

Revised collection entitled: Consumer Fixed Broadband Services Testing and Measurement

Part B: Collections of Information Employing Statistical Methods:

1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection methods to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.

The target population for this study is the American consumer with broadband. The survey delivers reliable national estimates of broadband speed performance for the US Internet Service Providers (ISP) delivering broadband service to the majority of fixed and mobile consumers. The program current supports data collection from a set of roughly 10,000 fixed broadband hardware devices and is set to launch a crowd-sourced mobile broadband measurement using Android Application clients in Q4 2013.

Fixed Broadband Measurement Statistical Methods

In response to RFQ-10-000013 issued on March 24, 2010, a contract award was made to SamKnows to conduct the FCC’s first fixed broadband performance data collection. SamKnows developed a panel of some 10,0000 fixed broadband subscribers, geographically dispersed across the United States, to conduct measurements to determine the national estimates of actual broadband speed performance per ISP for 34 services tiers, and a reliable estimation of performance by ISP packages within geographical region was conducted. Three annual broadband speed measurements have been collected through hardware devices placed in consumers’ panelists homes in March 2011, April 2012, September 2012, and September 2013. For 2014 we expect an even larger set of panelists that may allow us to measure broadband speeds across the USA with a much higher degree of confidence.

Internet Service Provider	Type	Estimated Subscribers [†]
Comcast	Cable	15,930,000
AT&T	DSL	15,789,000
Time Warner	Cable	9,289,000
Verizon	DSL	9,220,000
Cox	Cable	4,200,000
Charter	Cable	3,062,300

Internet Service Provider	Type	Estimated Subscribers [†]
Qwest	DSL	2,974,000
Cablevision	Cable	2,568,000
CenturyLink	DSL	2,236,000
Windstream	DSL	1,132,100
Mediacom	Cable	778,000
Frontier	DSL	635,947
Insight	Cable	501,500
Cable ONE	Cable	392,832
RCN	Cable	312,000
FairPoint	DSL	295,000
Cincinnati Bell	DSL	244,000

[†] Estimated at Q4 2009.

The various media channels used for recruitment have been selected so as not to correlate with the broadband performance (as long as the media channels reach the population without pre-selection from the media itself). Media used will be kept on record to ensure at the entirety of the respondents from a given media does not differ from the performances from respondents from other medias.

Note that broadband speed performance may be considered to correlate with

- Distance between premises and exchange,
- Contention in the ISP's own network particularly at peak time,
- Location density, such as rural or urban; typically in urban areas there are greater availability of higher speed services in urban area, further more in rural areas the average line length from local exchange to premises is longer than in urban areas.
- The technology used. There are different technologies available to delivered broadband services.

Therefore although this sample is based on pool of volunteers, there is no good reason to believe that this particular sample would be different to a random sample as SamKnows control the region, U/S/R split and services tiers. The panelist behavior, as an individual, has no impact on the technical performance of broadband received nor on the main influencers of performances.

Source Quality and Size of Frame

The size of 10,000 has been designed to provide broadband speed performance at

- geographic region level;
- ISP's level; and
- service tiers level by region.

To achieve requirement at regional service tiers level, we require a minimum of 125 panelists for each service tier present at regional level. There are 34 services tiers, which equates for the 4 regions to 80 services tiers, since not all 34 services are present in the 4 regions. Therefore the total sample of 10,000 is required (80 x 125).

Strata Definition and Proposed Allocation

The sample is split using census 4 regions: West, Midwest, Northeast and South.

For each services tiers, SamKnows will target 125 panelists per region when the service tiers is present (see table below, a bucket=125 panelists). Quotas will also be set on Urban-Suburban-Rural (U/S/R) continuum in order for each region to be representative of the Urban-Suburban-Rural universe per region. The classification of the data will be based on an analysis of the location of the exchange using the US Census bureau definition Urban-Suburban-Rural. Using the exchange location, each panelist will be categorized in terms of Urban-Suburban-Rural.

Within that primary split, we subgroup the service tiers further by region, with other subgroups relating to Density and Geographical region.

Region breakdowns are:

Northeast Region (including the New England division): Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont; and the Middle Atlantic division: New Jersey, New York, and Pennsylvania.

Midwest Region (including the East North Central division): Illinois, Indiana, Michigan, Ohio, and Wisconsin; and the West North Central division: Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota.

South Region (including the South Atlantic division): Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia; the East South Central division: Alabama, Kentucky, Mississippi, and Tennessee; and the West South Central division: Arkansas, Louisiana, Oklahoma, and Texas.

West Region (including the Mountain division): Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, and Wyoming; and the Pacific division: Alaska, California, Hawaii, Oregon, and Washington.

Overall Precision Requirement

Our overall precision requirements are an error margin of $\pm 4.4\%$ or $\pm 8.8\%$ (depending on whether or not the sample size is 125 or 500) with 95% confidence. (Typical minimum market research standards are a 90% confidence with an error margin of $\pm 7.5\%$).

Sample error at service tiers by region, assuming the sample is comparable to a random sample, is 8.76% for 125 panelists, and 4.38% for 500 panelists at 95 % confidence.

This is calculated using the formula:

$$e = z\sqrt{(p\%(100-p\%))/s}$$

Where:

e = sampling error (the proportion of error we are prepared to accept)

s = the sample size

z = the number relating to the degree of confidence required

p = an estimate of the proportion of people falling into the group in which we are interested in the population

2. Describe the procedures for the collection of information including:

- *Statistical methodology for stratification and sample selection,*
- *Estimation procedure,*
- *Degree of accuracy needed for the purpose described in the justification,*
- *Unusual problems requiring specialized sampling procedures, and*
- *Any use of periodic (less frequent than annual) data collection cycles to reduce burden.*

Building on its experience completing a similar fixed broadband performance project for the United Kingdom telecommunications regulator, Ofcom, SamKnows recruited a US-based panel and deployed its technology throughout all fifty states to cover all regions. Sample quotas were set for 4 regions, West, Midwest, Northeast and South. The quotas, defined by geography, technology and service level, have been named buckets. Within each bucket, a soft quota was defined by population density using Census Bureau definitions.

SamKnows best practices in panel recruitment developed in its global efforts have been reviewed by 3rd parties during the course of the project. The panel recruitment consisted of two steps, using a multi mode recruitment effort to build a large pool of volunteers. Volunteers were encouraged to sign-up to a project specific website (www.testmyisp.com), providing an initial panel of volunteers. This was supplemented by more targeted participants who were recruited with the support of the ISPs, who targeted specific customer groups with mailings. The approach provides flexibility in building the panel to ensure geographical representation, excluding outliers and, as much as possible, maximizing the multi mode recruitment selection to avoid pitfalls such as selection bias, coverage error, and panel attrition.

SamKnows provides report data when deemed accurate enough to be useful. Accuracy is reflected by sample size and variance. In order to reflect the limitation of the sample SamKnows shows the speed results as an interval using the 95% interval around the mean. With a sample of up to 500 national wide panelists, the associated sample error is 4.4% and from experience the confidence interval for broadband speed interval is typically less than 0.5 Mbit/s.

Estimates for sub groups – including buckets – are subject to higher sampling error margins based strictly on the smaller number of panelists in a specific subgroup e.g. for each bucket - 125 panelists the sample error is 8.8% and the expected confidence interval is around 0.8 Mbit/s.

Weighting

To insure representativeness of the US broadband population, and to obtain national average speed, the results are weighted by density: urban and rural populations consistent with standard Census Bureau definitions. The penetration data by region is made available from third parties.

To compare ISPs' performances for DSL, SamKnows normalizes the data by copper loop length (sometimes called "last mile").

Normalization by loop length is critical for DSL, because with this technology, speed degrades as the length of the "last mile" copper line to the premises increases. Therefore operators that have a higher proportion of customer in rural areas where loop length is typically higher may be

expected to deliver lower speed than those who focus on towns and cities because they have different customer profile. In order to normalize the distance, we will use the “last mile” loop length.

A weight adjustment is calculated to the contribution to the average speed made by each respondent based on their “last mile” copper loop length by matching the percentage observed in distance bands to the percentage in the distribution of the total sample distribution.

To reduce burden on the panelist, information technology is used extensively, all data collection is automated after the initial installation of a hardware device. The speed and performance is monitored through the hardware devices in consumers’ homes, to accurately measure the performance of fixed line broadband connections based on real-world usage. These hardware devices are controlled by a cluster of servers, which host the test scheduler and reporting database. The data is collated on the reporting platform and accessed via a reporting interface and secure FTP.

Further Information on DSL Weighting Methodology

DSL performance results are not weighted based on loop length. Technical properties of DSL technology dictates that performance can degrade as the length of the line from the exchange to the consumer’s modem increases. However, the study focuses on measuring broadband performance as presented by the carrier to the consumer, and varying technology limitations are or should be reflected in the speed tier and other performance offered to the consumer. Weighting of performance by distance would introduce a potential distortion of the broadband performance offered to consumers of a given tier.

Mobile Broadband Measurement Statistical Methods

In 2012, the FCC announced its intent to expand the existing program to include measurement of mobile broadband performance using a crowd-sourced application deployed on Android smartphone devices. In a four-month privacy by design process, the FCC developed a privacy policy to guide its data collection and Open Data policies. The process brought together privacy experts, government agencies, academics, manufacturers, carriers, public interest and other interested stakeholders to address technical, legal and other policy concerns resulting in a completely anonymized data collection process that minimized risk of any individual being able to be identified from the pool of anonymous performance data.

The FCC instituted a policy to collect anonymous data and make no public use without first undergoing a technical privacy review to delete or process datum to minimize re-identification risks to participant volunteers submitting data. The technical privacy review will employ clustering and other cutting edge statistical techniques in analyzing the bulk raw collected data. The privacy review will produce recommended business rules for the deletion of particular results or the coarsening of data elements located in areas limited total number of samples for the defined geographic area. The review will also provide recommendations for allowable combinations of data elements for data sets to be released to the public. The technical privacy review will be performed after data collection and the statistical analysis will make

recommendations that take into account the total number of samples collected as well as the variance found in the collected data.

The FCC released the FCC Speed Test App for Android in November 2013, and the FCC Speed Test App for iPhone in February 2014. The smartphone measurement application is designed to be installed on a user's smartphone either directly or via an app store, such as Google Play or iTunes. The application is free to download, but carrier charges may apply. The application will run continuously in the background, periodically performing measurements. Volunteers with Android phones are configured to use automated testing. The FCC Speed Test app automated testing function can be disabled and the app can be configured to start a test only when manually executed, but due to limitation of the operating system iPhone devices do not have automated testing capability and can only execute the speed test manually. Android devices also support collecting more information about cellular performance that is not supported by iPhone devices.

3. Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.

Maximizing Response Rates in Fixed Test Panel

Once the fixed volunteer panel is recruited and the SamKnows 'Whiteboxes' deployed, the tests will run on a pre-configured schedule, subject only to changes in the schedule and the volunteers own use of their broadband connection. Therefore, 'response rates' will always be at a 'maximum'. Details on the proprietary framework are provided below:

Panel Recruitment

The recruitment strategy will solicit volunteers primarily through a media campaign using social and traditional media, such as consumer and technology press, alongside Twitter and independent bloggers and opinion formers. The FCC and SamKnows will both conduct this outreach. We are confident in doing this based on significant public interest in this study and past success that SamKnows had with this approach when conducting a similar project in the UK. These efforts would outline the project and direct interested volunteers to a URL where they could sign-up.

Once the volunteer panel is recruited and the SamKnows 'Whiteboxes' deployed, the tests will run on a pre-configured schedule, subject only to changes in the schedule and the volunteers own use of their broadband connection. Therefore, 'response rates' will always be at a 'maximum'. Details on the proprietary framework are provided below:

Test Scheduling

Tests will be run every hour on a randomize timing in the hour - 24 hours a day, 7 days a week. Within a month period, the software within the white box unit will perform over 700 separate

speed tests. The unit is installed immediately in front of their home internet connection. This ensures that tests can be run at any time, even if all home computers are switched off.

Testing Locking

No two tests may operate simultaneously. The first test to begin will have priority, and the other will block until the first finishes (or a timeout is reached).

Threshold Manager

Both before and during testing, the Threshold Manager analyzes the UDP & TCP packet data going over the WAN interface on the unit to check whether the Internet connection is in active use, The Threshold Manager is configurable by design, but is set by default at 400Kbps. If this threshold is breached before or during the tests, activity is and the threshold test is then repeated for one minute, for a maximum of 5 times until either the amount of traffic returns to a level below the threshold or the tests are suspended and that time period is marked as having a busy line.

Order of Tests

The tests are run in the following order:

- ping - for all target hosts serially
- dns - for all target hosts serially
- www - for all target hosts serially
- single-threaded http get
- multiple-threaded http get
- single-threaded http post

Required number of data points

For reporting purpose, the data will be aggregated first per hour for the reporting period and then overall to minimize the effect of missing data. If less than 5 data points are recorded for one time slot within a month, the data will be discarded.

Tests

Web browsing

Measures the time taken to fetch the HTML and referenced resources from a page of a popular website. This test does not test against centralized testing nodes; instead it tests against real websites, ensuring that content distribution networks and other performance enhancing factors may be taken into account.

By default, each Whitebox will test three common websites on every test run. The time taken to download the resources, the number of bytes transferred and the calculated rate per second will

be recorded. The primary measure for this test is the total time taken to download the HTML page and associated resources.

The results include the time taken for DNS resolution. The test uses up to eight concurrent TCP connections to fetch resources from targets. The test pools TCP connections and utilizes persistent connections where the remote HTTP server supports them.

The test may optionally run with or without HTTP headers advertising cache support (through the inclusion or exclusion of the “Cache-Control: no-cache” request header).

Video streaming

This generic streaming test can be configured to model the characteristics of a variety of voice and video protocols. For the purpose of the video streaming test, the intention is to simulate an end user viewing a streaming video over one of the many popular websites that provide this service (e.g. YouTube).

The test operates over TCP and uses a proprietary client and server side component. The client and server negotiate the test parameters at the start of each test.

A playout buffer is configured and the client will attempt to download data from the server at the maximum rate necessary to ensure that this buffer is never empty. A separate thread is reading data from this buffer at a fixed rate, looking for buffer underruns (which would manifest themselves to users as a pause in video). The client will record the time to initial buffer, the total number of buffer underruns and the total delay in milliseconds due to these underruns.

It is expected that the bitrate of the streaming will vary according to the access line speed being tested.

Voice over IP

This test utilizes the same generic streaming test as the video test, albeit with different configuration. The test operates UDP and, unlike the video streaming test, utilizes bi-directional traffic.

The client initiates a UDP stream to the server and a fixed-rate stream is tested bidirectionally. A de-jitter buffer of 25ms is used to reduce the impact of jitter. The test measures this disruption by monitoring throughput, jitter, delay and loss. These metrics are measured by subdividing the stream into blocks, and measuring the time taken to receive each block (as well as the difference between consecutive times).

By default, the test uses a 64kbps stream with the same characteristics and properties (i.e. packet sizes, delays, bitrate) as the G.711 codec.

Jitter is calculated using the PDV approach described in section 4.2 of RFC5481.

Availability Test

Measures the availability of the network connection from the Whitebox to multiple target test nodes by sending and receiving TCP segments to a receiving server located on each test node.

The client establishes long lived TCP connections to the server on each test node, periodically sending TCP packets containing:

- A Magic number to enable the server to differentiate multiple clients
- Send timestamp (microseconds)

The server echoes back the same data to the client and if it fails to respond or the connection is reset via TCP RST or FIN then the client will attempt to re-establish the connection. If the client is unable to re-establish the connection to all 3 servers simultaneously, it is inferred that Internet connectivity is at fault, the test records a failure locally, along with a timestamp to record the time of failure.

To aid in diagnosing at which point in the route to the target test nodes the connectivity failed, a traceroute is launched to all target test nodes, the results of which are stored locally until connectivity is resumed and the results can be submitted.

This test is executed when the Whitebox boots and runs permanently as a background test.

Latency / Ping and packet loss

This test uses ICMP pings to measure latency and packet loss to popular in country websites. Three target hosts will be tested, each receiving three pings (the first of which will be ignored). The round trip time for each is recorded individually, as well as the number of pings that were not returned.

Whilst ICMP packets may be dropped by routers under heavy load, their simplicity still provides one of the most accurate measures of latency. Extended periods of packet loss over many units on a single ISP may indicate a congested network, which would become another metric that could be tracked.

DNS Resolution Time and Failure Rate

DNS resolution of the ISPs' recursive DNS resolvers is tested by querying their DNS servers for popular USA websites. A standard "A Record" query is sent, and resolution time and success/failure results are recorded.

Two of the ISP's recursive DNS resolvers are tested directly from the monitoring unit. Queries are sent for three popular USA websites, with each DNS server being tested independently.

Note that these tests do not rely on the DNS servers set on the user's router.

Web page loading

This test fetches the main HTML body of a website. Note that additional resources, such as images, embedded media, stylesheets and other external files are not fetched as a part of this test.

The time in milliseconds to receive the complete response from the web server is recorded, as well as any failed attempts. A failed attempt is deemed to be one where the web server cannot be reached, or where a HTTP status code of something other than 200 is encountered.

We will use the home page of three popular country-hosted websites and tests will be run every hour against these. Note that tests were designed to ensure that pages were not cached.

HTTP test methodology

All of the tests are run against two or three in country managed servers dedicated purely to this task.

All servers have at least 1Gbps connectivity and have diverse routes through a multiple of transit providers. All servers reside on networks that peer directly or indirectly within one hop to the core network

Units attempting to perform a speed test will request the target speed test server from the data collection server, this allows us to remove speed test servers from the equation should they be temporarily overloaded or down for maintenance. Under normal operation the process will round robin between the servers.

The speed test servers are all configured to return immediate content expiry in the HTTP headers, ensuring that compliant proxy servers should not cache the speed test content.

The HTTP tests make use of a SamKnows designed test that includes a built in HTTP compatible client following RFC 2616

UDP Latency and Packet Loss

Measures the round trip time of small UDP packets between the Whitebox and a target test node. Each packet contains consists of an 8-byte sequence number and an 8-byte timestamp. If a packet is not received back within three seconds of sending, it is treated as lost. The test records the number of packets sent each hour, the average round trip time of these and the total number of packets lost.

As with the availability test, the test operates continuously in the background and will perform multiple randomly distributed tests within a one hour period.

Speed Tests

Measures the download and upload speed of the given connection in bits per second by performing single and multi-connection GET and POST HTTP requests to a target test node.

Binary non-zero content herein referred to as the payload is hosted on a web server on the target test node. The test operates for a fixed duration (5 seconds by default). The client will attempt to download as much of the payload as possible for the duration of the test. The payload and all other testing parameters are configurable and may be subject to change in the future.

Four separate variations on the test are supported:

- Single connection GET
- Multi connection GET
- Single connection POST
- Multi connection POST

Note that only the multi connection tests are currently intended for use in the FCC project.

Each connection used in the test counts the numbers of bytes of the target payload, transferred between two points in time and calculates the speed of each thread as Bytes transferred / Time (seconds).

Factors such as TCP slow start and congestion are taken into account by repeatedly downloading small chunks (default 256KB) of the target payload before the real testing begins. This “warm up” period is said to have been completed when three consecutive chunks were downloaded at the same speed (or within a small tolerance (default 10%) of one another). In a multi connection test, three individual connections are established (each on its own thread) and are confirmed as all having completed the warm up period before timing begins.

Content downloaded is output to /dev/null or equivalent (i.e. it is discarded), whilst content uploaded is generated and streamed on the fly from /dev/urandom.

Maximizing Response Rates in Mobile Test Panel

The accuracy and reliability of the data collected are achieved by employing the same techniques show to be effective in the fixed program. The crowd-sourced Android application collection automates testing using the same centrally managed test schedule methodology and randomizes tests in three periods daily, namely two peak periods and one off-peak period. The mobile clients execute upload, download, latency, and packet loss tests using the same technology developed for the fixed testing panel. The program employs a variety of efforts to ensure that potential participants are alerted to the privacy terms of the program efforts made to minimize risks that any individual may be identified from the anonymous pool of data. The panel will be developed using a multi-prong outreach effort in using social and traditional media to maximize the geographic distribution of the volunteers contributing data to the program. These efforts will help ensure that valuable data will be collected to maximize the ability to report on the geographic nature of mobile broadband performance while minimizing the risk to any individual volunteer’s privacy.

4. Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of test may be submitted for approval separately or in combination with the main collection of information.

The SamKnows tests and methodology have been evaluated and approved by Government Regulators, Internet Service Providers, leading Academics and researchers.

The stated service level name/identifier, headline speed, price and download limit will be used to check and assign panelist to the correct broadband packages. Different tests and checks will be held to confirm ISP and flag incorrect packages in the following manner:

At the recruitment stage

1. The ISP allocation will be validated using customer's IP against those held in ISP IP libraries. If inconsistent the record is flagged
2. Actual speed will be validated in relation to package and distance from premises to exchange distribution. Outliers will be excluded
3. Exclusion of household when distance to exchange is untypical to maximize the effective sample size after normalization

Test on continuous basis

4. Panelists have the ability to update their information through SamKnows web interface
5. Automatic flag of change of ISP and geographical location through the reporting system
6. Automatic flag of untypical maximum download speed

Independent Review: As there are a large number of stakeholders in this project, SamKnows has dedicated resource to liaising with interested 3rd parties who wish to participate in validating/enhancing the SamKnows methodology. These discussions are open to anyone who wishes to participate.

The mobile program has employed the same collaborative review process with diverse mobile broadband stakeholders in securing evaluation and approval of the collection approaches. The program has enlisted specialists in the mobile data privacy to address aspects of mobile broadband data that raise novel issues. The geographic distribution of data will be reviewed on a continuous basis to develop new outreach program targets.

5. Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.

Walter Johnston
Federal Communications Commission
Email: walter.johston@fcc.gov
Phone: 202-418-0807

James Miller
Federal Communications Commission
Email: james.miller@fcc.gov
Phone: 202-418-7351

Nathalie Sot
Statistician, SamKnows
Email: nathalie@samknows.com
Cell: +44 7715 484 803