SUPPORTING STATEMENT

Part B.   Statistical Methods:

B.1   *Survey Purpose*

The purpose of the annual Urban Rate Survey (URS) is to collect price data separately for standalone residential telephone service (voice URS) and internet access service (broadband URS) currently charged by a representative sample of fixed voice and broadband providers in urban Census tracts in the United States.

The goal of the URS is to publish "reasonable comparability benchmarks" that serve as price caps for new residential fixed voice and broadband services to be offered the following year by providers that participate in at least one High Cost program under the Universal Service Fund. Because all such providers will have to offer these new services to non-urban areas, the benchmarks ensure that these providers offer these services at reasonably comparable rates for the same or similar services in urban areas.

There is only one voice benchmark, but there are as many broadband benchmarks as the number of internet service plans (or "tiers") defined by three parameters: minimum download speed, minimum upload speed (both in megabits per second, Mbps), and monthly capacity allowance (in gigabytes, GB). Moreover, the broadband benchmarks are produced separately for the continental United States and Alaska.

For both voice and broadband services, the definition of reasonable comparability benchmark is the same: national average price, plus twice the standard deviation. While this definition refers to a "national average," the URS does not produce national average fixed voice or broadband price estimates that can be regarded as official statistics in any context.

B.2   *Respondent Universe*

The universe of respondents for the URS consists of providers of fixed voice or broadband services that are required to file, and did file, the Form 477 in December of the year prior to data collection. This universe is limited to providers of residential services, and for broadband is further limited to providers of terrestrial services. This definition of the respondent universe will not change even after the Form 477 is decommissioned and replaced by the Broadband Data Collection (BDC).

The URS collects voice or broadband price data at the Census tract level. By its very nature, the URS limits its scope only to urban Census tracts, defined historically as those 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 (MSA). In the latest survey cycle, the voice URS universe consisted of more than 650 providers in more than 54,700 urban Census tracts, and the broadband URS consisted of about 1,600 providers in more than 58,200 urban tracts.

The release of 2020 Census data requires a change in the definition of urban tracts because (1) Census tract boundaries have been redrawn; (2) the definition of urban areas has been changed; and (3) the boundaries of urban areas have not yet been published. Thus, starting from the 2023 URS, urban tracts are defined as 2020 Census tracts in which at least 80% of housing units are within a census 2010 urban area with a population of 50,000 or more. This definition will be further modified once the Census Bureau

publishes the boundaries of 2020 urban areas. The modified definition of URS urban tracts excludes tracts within MSAs that are in smaller urban agglomerations (urban clusters), as well as tracts that are not within any urban area at all. However, excluding these areas ensures that the resulting urban tracts are actually urban, which is consistent with the intent of the URS data collection.

B.3 *Sample Design*

B.3.1 Sampling Frame

For both voice and broadband URS, the frame consists of primary sampling units that are pairs (X,Y), where X is a voice or broadband provider as described in the section on *Respondent Universe*, and Y is an urban Census tract in which X provides service. In the latest survey cycle, the voice URS frame consisted of more than 111,500 sampling units, and the broadband URS frame consisted of about 190,500 sampling units.

The current construction method for the URS sampling frames will be retained in future survey cycles.

B.3.2 Stratification

Both voice and broadband URS use a stratified design. However, the exact strata definitions differ between the two.

Voice
Departing from what had been done in previous rounds, the 2022 voice URS carved out separate strata for AT&T in addition to the usual distinction between incumbent and competitive local exchange carriers (ILECs and CLECs/non-ILECs). This change was adopted based on what had been done on the broadband side, where separate strata are defined for holding companies that operate nationwide (see the discussion of stratification for *Broadband*, below), for the purpose of reducing burden on carriers with nationwide footprint.

The desired result of reduction in reporting burden for AT&T was nominally achieved, but at the expense of a reduction in the quality of AT&T price data.

In future voice URS cycles, stratification will still retain the ILEC and non-ILEC strata, and possibly carve out separate strata for some holding companies, if doing so will reduce the reporting burden for any respondent.

Broadband
In the last five survey cycles of the broadband URS prior to 2022, the stratification of the frame was based on combinations of the following five major factors:
   (a) Continental United States (the 50 states except Alaska, the District of Columbia, and Puerto Rico) versus Alaska;
   (b) Provider's affiliated holding company;
   (c) Low- versus high bandwidth, where the distinction lies in whether the provider's service tier in the Census tract has minimum download speed lower than, or at least equal to, 500 Mbps;
   (d) Providers of terrestrial fixed wireless (TFW) versus non-TFW (i.e., wired) service; and
   (e) "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.

In the 2022 URS, strata definitions were modified either to conform with the design proposal submitted in 2019 to the Office of Management and Budget (as part of the request for a four-fold increase in the broadband sample size), or to remove redundancies in the sample design. The former resulted in a more granular division of the range of download speeds that define strata along this factor, and the latter resulted in the removal of "major" and "minor" strata. While the latter removed four strata, the former added more than a dozen. The net result was an almost doubling in the total number of broadband strata, without a corresponding increase in the diversity of broadband service tiers these modifications were expected to yield.

In future URS cycles, the following changes to the broadband stratification will be implemented:

- Strata based on holding company will be defined only if absolutely necessary to reduce the reporting burden for any respondent

  Stratification by holding company was adopted in prior broadband URS rounds primarily to reduce the reporting burden for nationwide broadband providers that had consistently been assigned significantly more samples than other providers, because of their extensive footprint. This burden is exacerbated by the manual method for entering the required price data, which not only are broken down by urban tract, but also by technology for delivering broadband service.

  As a direct result of modernization of the URS data collection instrument, the option to upload price data will be available starting with the 2023 URS. Because this option eliminates the need to manually enter price data, there may not be any need to carve out separate strata for holding companies that operate nationwide. Moreover, when such companies certify that they have a uniform pricing structure in some, if not all, of their service areas in the country, we can accept a smaller volume of price data and use these prices where uniform pricing apply.

- TFW and non-TFW providers will not be distinguished as separate strata

- Geography will be used as a stratification variable

  For the broadband benchmarks, which are model-based and not directly estimated from survey data, the models include a geographic variable (specifically, the state where the service is provided). Because state had never been used for stratification in prior broadband URS cycles, the number of price data points collected and used for modeling vary widely by state.

  Stratifying by state, however, results in a lot of strata. For this reason, stratification will instead be done by Census Region or Census Division.

- Strata defined by minimum download speed will be adjusted based on price data from the last five URS cycles

B.3.3   Sample Size and Allocation

Both voice and broadband URS use a fixed sample size: about 500 provider-urban tract pairs for voice, and, starting from the 2022 URS, around 2,000 such pairs for broadband. These sample sizes will be retained in future URS rounds.

Consistent with the standard practice in other federal surveys, sample allocation in the broadband URS has traditionally been based on stratum-level estimates of variation in the variable of interest (here, price) as observed in the previous survey round. That is, optimal allocation has been employed to produce broadband URS sample counts in such a way that more samples are allocated to current strata with higher price variance estimates based on the prior strata they map to.

However, because this method requires carefully mapping the current strata onto the prior year strata, changes in frame stratification, such as the ones outlined above, will result in misalignment in strata mapping. For this reason, proportional allocation, which is the allocation method used in the voice URS, will be used instead. Additionally, in no case will a given stratum receive a sample size less than 20. This minimum stratum sample size will apply to future voice and broadband URS rounds. Also, we may specify a stratum-level cap in allocated sample size for holding company-based strata, if these strata have to be defined in order to reduce the reporting burden on any respondent.

B.3.4   Measure of Size

The current method for calculating the measure of size (MOS) in both voice and broadband URS relies on the concept of a provider's "presence ratio" in an urban tract it serves. In particular, the MOS represents an estimated count of the provider's potential subscribers in the tract. However, the exact calculation of this number differs between the voice and broadband frames.

For the voice URS, the MOS is calculated differently for ILEC and non-ILEC providers. The MOS calculation for ILEC providers, while straightforward, depends on whether the provider is the sole carrier, or one of at least two carriers, in the Census tract. The calculation for non-ILEC providers, on the other hand, requires a log model be fit using information from the broadband frame but applied to data in the voice frame, and assumes that "the relationship of voice provider presence to voice subscribership is similar to that of broadband provider presence to broadband subscribership."

For the broadband URS, the MOS is calculated differently for TFW and non-TFW providers. For TFW providers, the MOS is calculated as twice the number of subscribers it reported on the Form 477 for an urban tract (because TFW providers "could no more than double their residential customers within a few months"). For non-TFW providers, the provider's presence ratio in an urban tract (the ratio of the number of subscribers it reported on the Form 477 for the tract, relative to the total subscribers for all providers in the tract) is applied to the total housing population in the tract (a static number tied to the most recent decennial Census data on tract-level housing unit counts) to obtain the MOS.

Future voice and broadband URS will adopt a common, simpler approach that uses data directly from the Form 477 to calculate the MOS. In particular, tract-level residential service deployment counts reported by providers in their latest December filing of the Form 477 will be used as the new MOS.

B.3.4   Sample Selection

Both voice and broadband URS implement probability sampling, which means that every sampling unit has *some* chance of being selected in the sample. However, neither uses equal probability sampling, where every sampling unit has an *equal* chance of selection. Instead, both voice and broadband URS calculate a measure of size (MOS) for every sampling unit in their frame. The sample is then selected independently within each stratum based on this MOS. In this way, sampling units with higher MOS values have a higher chance of selection. This type of unequal probability selection is called probability proportional to size (PPS) sampling.

Future voice and broadband URS will continue to use PPS sampling without replacement and select the sample using SAS proc surveyselect.

B.3.5 <u>Weighting</u>

Selection probabilities are automatically calculated by proc surveyselect based on standard PPS formulas. To be specific, if n_h is the sample size allocated to stratum h, and $z_h = \sum_{i=1}^{n_h} z_i$ is the total MOS for the stratum, where z_i is the MOS for the ith sampling unit, this unit's selection probability, π_i, is given by

$$\pi_i = \frac{z_i}{z_h} * n_h.$$

Depending on the distribution of the MOS in the stratum, this equation may calculate selection probabilities greater than one for some sampling units. This is an artifact of PPS sampling and when it happens, a common procedure is to first set the selection probability manually to one and treat these sampling units as certainties and separate them from the rest of the units. Then, we calculate selection probabilities again for the remaining sampling units in the stratum. We repeat this process until there are no certainty units in the stratum.

Base sampling weights w_i are just the reciprocal of selection probabilities:

$$w_i = \frac{1}{\pi_i} = \frac{z_h}{z_i * n_h}.$$

In particular, certainty units have base weights equal to 1.

After data collection, both voice and broadband URS currently apply additional weight adjustments to account for multiple levels or multiple rates in broadband services, or for simultaneous provision of iVoiP and circuit voice services in the same urban tract.

To ensure that the sample represents the stratum, the same number currently used as the MOS is also applied. In addition, non-response adjustments are commonly applied to remedy unit non-response.

Future voice and broadband URS will use the same weighting methodology.

B.4 *Estimation*

The voice benchmark is currently calculated based on survey data alone, consistent with the definition as given in the section on *Survey Purpose* above. Future voice URS rounds will continue to follow this benchmark calculation method.

The broadband benchmark, on the other hand, has never been calculated based on survey data alone. As discussed in the same section on *Survey Purpose,* it is impossible to get enough price data points to calculate benchmarks in this manner, for all possible combinations of the three parameters that define a broadband service tier. To produce benchmarks for all possible service tiers realized in the broadband market, and especially for services from providers that participate in the High Cost programs, a model

that includes at least these three parameters as independent variables, and price as dependent variable, must be used instead.

Since the 2019 URS, an average rate model for broadband has been consistently developed using a weighted Generalized Boosted Model (GBM), and used to calculate the broadband benchmarks. For details about this model, the reader is referred to the most recent broadband URS methodology report. Future broadband URS cycles will continue to use this model for benchmark calculation.

B.5 *Response Rates*

The voice and broadband URS generally yield response rates above 80%. These high response rates from both surveys are expected to continue. Staff directly contact any provider that is sent a survey notification that does not complete the online survey form within 30 days. Because compliance is mandatory, failure to comply may lead to enforcement action, including forfeiture penalties, pursuant to the Communications Act of 1934, as amended, and other applicable law. Based on the sampling methodology described above and an anticipated high rate of compliance, the information collected can be expected to be both sufficiently accurate and reliable for the purpose of calculating the reasonable comparability benchmarks.

B.6 *Statistical Consultations*

For additional information concerning the Urban Rate Survey, please contact the following subject matter experts:
- Craig Stroup at (202) 418-0989, or craig.stroup@fcc.gov;
- Yongping Hao at (202) 418-2987, or yongping.hao@fcc.gov; or
- Edgardo Cureg at (202) 418-4537, or edgardo.cureg@fcc.gov.