491	9	48,434	80	8	80
WideOpenWest	2,636	2	1,318	80	2	80
New England	1,157	33	758	20	8	20
Middle Atlantic	1,680	63	1,423	46	12	46
East North Central	4,737	178	3,405	140	24	139
West North Central	3,342	156	1,540	136	27	134
South Atlantic	3,664	135	2,727	152	41	151
East South Central	1,328	69	1,018	56	15	55
West South Central	3,458	164	2,154	60	21	60
Mountain	5,464	152	3,006	76	28	75
Pacific	5,575	156	4,134	90	23	90
Puerto Rico	4,530	39	784	88	7	85
Alaska	176	6	77	176	6	77
Overall	253,802	1,065	58,568	2,000	219	1,879

[8] The SURVEYSELECT Procedure, SAS User's Guide, https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.3/statug/statug_surveyselect_toc.htm.

Survey Response

This section describes how the sample of 2,000 sampling units responded to the 2024 broadband URS.

This year, there are two ineligible sampling units because the selected provider did not offer, or stopped offering, fixed broadband service in the selected census tract. Of the remaining 1,998, 28 did not respond to the survey. Thus, the overall response rate is 99% = (100 x (1,998 – 28) / 1,998).

The table below shows the number of responses, the number of different service providers, and the number of different census tracts requested and received at the close of data collection for the 2024 broadband URS.

Survey Status	Responses	Service Providers	Census Tracts
Requested	2,000	219	1,879
Received	1,970	203	1,851
Service Provided	1,964	202	1,846

After conducting edit checks, including the removal of submitted rate data for business instead of residential plans and those rates where download speed is either less than 2 Mbps or less than upload speed, a total of 1,963 responses had useable unique monthly rates. Monthly rates were treated as unique for a combination of census tract, FCC Registration Number (FRN), service name, technology, download speed, upload speed, and capacity allowance. A total of 13,332 unique monthly rates were used to estimate the 2024 broadband comparability benchmarks.

The next table shows summary information on how these 13,332 unique monthly rates used to fit the average rate model distribute by technology.

Technology	Responses	Service Providers	Census Tracts	Rates
Cable	949	69	949	5,190
DSL	406	42	406	1,898
Fixed Wireless	217	29	193	1,603
FTTH	1,180	157	1,169	4,641
Total	**1,963**	**202**	**1,845**	**13,332**

Monthly Rates and Rate Spreads

The main analysis variable for the broadband URS is the monthly rate which broadband providers charge their customers in urban census tracts. It is common, however, for providers to offer multiple service tiers in the same census tract at different monthly rates. For this reason, the survey asks for the minimum and maximum of these rates and calculates an “average” monthly rate based on these two extreme values. Specifically, the following equations are used to calculate this average monthly rate, if the service provider offered multiple rates in the census tract:

- Minimum Rate = Minimum Monthly Charge + Minimum Other Mandatory Charge + Minimum Surcharge

- Maximum Rate = Maximum Monthly Charge + Maximum Other Mandatory Charge + Maximum Surcharge
- Rate Spread = Maximum Rate - Minimum Rate
- Average Rate = (Minimum Rate + Maximum Rate)/2

The following equations were used if the service provider did not offer multiple rates in the census tract:

- Average Rate = Minimum Monthly Charge + Minimum Other Mandatory Charge + Minimum Surcharge
- Rate Spread = 0

Weights

The broadband URS uses weights to ensure the contributions of each response properly represent the offers that consumers possibly receive nationwide. Weights are also used to ensure that a service provider's rates do not exert extra influence on the estimate only because the provider offers different services using multiple technologies.

The 2024 broadband URS weight construction is consistent with the method used in previous years. That is, each rate was assigned a weight based on the following equation:

Weight = Sampling Weight x Nonresponse Weight x Same Rate Weight x Service Level Weight x Number of Service Subscribers

Sampling Weight is the inverse of the selection probability for each sample unit. The selection probability is determined by the sample size in each stratum, the units' number of service subscribers described in the sample selection section earlier, and the total number of service subscribers in each stratum. Each sample is assigned a sampling weight to reflect its selection probability.

Nonresponse Weight is assigned to each stratum in order to compensate for unit nonresponse in each stratum. It is the total number of service subscribers sampled over the total number of service subscribers in the sampled census tracts of a given provider who has provided rate responses in each stratum.

Same Rate Weight is assigned to the respondents that provided i) multiple service levels or ii) equal service levels via different technologies for the same rate in the same census tract.[9] In such cases, the rate was assigned a Same Rate Weight equal to 1/R, where R is the number of rate responses provided by a service provider at the same rate in the census tract.

Service Level Weight is assigned to the respondents who provided multiple rates for the same service level offered via different technologies and/or service names. Each rate was assigned a Service Level Weight equal to 1/L, where L is the number of responses with different rates provided by a service provider for the same service plan (same download bandwidth, upload bandwidth, and monthly capacity allowance) in the census tract.

Number of Service Subscribers is the number of service subscribers in the tract.

The final weight is the product of Sampling Weight, Nonresponse Weight, Same Rate Weight, Service Level Weight, and the Number of Service Subscribers.

[9] Such a situation could arise when a provider uses different technologies to provide similar services to customers in different parts of a census tract.

Average Rate Model

The 2024 URS shows that broadband rate is nonlinear in proportion to download speed and upload speed (see Appendix A). To estimate an average rate for every possible combination, we applied a weighted Generalized Boosted Model (GBM),[10] which is an algorithm allowing nonlinearity[11] to all terrestrial fixed broadband services with download speeds between 2 Mbps and 1 Gigabit per second (Gbps), inclusive.[12]

The 13,332 rates used in the final analysis ranged from $11.40 to $1,000.00, with a weighted standard deviation ranging from $8.78 to $90.43. The rates vary widely across technologies. The following table shows the range, weighted mean, and weighted standard deviation of these rates, as well as the weighted mean download speed for different technologies.

Technology	**Min**	**Max**	**Rate, weighted mean**	**Rate, weighted standard deviation**	**Download speed, weighted mean**
Cable	14.95	315.00	95.26	24.54	651.81
DSL	24.95	189.98	71.71	8.78	47.39
Fixed Wireless	11.40	300.00	52.37	36.58	75.71
FTTH	15.00	1000.00	151.21	90.43	2627.24
All	**11.40**	**1000.00**	**70.72**	**57.15**	**473.46**

We undertook a weighted GBM[13] based on the following form:[14]

---

[10] *See* Appendix B.

[11] Ideally, we would calculate directly the weighted means and the weighted standard deviations of rates for all services. However, our samples do not cover all possible combinations of services provided to consumers nationwide. Therefore, we use a statistical model to estimate rates for all possible services.

[12] The 2018 broadband average rate model was the first year to include data with download bandwidths between 2 and 1000 Mbps. The 2017 broadband linear regression only models average rates between 2 and 50 Mbps.

[13] The average rate model based on a weighted GBM for the 2024 URS allows nonlinearity in rate per download bandwidth and rate per upload bandwidth by stratum. For further information, see Appendix B.

[14] We used the R package "gbm: Generalized Boosted Regression Models" to perform model fitting. We used random 50% of data as training set and 50% of data as validation set for each regression tree phase. Multiple GBM

$$\text{Average Monthly Rate (\$)} = Y = f(D, U, A, ST)$$

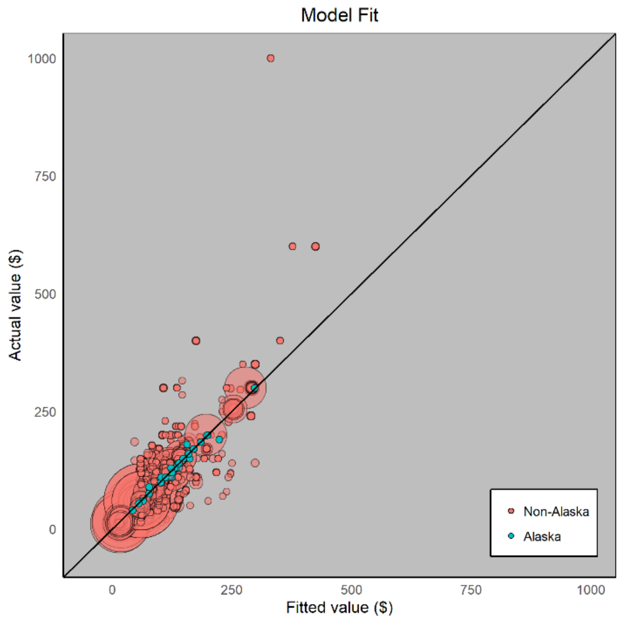
where D is download speed in Mbps, U is upload speed in Mbps, and A is the inverse of monthly capacity allowance in GB. ST includes the 23 strata as shown in the sample design summary table on page 4. The average monthly rate estimate is a function of D, U, A, and ST.

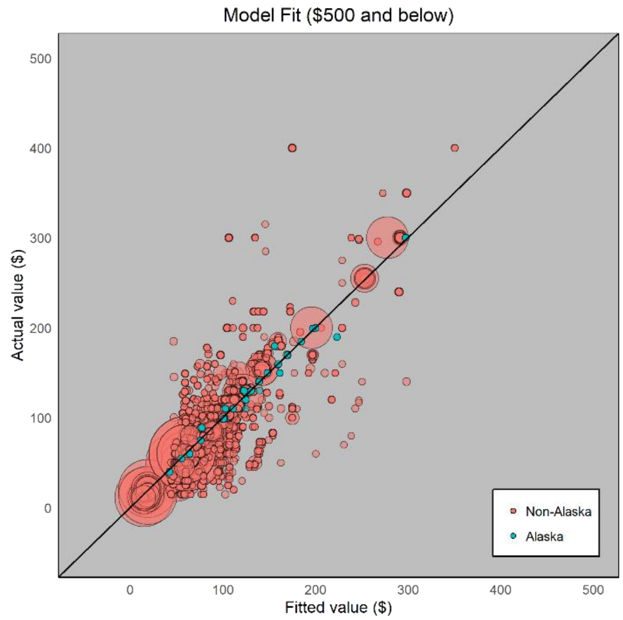
We estimated the U.S. average monthly rate as:

$$\text{U.S. Average Monthly Rate (\$)} = \sum_{i=1}^{n} \gamma_i \, \mathrm{E}(\mathrm{Y} \mid \mathrm{D}, \mathrm{U}, \mathrm{A}, \mathrm{ST} = \mathrm{ST}_i)$$

where $n = 22$, which represents 22 strata in the continental U.S. $E(Y \mid D, U, A, ST = ST_i)$ is the expected value conditioned on combinations of download speed, upload speed, and capacity allowance for a given stratum. The γ_i is the proportion of total continental U.S. residential service subscribers in a given stratum.

The plots below show how the model fits the raw data. The closer the dots are to the 45-degree line, the better the fit. The size of the circles represents the weights of the sample rates.





models were constructed and compared. Our final model was selected based on root mean square errors. The optimal number of trees of our final model is 462 for Alaska and 459 for Non-Alaska based on the out-of-bag error statistic, which is a method of measuring the prediction error of boosted decision trees.

U.S. Reasonable Comparability Benchmark

Under the methodology previously adopted by the Bureau, the reasonable comparability benchmark is the estimated average monthly rate plus twice the standard deviation of rates for terrestrial fixed broadband service plans with download bandwidths of 10 Mbps or greater, upload bandwidths of 1 Mbps or greater, and meeting or exceeding the minimum monthly usage allowance. The root weighted mean squared residual (RWMSR) is an estimate of the standard deviation of rates for service plans meeting the reasonable comparability benchmark criteria.[15]

The 2024 broadband URS average rate model approximates rate per download bandwidth and upload bandwidth closely. Therefore, the RWMSR of rates does not show a trend by download bandwidth and upload bandwidth. As before, we calculated the RWMSR values separately for the continental U.S. and Alaska. The table below shows the final RWMSR values.

	RWMSR
Continental U.S.	7.49
Alaska	3.62

Following the definition adopted by the Bureau, the U.S. reasonable comparability benchmark is calculated as follows:

U.S. reasonable comparability benchmark ($) = U.S. Average Monthly Rate + 2 ($RWMSR_{ContinentalUS}$)

= U.S. Average Monthly Rate + 14.98

The U.S. average monthly rate estimator is described in the previous section.

Alaska Reasonable Comparability Benchmark

For the Alaska reasonable comparability benchmark, the average monthly rate model is defined as follows:

$$\text{AK Average Monthly Rate (\$)} = \sum_{j=1}^{m} \gamma_j \, E(Y \mid D, U, A, ST = ST_j)$$

where m = 176, which represents the 176 strata (PSUs) in Alaska. $E(Y \mid D, U, A, ST = ST_j)$ is the expected value conditioned on combinations of download speed, upload speed, and monthly capacity allowance for a given stratum in Alaska. The γ_j is the proportion of total Alaska residential service subscribers in a given Alaska stratum.

The AK reasonable comparability benchmark is the Alaska average monthly rate plus two times its RWMSR:

AK reasonable comparability benchmark ($) = AK Average Monthly Rate + 2 ($RWMSR_{Alaska}$)

= AK Average Monthly Rate + 7.24

[15] RWMSR is the square root of the weighted average of the square of residuals (observed rate minus average rate as defined by the Average Monthly Rate equation) plus the square of the spreads divided by 12.

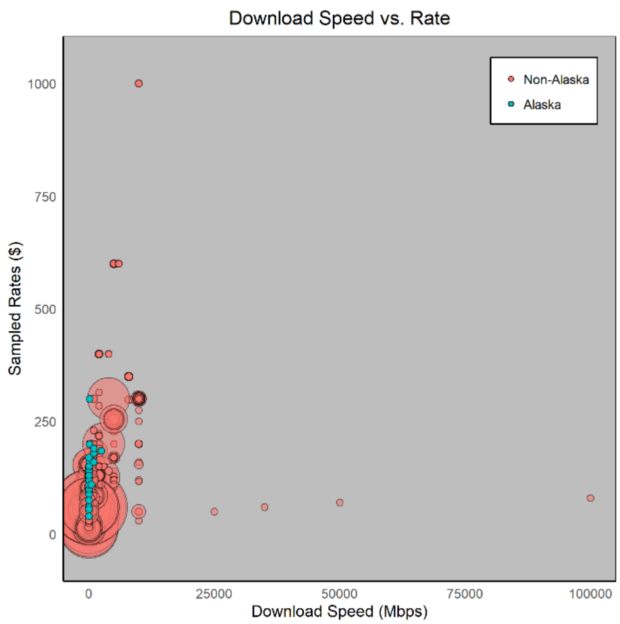
Reasonable Comparability Benchmark Results

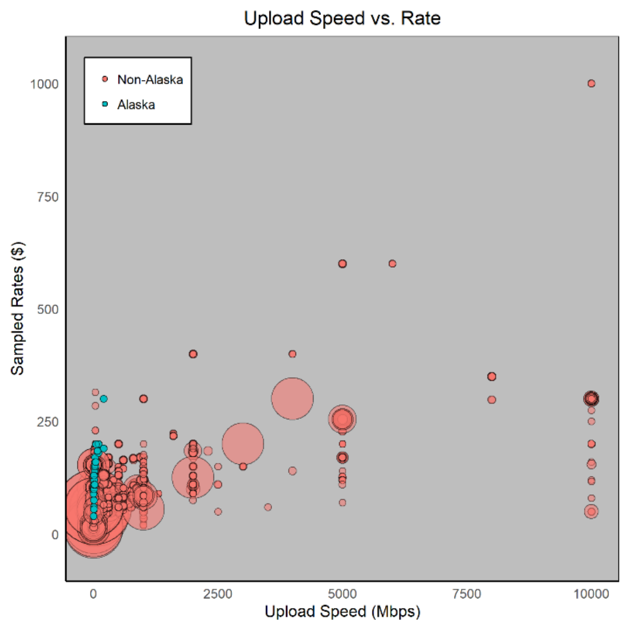
The table below provides examples of reasonable comparability benchmarks (rounded up to the nearest cent) for several service plan levels. The estimates are available for a reasonable comparability benchmark for lower download speeds (greater than or equal to 4 Mbps) if needed and up to download speeds of 1,000 Mbps. Upload speed may not exceed download speed.

Download Bandwidth (Mbps)	Upload Bandwidth (Mbps)	Capacity Allowance (GB)	2024 U.S.	2024 Alaska
4	1	660	$89.76	$62.62
4	1	Unlimited	$89.94	$64.22
10	1	660	$89.17	$82.24
10	1	Unlimited	$89.35	$84.38
25	3	660	$87.83	$83.24
25	3	Unlimited	$87.83	$85.18
50	5	660	$88.38	$71.24
50	5	Unlimited	$88.38	$71.29
100	20	660	$92.26	$125.89
100	20	Unlimited	$92.26	$145.75
1000	500	660	$118.24	$240.58
1000	500	Unlimited	$118.24	$240.58
25	5	660	$87.93	$83.24
25	5	Unlimited	$87.93	$85.18
100	10	Unlimited	$88.07	$129.41
250	25	Unlimited	$89.98	$172.75
500	50	Unlimited	$106.33	$196.78
1000	100	Unlimited	$117.26	$167.09

APPENDIX A

The 2024 URS modeled rates by download speed and by upload speed. Over this large range, the rates are not linear functions of either quantity. The size of the circles in the plots below represents the weights of the sample rates. Sampled rates represent common services provided to the customers and do not include all possible combinations of download bandwidth, upload bandwidth, and monthly capacity allowance.





APPENDIX B

A Generalized Boosted Model (GBM) is a machine learning algorithm that combines regression trees and gradient boosting techniques. The GBM framework does not assume a specific pattern between the independent variables and the dependent variable. It illustrates nonlinearity and interactions well without the need to define complex mathematical equations.

The algorithm first selects a portion of data to "train" a regression tree model (regression tree phase). The regression tree model used in GBM is usually a stump-only model or with only very few branches. Then, it uses the unselected data to "validate" the model and output a user defined performance statistic or loss function (validation phase). The algorithm repeats the same procedure on the residuals from the previous modeling phases until the performance gain stabilizes or loss function optimizes (gradient boosting phase). The outputs of a GBM are model fits from a series of regression tree models. Therefore, conventional coefficients are not applicable. Independent variable collinearity and data outliners have very little impact on the model fit because only the most influential variables are selected during each regression tree phase (only one most influential variable is selected if fitting a stump-only model). The interactions are naturally embedded in the structure of a series of regression tree models. Overfitting is safeguarded by inserting a cross-validation technique. Therefore, the GBM algorithm is considered to have high predictive accuracy. However, its predictive performance is weakened when the relationship between an independent variable and the dependent variable is very linear. More information about GBM can be found in the following references:

Y. Freund and R.E. Schapire. 1997. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences. 55(1):119-139.

G. Ridgeway. 1999. The state of boosting. Computing Science and Statistics. 31:172-181.

J.H. Friedman, T. Hastie, and R. Tibshirani. 2000. Additive Logistic Regression: a Statistical View of Boosting. Annals of Statistics. 28(2):337-374.

J.H. Friedman. 2001. Greedy Function Approximation: A Gradient Boosting Machine. Annals of Statistics. 29(5):1189-1232.

J.H. Friedman. 2002. Stochastic Gradient Boosting. Computational Statistics and Data Analysis. 38(4):367-378.