National Telecommunications and Information Administration

Broadband Technology Opportunities Program Evaluation Study

Order Number D10PD18645

Raw Data Delivery

Submitted September 26, 2014

ASR Analytics, LLC
1389 Canterbury Way
Potomac, MD 20854

Federal TIN: 20-1204680
DUNS: 15-108-3305
GSA Schedule #: GS-10F-0062R

Submitted to:
Shelita Saint-Louis, Contracting Officer
Cassandra Sterba, Contract Specialist
Acquisition Services Directorate
National Business Center
Department of the Interior

Table of Contents

Section 1. Introduction

1.1 About the Evaluation Study

On September 17, 2010, the National Telecommunications and Information Administration (NTIA) awarded a task order to ASR Analytics, LLC (ASR) to complete an evaluation study of the Broadband Technology Opportunities Program (BTOP).[1] The scope of work includes an assessment of the benefits that BTOP grants are having on broadband availability and adoption, and in achieving social and economic benefits in areas served by the grantees.[2] A complete description of the methodology used in the evaluation study is available in the *BTOP Evaluation Study Design*.[3]

1.2 About this Document

In September 2014, ASR submitted a *Final Report* to NTIA that quantitatively and qualitatively assesses the social and economic impact of BTOP grants and discusses how NTIA's implementation of BTOP has encouraged the fulfillment of the goals of the American Recovery and Reinvestment Act of 2009 (Recovery Act).

The Statement of Work (SOW) requires ASR to prepare a *Raw Data Delivery*. Specific requirements include the following:

- "The Contractor shall also deliver any disaggregated, underlying data to the extent practicable, which will be made available by the Government for public use and consumption" (page 12 of 31).
- "Raw Data with an explanation of the data formats and access protocols the contractor plans to use in order to deliver the data to NTIA such that it is suitable both for NTIA and public use" (page 12 of 31).
- "This includes all data that created a foundation for later analysis & conclusions, as well as all data that could be utilized by future researchers. Data content level of detail, storage format, word processing and/or image format, and accessibility to multiple computer stations to be determined by the COTR based on program needs" (page 17 of 31).
- "Explanation of what data will be used to conduct the study, including how data inaccuracies or gaps will be managed, and what quality assurance processes will be employed" (page 19 of 31).
- "The Contractor is expected to amass a large amount of raw data in the execution of Case Studies and Longitudinal Study. All raw data shall be turned over to NTIA at the conclusion of the effort with an explanation of the data formats and access protocols the contractor plans to use in order to deliver the data to NTIA such that it is suitable both for NTIA and public use" (page 22 of 31).

This document supplements the *Final Report* and describes the data used and how the evaluation study team arrived at the results presented. The *Final Report* included the following types of analysis:

- For Public Computer Centers (PCC), Sustainable Broadband Adoption (SBA), and Comprehensive Community Infrastructure (CCI) grants, ASR summarized grant budgets, activities, and outcomes over time and cumulatively.

- ASR used budget data for PCC, SBA, and CCI grantees to estimate short-term economic impacts due to expenditures
- ASR applied findings in the broadband literature to grantee-reported and publicly available data to estimate the intermediate quantitative impacts of PCC and SBA grants
- ASR collected pricing and subscription speed data for community anchor institutions (CAI) in the evaluation study sample and applied changes due to BTOP to all connected CAIs, estimating broadband cost reductions due to CCI grants
- ASR used National Broadband Map (NBM) and publicly available data to estimate the effect of CCI grants on broadband availability
- ASR used estimated effects of CCI grants on broadband availability to extrapolate the long-term quantitative benefits of BTOP

1.3 About the Raw Data Delivery

The Raw Data Delivery includes the following components:

- "Raw Data Delivery": this document describes the processes used to create the results presented in the *Final Report*. There are two types of tables and figures in this document:
  1. Explanatory tables and figures are included in this document to provide background details and other information about the methodology used in ASR's analysis. These tables and figures do not mirror any tables or figures in the *Final Report*. In this document, explanatory tables and figures have captions that are numbered alphabetically (i.e., Table A, Table B).
  2. *Final Report* tables and figures included in this document mirror tables and figures in the *Final Report* that contain quantitative information. Not all tables and figures in the *Final Report* contain quantitative information, so not all tables in the *Final Report* are included in this document. In this document, *Final Report* tables and figures have identical captions to those in the *Final Report* (i.e., Table 2 in this document mirrors Table 2 in the *Final Report*).
- "File List and Descriptions": an Excel spreadsheet listing all individual files provided to NTIA. Individual sheets contain the following files and descriptions:
  - "Shared Source Code": snippets of code that are referenced by all R scripts used to perform any kind of data management or analysis for the *Final Report.* Section 2 discusses "Shared Source Code."
  - "Input Files": disaggregated, underlying data used in the *Final Report*, provided to the extent practical and extent allowed by terms of use. Section 3 discusses "Input Files."
  - "Prepared Data": data derived by ASR from "Input Files" to provide consistent methods of using primary data for multiple purposes. All provided "Prepared Data" are derived from provided "Input Files." Section 4 discusses "Prepared Data."
  - "Statistical Analysis": results of long-term statistical analyses performed by ASR. Statistical analyses are saved, then loaded when required to avoid unnecessary re-estimation. Section 5 discusses "Statistical Analysis" files.
  - "0. Executive Summary": tables and figures created for *Final Report Executive Summary* and all programs used to generate them. Section 6 discusses "0. Executive Summary" files.
  - "1. Introduction": tables and figures created for *Final Report Section 1. Introduction* and all programs used to generate them. Section 7 discusses "1. Introduction" files.
  - "2. Short-Term Impacts": tables and figures created for *Final Report Section 2. Short-Term Economic Impacts* and all programs used to generate them. Section 8 discusses "2. Short-Term Impacts" files.
  - "4. Long-Term Impacts": tables and figures created for *Final Report Section 4. Long-Term Impacts* and all programs used to generate them. Section 9 discusses "4. Long-Term Impacts" files.

- "6. Progress towards Goals": tables and figures created for *Final Report Section 6. Progress towards Recovery Act Goals* and all programs used to generate them. Section 10 discusses "6. Progress towards Recovery Act Goals" files.
- "C. Quant Intermediate Impacts": summaries of the quantitative intermediate impacts of BTOP for the *Final Report*, particularly *Final Report Appendix C. Quantitative Intermediate Impacts,* and all programs used to estimate them. Section 11 discusses "C. Quant Intermediate Impacts" files.
- "D. Long-Term Quant Analysis": tables and figures created for *Final Report Appendix D. Long-Term Quantitative Analysis* and all programs used to generate them; and summaries of the long-term benefits due to BTOP for the *Final Report,* particularly *Final Report Appendix D. Long-Term Quantitative Analysis,* and all programs used to extrapolate them. Section 12 discusses "D. Long-Term Quant Analysis" files.

- Data files: a series of folders containing all individual files listed in "File List and Descriptions," summarized above and discussed in this document

1.4 Software

1.4.1 R

R is a widely used open source, cross-platform statistical language and environment.[4] R scripts described in Section 2, Section 4, Section 5, Section 6, Section 7, Section 9, Section 10, Section 11, and Section 12 read, manipulate, aggregate, summarize, and visualize data and perform advanced statistical analysis. The output of these procedures were written to external files for inclusion in the *Final Report*.

The open-source nature of R fosters a large user community. This community tests and validates of software components and contributes extensions to R called packages. Packages provide functionality not included in the base R software, or alters the functionality in some way to improve the user experience. ASR used the following packages to import and manipulate data, create summary tables and figures, and perform statistical analysis:

- "data.table": enhances operations and improves manipulation performance with tabular data in R[5]
- "xlsx": reads from and writes to Microsoft Excel files, which are not natively supported by R.[6] "xlsx" depends on the following packages: "xlsxjars" and "rJava."[7] "xlsx" and its dependencies require that a recent version of the Java Runtime Environment (JRE) is installed and is defined in environment variables or (Windows only) the registry.
- "reshape2": provides flexible and powerful reshaping (transposing) capabilities for tabular data and arrays.[8]
- "plyr": includes a function to concatenate tabular data with columns that do not align[9]
- "ggplot2": provides a flexible, advanced approach to data visualization (graphics)[10]
- "scales": functions to manipulate scales and labels in graphics[11]
- "tables": methods to tabulate data (i.e., apply one or more summary functions by zero, one, or more grouping variables to create multidimensional tables).[12] "tables" depends on the "Hmisc" package.[13] "Hmisc" depends on the following packages: "lattice," "survival," and "Formula."[14]
- "lubridate": improves mathematical operations on formatted data values[15]
- "Matching": includes functions for multivariate nearest-neighbor matching based on Mahalanobis distance.[16] "Matching" depends on the "MASS" package.[17]
- "boot": functions to bootstrap statistical estimates and calculate confidence intervals[18]

1.4.2 Microsoft Word

Microsoft Word is a widely used word processor. ASR prepared the *Final Report* and this document using Microsoft Word. Microsoft Word is not required to reproduce any of the analysis or results discussed in the *Final Report*. However, summary and results tables produced through the methods described in this document were formatted in Microsoft Word for the *Final Report.* Therefore, the tables saved by R will lack the formatting of the tables in the *Final Report*.

1.4.3 Microsoft Excel

Microsoft Excel is a widely used spreadsheet tool. ASR transferred results from output files to the *Final Report* and this document using Microsoft Excel. Specific graphics in *Final Report Section 2. Short-Term Impacts* were created in Microsoft Excel. Aside from these graphics, Microsoft Excel is not required to reproduce any of the analysis or results discussed in the *Final Report*.

1.4.4 SAS

SAS is an industry-standard software tool for data management, manipulation, and analysis. SAS is used only for *Final Report Section 2. Short-Term Impacts*. SAS scripts described in Section 8 prepare BTOP data for economic impact analysis and manipulate the results of the impact analysis. The functionality of these scripts is not unique to SAS and could be implemented in another software system or language.

1.4.5 IMPLAN

Impact Analysis for Planning (IMPLAN) is an industry-standard tool for economic impact estimation based on input-output analysis. IMPLAN is used only for *Final Report Section 2. Short-Term Impacts*. IMPLAN files and procedures described in Section 8 estimate short-term economic impacts. IMPLAN is proprietary software owned by the Minnesota IMPLAN Group, Inc. (MIG). IMPLAN software is required to reproduce ASR's short-term economic impact estimates.

MIG is the creator of the IMPLAN software and data tools used for economic impact analysis. By categorizing expenditures that describe the structure and function of a particular economy, IMPLAN is able to create a model that calculates the extent of projected economic transactions in the geographic region. More than 2,000 public and private institutions use IMPLAN.[19]

The IMPLAN software calculates predicted impacts based on an input-output model. An input-output model categorizes the flow of dollars through an economy and assumes fixed relationships between producers and their suppliers based on demand. In an input-output model, the inter-industry relationships largely determine how an economy will respond to change. An increase in demand of a particular product or service causes a multiplier effect. In other words, increased demand of a product affects the producer of that product, the producer's employees, the producer's suppliers, the supplier's employees, and so on. Ultimately, the total effect on the economy is larger than the initial change in demand. These effects, or impacts, are categorized into the three groups described below:

- Direct impact: jobs, wages, and output created by the BTOP project itself
  - Example: a manager at a BTOP-funded public computer center
- Indirect impact: jobs, wages, and output created by the businesses that supply goods and services to the project ("supplier impacts")
  - Example: a concrete manufacturer providing materials to a CCI construction site
- Induced impact: the result of employees' (of direct and indirect impact jobs) spending of wages and salaries on food, housing, transportation, and other sectors

- Example: the employees of the construction firm contracted by the CCI grantee spend a portion of their wages at nearby restaurants (induced or secondary impacts occur in nearly all sectors of the economy, although primarily in the service sector)

Economic input-output models, like IMPLAN, are the primary tools to measure the total economic impact of a policy or event. Recently, the federal government chose IMPLAN to measure economic impacts in the following published studies: *Economic Impact of Recovery Act Investments in the Smart Grid*; *Estimating the Impact on Employment of USDA's Programs in ARRA*; and *Economic and Fiscal Impacts of the Corps of Engineers' Trinity River Vision Project in Tarrant County Texas*.[20]

Tables, figures, and analysis results in the *Final Report* were generated from a consistent set of file types and software. There is one exception: *Final Report Section 2. Short-Term Impacts* is a summary of the findings of the *Short-Term Economic Impacts Report* submitted September 30, 2013. Those results are considered final and none of the analytical techniques, data, or results have been updated in any way since submission. Short-term analysis requires specialized impact analysis software that are not used for any other purpose. Therefore, results presented in *Final Report Section 2. Short-Term Impacts* are generated by different software than results in the remainder of the *Final Report*.

1.4.6 File Types

The tables, figures, and analysis in the *Final Report* use the following file types:

- Microsoft Excel (XLSX or XLS, *.xlsx or *.xls): All "Input Files" (with one exception), "Prepared Data," and results are stored in Microsoft Excel format. Data and results are stored in Excel to allow users to access data and inspect visually the results in any environment they choose.
  - Comma Separated Value (CSV): One "Input File" is stored in CSV format. The file was too large to be read consistently when it was stored in Excel format. The CSV format is natively supported by R and Microsoft Excel.
- R script (R, *.R): R scripts are plain text files containing sets of instructions written in the R statistical programming language. R scripts can be executed in the R statistical package.
- R data file (RDA, *.rda): R data files are compressed, binary representations of one or more R objects. Since it is the native R file format, an RDA file maintains the attributes of all objects it contains and can be quickly loaded.
- SAS script (SAS, *.sas): SAS scripts are plain text files containing sets of instructions written in the SAS programming language. SAS scripts can be executed in SAS software.
- Portable Network Graphics (PNG, *.png): The PNG format is a widely used cross-platform raster graphics format. The tools used to create the *Final Report* (i.e., R and Microsoft Word) natively support the PNG format.
- IMPLAN database (*.impdb): IMPLAN database files store the input data and procedures used to estimate impacts. IMPLAN database files are proprietary and must be opened with IMPLAN software.

Section 2. Shared Source Code

The “Shared Source Code” directory contains three simple R script files. R scripts are sets of instructions written in the R statistical programming language that can be executed in the R statistical package. The scripts in “Shared Source Code” include:

- “Define Child Directories”: defines folder paths within the parent directory that contain output for inclusion in the *Final Report*
- “Define Functions”: user-defined functions that are used frequently across the R scripts described in this document. All of the functions in this file are wrapper functions. In other words, they take the exact arguments of an existing R function, execute that function, then manipulate the results in some way:
  - The “Read.xlsx2,” “Dcast,” and “Rbind.fill” user-defined functions are equivalent to calling the “read.xlsx2,” “dcast,” and “rbind.fill” R functions and converting the results from data frames to data tables.
  - The “write.xlsx.tabular” user-defined function provides a method to write a tabular object to an Excel file, not natively supported by R. The user-defined function uses the “write.csv.tabular” R function to save a tabular object to a temporary file, reads the temporary file as a data frame, and writes the data frame to an Excel file.
  - The “theme_btop” user-defined function is a graphical theme applied to the figures included in the *Final Report* that were generated in R. This ensures uniform formatting across these figures.
- “Install and Load Packages”: lists all packages required for the *Final Report,* installs the packages if they were not previously installed, and loads all packages.

The three scripts are called by every R script in the remaining sections of this document. This ensures that every script uses the same directory definitions, eliminates the need to redefine user functions, and requires that all packages used across all R scripts are available for every R script, ensuring required packages will always be available.

Section 3. Input Files

The processes described later in this section rely on a number of input files and programs/scripts to perform the analysis. All input files contain data taken directly from primary sources. Unless otherwise noted, ASR did not modify the primary data contained in input files. Input files allow future researchers or data users to access the primary sources of data. The following subsections describe the input files in alphabetical order.

3.1 Awarded BIP Service Areas

The "Awarded BIP Service Areas" Excel file identifies the service areas of awarded Broadband Initiatives Program (BIP) grants. A service area is a group of one or more counties throughout the United States. "Awarded BIP Service Areas" contains nearly all counties impacted by awarded BIP infrastructure grants, including both last mile and middle projects as defined by BIP. For each grant, "Awarded BIP Service Areas" reports the grantee and the county Federal Information Processing Standards (FIPS) code for each county in the grantee's service area.

"Awarded BIP Service Areas" may not exactly match the service area definition of the grantees, as ASR was not able to identify complete service areas for some awarded BIP grants. Partial service areas are included in "Awarded BIP Service Areas" in these cases. Awarded technical assistance, satellite, and rural library grants have been excluded from this list. Awarded BIP projects in Hawaii, Puerto Rico, and the Island Areas of the United States were also not recorded.

The primary source used to create "Awarded BIP Service Areas" is the United States Department of Agriculture's (USDA) 2011 *Advancing Broadband* report.[21] This report contains all awarded BIP grants as well as a brief description of each project. Where the proposed funded service areas are not clear from the report, the definitions of the service areas included in "Awarded BIP Service Areas" are based on the following factors:

1. tower site locations;
2. fiber to the premises (FTTP) or fiber to the home (FTTH) locations;
3. communities to be served by wireless broadband services;
4. impacted tribal entities and federal and state American Indian Reservations.

Most of the service area counties included in "Awarded BIP Service Areas" were identified using a combination of the four factors listed above. In circumstances where USDA (2011) defines the service area as only the grantee's service territory but does not mention an exact area (as is sometimes the case with incumbent local exchange carriers), the service territory specified on the grantee's website defines the service area.[22] Service areas in "Awarded BIP Service Areas" are described at the county level. For grantees with service areas at the town, community, or tribal reservation level, service areas were converted to counties based on location data.

For grants with service areas not specified in the project description of USDA (2011), ASR examined ProPublica (2012).[23] In addition to recording all awarded BIP grants, this database tracks the stimulus funds received by each county. ASR searched for counties in ProPublica (2012) that received BIP funding.[24]

3.2 Awarded BTOP Service Areas

The "Awarded BTOP Service Areas" Excel file identifies the service areas of awarded BTOP grants. A service area is a group of one or more counties throughout the United States. "Awarded BTOP Service Areas" contains nearly all counties impacted by awarded BTOP infrastructure grants, including last mile and middle mile projects. For each awarded BTOP grant, "Awarded BTOP Service Areas" lists the award number and the proposed funded service area counties. Due to redactions, only partial service areas for some of the awarded grants were established. "Awarded BTOP Service Areas" excludes awarded infrastructure grants in Alaska, Hawaii, and the outlying areas of the United States, public safety networks, and satellite-based service grants. In addition, awarded grants that were later defunded are not included.

Grantee applications, attachments, project fact sheets, Quarterly Performance Progress Reports (PPR), and websites were used to define service areas.[25] In some circumstances, the proposed funded service area was explicitly stated in the broadband application. When the proposed funded service area was not stated in the application, impacted counties were identified using the following methods:

1. lists or declarations of CAIs in applications or attachments;
2. lists or declarations of interconnection points in applications or attachments;
3. lists or declarations of POPs in applications or attachments;
4. maps of proposed infrastructure routes.

"Awarded BTOP Service Areas" uses county-level service areas. When the sources above identify only select cities, ZIP Codes, towns, or census tracts, ASR converted these locations to counties. Further, if application materials emphasized that a project was statewide, all of the counties within that state were defined as the service area.

3.3 Awarded BTOP Service Areas of with Redacted Locations

The "Awarded BTOP Service Areas of Redacted Locations" Excel file contains a list of county FIPS codes that correspond to counties with a community anchor institution (CAI) or Point of Presence (POP) redacted from publicly available data. As described in Section 4.28, ASR expanded service areas determined through research and case study reports using location data reported in CAC.[26] NTIA redacted CAI and POP locations for several grantees from Connecting America's Communities (CAC). To expand service areas for grantees with redacted locations, ASR pulled CAI and POP locations from 2012 Annual Performance Progress Report (APR) attachments for these grantees.[27] Grantees report geographic coordinates for all locations included in APR attachments. ASR related these coordinates to Census Bureau county shapefiles to determine the counties that contain a redacted CAI or POP location.[28] Since these locations were redacted, "Awarded BTOP Service Areas of Redacted Locations" only contains the counties that include at least one redacted location and does not include individual CAI or POP names and locations.

3.4 BEA County Personal Income

The "BEA County Personal Income - 2011" Excel file contains county-level estimates of Local Area Personal Income (LAPI) published by the United States Bureau of Economic Analysis (BEA) for all available counties in 2011.[29]

3.5 BEA Gross State Product

The "BEA Gross State Product - 2011" Excel file contains state-level estimates of Gross Domestic Product (GDP) published by BEA.[30] State-level GDP is referred to as Gross State Product (GSP). "BEA Gross State Product - 2011" contains all industry estimates (total GSP) data for all available states in 2011.

3.6 BLS Census of Employment and Wages - 2011

The United States Bureau of Labor Statistics (BLS) publishes county-level estimates of average annual wage each year as part of the Quarterly Census of Employment and Wages (QCEW).[31] The "BLS Census of Employment and Wages - 2011" Excel file contains average annual wage estimates for all available counties in 2011.

3.7 BLS Local Area Unemployment Statistics

BLS publishes state- and county-level annual estimates of labor force size, employment, unemployment, and the unemployment rate as part of the Local Area Unemployment Statistics (LAUS).[32] ASR uses the following LAUS input files:

- The "BLS Local Area Unemployment Statistics - 2009" Excel file contains county-level unemployment statistics published by BLS for all available counties in 2009.
- The "BLS Local Area Unemployment Statistics - 2011" Excel file contains county-level unemployment statistics published by BLS for all available counties in 2011.
- The "BLS Local Area Unemployment Statistics - 2012" Excel file contains county-level unemployment statistics published by BLS for all available counties in 2012.

3.8 BLS-DOL Unemployment Duration

The "BLS-DOL Unemployment Duration - 2009" Excel file contains state-level estimates of average unemployment duration jointly published by the United States Department of Labor (DOL) and BLS for all available states in 2009.[33]

3.9 CAC Data

NTIA annually publishes grantee activity and other information in Connecting America's Communities (CAC).[34] NTIA redacts and validates these data prior to publication.[35] The 2011 and 2012 CAC data sets contain data for PCC, SBA, and CCI grantees by year. ASR downloaded these data sets from the CAC website as Excel files and renamed them "CAC Data - 2011" and "CAC Data - 2012."

ASR makes use of the following Excel sheets:

- CAI Data: A table of connected CAI locations. For each CAI, data in the table include institution type and geographic coordinates.
- PCC Sites Data: A table of BTOP-funded PCC locations. For each location, data in the table include institution type, PCC type (new or improved), number of workstations, weekday and weekend hours, connection speed, and weekly users.

- SBA Programs Data: A table of locations hosting BTOP-funded SBA programs. For each location, data in the table include location name, institution type, and address.

3.10 CAC Data - CAI Locations with Counties

NTIA annually publishes grantee activity and other information in CAC.[36] NTIA redacts and validates these data prior to publication.[37] The 2011 and 2012 CAC data sets contain data for CCI, PCC, and SBA grantees by year. CCI data include geographic coordinates of CAI locations. ASR related these coordinates to county shapefiles published by the Census Bureau to determine the counties in which each location lies.[38] The "CAC Data - CAI Locations with Counties" Excel file includes every CAI location reported in the 2011 and 2012 data sets with the county FIPS code that corresponds to each location's coordinates.

3.11 CAC Data - POP Locations with Counties

NTIA annually publishes grantee activity and other information in CAC.[39] NTIA redacts and validates these data prior to publication.[40] The 2011 and 2012 CAC data sets contain data for CCI, PCC, and SBA grantees by year. CCI data include geographic coordinates of POP facilities throughout the United States.[41] Each record in the POP table contains geographic coordinates for two facilities. ASR related both sets of coordinates to county shapefiles published by the Census Bureau to determine the county in which each facility lies.[42] The "CAC Data - POP Locations with Counties" Excel file includes every pair of POP facilities reported in the 2011 and 2012 data sets with the county FIPS codes that corresponds to each pair's coordinates.

3.12 CCI Annual PPR Data

ASR used the cumulative total CAIs connected or passed by CCI grantees by CAI category from CCI APRs for further analysis in the *Final Report.*[43] APRs include the following CAI categories for connected or passed institutions: Schools (K-12); Libraries; Community Colleges; Universities and Colleges; Medical or Healthcare Facilities; Public Safety Entities; Other Community Support (Governmental); Other Community Support (Non-Governmental); and Public Housing. The "CCI Annual PPR Data" Excel file contains these data.

3.13 CCI Categorized Connected CAIs

CCI grantees reported connected CAIs individually in PPRs, along with their institution types and location names. ASR retrieved these data for all available PPRs submitted by CCI grantees.[44] ASR then cleaned these data by verifying the CAI type assigned to each institution using institution names. ASR manually classified institutions where classifications were missing or could not be confirmed. ASR eliminated all records without specific data to support confirmation or manual classification with any confidence. The "CCI Categorized Connected Locations" Excel file contains these data.

3.14 CCI Connected CAIs in Case Study Reports

During the development of CCI case study reports, ASR manually reviewed CAIs reported in PPRs, APRs, and grantee-provided lists to match connected CAIs to publicly available data sets with as much confidence as possible. The most common adjustments to these data were splitting reported locations that represented multiple institutions into the representative number of records

and recategorizing institution types. Any adjustments made are explained in detail in the notes of each case study report.

As a result of these adjustments, the total number of CAIs connected by projects reported in the case study reports is often different from the available APR and PPR totals. During the case study report review process, project representatives and NTIA program staff reviewed reported institution totals and in some cases made adjustments to the data presented in the draft CCI case study reports. To account for these adjustments, ASR stored the total number of connected CAIs and the last quarter of data included in this total for each evaluation study sample CCI grant in the "CCI Connected CAIs in Case Study Reports" Excel file.[45]

3.15 CCI Evaluation Study Sample Connected CAI Collected Speed and Pricing Data

During CCI site visits, ASR asked interviewees (grantees, CAIs, and service providers) for CAI subscription speed and pricing data for before and after BTOP-funded connections were established. The "CCI Case Study Connected CAI Collected Speed and Pricing Data" Excel file contains these data. Some interviewees provided estimates for typical sites when they lacked this information for all CAIs or for some CAIs. When interviewees provided estimates, ASR recorded details in the Notes column.

3.16 CCI Evaluation Study Sample Service Areas

During the CCI case study report process, ASR reviewed grantee applications, fact sheets, PPRs, APRs, and attachments to determine an initial list of service area counties for each grant in the evaluation study sample.[46] ASR also reviewed the service areas listed in the "Awarded BTOP Service Areas" table, described in Section 3.2, for any additional counties. Any other service area counties identified during site visit preparations, site visits, or correspondence with grantees were also added. This list was then confirmed with grantee representatives and refined as necessary.

ASR began its analysis in each case study with these lists. In some cases, ASR found that connected CAIs were located outside of these service areas. ASR used lists of connected CAIs in PPRs, APRs, and grantee-provided materials for this task.[47] When county names were not provided, ASR used geographic coordinates to determine counties by relating them to county shapefiles.[48] In cases when connected CAIs were located outside of the list of service area counties, the list was expanded to include counties containing connected CAIs. Any counties added due to a connected CAI outside of the initial service area were flagged.

The processes described above resulted in a list of service area counties for each of the twelve CCI grants in the evaluation study sample. These lists are provided in all CCI case study reports.[49] The "CCI Case Study Service Areas" Excel file contains the award numbers and service area county FIPS codes for all twelve evaluation study grants, as well as the flag if a county was added due to a connected CAI outside of the initial service area.

3.17 CCI Quarterly PPR Data

ASR used the following variables from quarterly PPRs for further analysis in the *Final Report*: cumulative deployed and upgraded network miles, cumulative new and upgraded leased miles, cumulative interconnection points, cumulative signed agreements, and cumulative connected CAIs.[50] The "CCI Quarterly PPR Data" Excel file contains these data.

CCI grantees also provide detailed information in quarterly PPRs on CAIs connected in a given quarter, including institution name, type, and service area town or county. ASR found that the

cumulative total CAI count does not always match the more detailed CAI information. Additional CAI information is often reported in PPR addendums, available to the evaluation study team as PPR attachments. However, CAIs are often duplicated across quarters and are sometimes reported with slightly different names. ASR assumes that the cumulative total connected CAI count is the most accurate CAI figure and uses this figure in the analysis.

3.18 Census Bureau ACS (2005-2009) Data

The United States Census Bureau annually publishes 5-year estimates of population, demographic, and economic statistics at granular geographic levels as part of the American Community Survey (ACS) program. ASR used the following census-tract-level variables from the ACS (2005-2009) Summary File for every available census tract in the United States: population (total, aged 15 to 64, 18 and older, 14 and older, and 65 and older), housing units (total and occupied), per capita income, average household size, labor force size, total employment, and total unemployment.[51] The "Census Bureau ACS (2005 - 2009) Data" CSV file contains these data.

3.19 Census Bureau ACS (2006-2010) Data

The United States Census Bureau's American Community Survey (ACS) collects demographic data on a sample of Americans to provide annually updated five-year estimates of population characteristics.[52] The "Census Bureau ACS (2006 - 2010) Data" Excel file contains the non-English speaking percentage of population and average household income for every available county in 2010.

3.20 Census Bureau County Business Patterns - 2011

The County Business Patterns (CBP) series provides county- and state-level annual estimates of economic activity for the North American Industry Classification System (NAICS) hierarchy of industry codes, including the establishment count, total employment, and total payroll.[53] CBP includes estimates for all industries combined in a county or state.

The "Census Bureau County Business Patterns" Excel file contains private nonfarm employment, total payroll, total establishments, and private nonfarm employment suppression flag data for all industries for all available counties in 2011. ASR added an additional variable representing the midpoint of the private nonfarm employment suppression flag.

3.21 Census Bureau County Shapefiles Selected Attributes

The TIGER/Line Shapefiles are extracts of selected geographic and cartographic information from the U.S. Census Bureau's Master Address File/Topologically Integrated Geographic Encoding and Referencing (MAF/TIGER) database.[54] TIGER/Line Shapefiles include attribute tables that describe the individual geographic areas included in the shapefiles. In county shapefiles, these attributes include county FIPS codes; Legal/Statistical Area Description (LSAD) codes, i.e., formatted county names; and land area. The "Census Bureau County Shapefiles Selected Attributes" Excel file contains these attributes.

3.22 Census Bureau State Business Patterns - 2011

The County Business Patterns (CBP) series provides county- and state-level annual estimates of economic activity for the NAICS hierarchy of industry codes, including the establishment count, total employment, and total payroll.[55] CBP includes estimates for all industries combined in a

county or state. The "Census Bureau State Business Patterns" Excel file contains private nonfarm employment data for all industries for all available states in 2011.

3.23 DOL National and State Minimum Wages - 2009

The DOL publishes a table of state and national minimum wage rates on its website.[56] ASR extracted the 2009 minimum wage rates from this table. Multiple minimum wage rates were published for several states. ASR used the highest minimum wage rate in these cases. The selected 2009 wage rates were saved to the "DOL National and State Minimum Wages - 2009" Excel file.

3.24 FCC Form 477 County Household Adoption Levels - 2011-06-30

The FCC publishes data on local telephone competition and subscribership to high-speed services, summarizing data gathered through FCC Form 477.[57] The "FCC Form 477 County Household Adoption Levels - 2011-06-30" Excel file contains household adoption levels for two different service levels: at least 768 kbps downstream and at least 200 kbps upstream; and at least 200 kbps either downstream or upstream. These data were formatted to present descriptive values for service level and adoption level, rather than numeric codes.

3.25 List of Awarded BTOP Grants

To reduce the amount of data stored in input files, ASR stored grant information in a single file that could be referenced whenever required. Data include the following variables: award number, internal identifiers (Award ID and EGID), recipient name, project title, BTOP round, applicant type, and grant type.[58]

ASR then created additional variables to identify whether the grant is included in the evaluation study sample, whether to exclude the grant from the evaluation study (public safety grants), and whether the project was defunded.[59]

The "List of Awarded BTOP Grants" Excel file contains this list of grants.

3.26 NBM Census Blocks with Availability Rate Decreases by County

NTIA provided NBM data, assembled by the NBM team at the FCC, to ASR for analysis in the *Final Report*.[60] Each release was provided at the census block level and included:

- block FIPS codes;
- population;
- population with 768 kbps downstream and 200 kbps upstream service available (Notice of Funds Availability [NOFA] definition of broadband);
- population with 3 Mbps downstream and 768 kbps upstream service available (NTIA definition of broadband).

Broadband availability data are collected at the state level by individual state grantees.[61] Broadband availability is determined geographically. Population data are annually estimated at the census block level by GeoLytics.[62] These data are licensed by NTIA for NBM analysis. The NBM team used geospatial analysis to relate broadband availability data with census blocks and their attributes, determining populations and households with broadband availability.

The "NBM Census Blocks with Availability Rate Decreases by County" Excel file summarizes the data described above for every county included in the June 30, 2011 and June 30, 2013 NBM data provided to ASR. ASR joined the June 30, 2011 and June 30, 2013 data by census block FIPS code, calculated the availability rate in each release, and subtracted the 2011 rate from the 2013 rate. Any census block with a negative difference experienced a decrease in broadband availability. ASR created a series of indicators for all blocks representing any decrease in availability, availability decreases of 1 percentage point or more, and availability decreases of 5 percentage points or more. ASR calculated differences and created indicators for the NTIA and NOFA definitions of broadband. ASR then aggregated the indicator variables by county FIPS code and whether or not the census block was populated in the June 30, 2011 release. The aggregated data were saved to the "NBM Census Blocks with Availability Rate Decreases by County" Excel file.

3.27 NBM Statistics 2011-06-30 and 2013-06-30

NTIA provided NBM data, assembled by the NBM team at the FCC, to ASR for analysis in the *Final Report*. Data were provided for the following releases: June 30, 2011; December 31, 2011; June 30, 2012; December 31, 2012; and June 30, 2013.[63] Earlier releases were removed from consideration for analysis at NTIA's approval. Each release was provided at the census block level and contained:

- block FIPS codes;
- urban/rural flag;
- population;
- housing units;
- population with 768 kbps downstream and 200 kbps upstream service available (NOFA definition of broadband);
- housing units with 768 kbps downstream and 200 kbps upstream service available (NOFA definition of broadband);
- population with 3 Mbps downstream and 768 kbps upstream service available (NTIA definition of broadband);
- housing units with 3 Mbps downstream and 768 kbps upstream service available (NTIA definition of broadband);
- minority population;
- population over 60 years of age;
- population in poverty.

Broadband availability data are collected by individual state grantees.[64] Broadband availability is determined geographically. Population, housing unit, and demographic data are estimated annually at the census block level by GeoLytics.[65] These data are licensed by NTIA for NBM analysis. The NBM team uses geospatial analysis to relate broadband availability data with census blocks and their attributes, determining populations and households with broadband availability.

ASR developed county-level measures of broadband availability, population and housing units, and demographics based on the provided NBM data. ASR's analysis of NBM data showed that, in some census blocks in the contiguous United States, the broadband availability rate decreased from the June 30, 2011 release to the June 30, 2013 release. This occurred for both definitions of broadband and for population and household availability rates. ASR developed two data adjustments to ensure that broadband availability rates do not decrease from June 2011 to June 2013 at the census block level:

- Forward looking: When the availability rate in a census block is higher in 2011 than in 2013, the 2011 availability rate is lowered to equal that of 2013.

- Backward looking: When the availability rate in a census block is lower in 2013 than in 2011, the 2013 availability rate is raised to equal that of 2011.

Both of these adjustments reduce the number of census blocks with decreasing availability rates to zero.

3.27.1 NBM Statistics, 2011-06-30 and 2013-06-30 - Housing Units

The "NBM Statistics, 2011-06-30 and 2013-06-30 - Housing Units" Excel file contains county-level total housing units and housing units with availability according to the two different definitions of broadband (NOFA and NTIA) plus two adjusted versions (forward looking and backward looking) for each in the June 30, 2011 and June 30, 2013 NBM releases. These county-level totals were calculated by aggregating the block-level housing unit statistics provided to ASR by NTIA.

3.27.2 NBM Statistics, 2011-06-30 and 2013-06-30 - Minority Population

The "NBM Statistics, 2011-06-30 and 2013-06-30 - Minority Population" Excel file contains county-level total minority population and minority population with broadband availability according to the two different definitions of broadband (NOFA and NTIA) plus two adjusted versions (forward looking and backward looking) for each in the June 30, 2011 and June 30, 2013 NBM releases. County-level total minority population was calculated by aggregating block-level minority populations. County-level minority populations with broadband availability were calculated by multiplying the block-level minority percentage of total population by the total population with availability and aggregating the result to the county level.

3.27.3 NBM Statistics, 2011-06-30 and 2013-06-30 - Over 60 Population

The "NBM Statistics, 2011-06-30 and 2013-06-30 - Over 60 Population" Excel file contains county-level total population over sixty years of age and population over sixty years of age with availability according to the two different definitions of broadband (NOFA and NTIA) plus two adjusted versions (forward looking and backward looking) for each in the June 30, 2011 and June 30, 2013 NBM releases. County-level total population over sixty years of age was calculated by aggregating block-level populations over sixty years of age. County-level populations over sixty years of age with broadband availability were calculated by multiplying the block-level percentage of total population over sixty years of age by the total population with availability and aggregating the result to the county level.

3.27.4 NBM Statistics, 2011-06-30 and 2013-06-30 - Population

The "NBM Statistics, 2011-06-30 and 2013-06-30 - Population" Excel file contains county-level total population and population with availability according to the two different definitions of broadband (NOFA and NTIA) plus two adjusted versions (forward looking and backward looking) for each in the June 30, 2011 and June 30, 2013 NBM releases. These county-level totals were calculated by aggregating the block-level population statistics provided.

3.27.5 NBM Statistics, 2011-06-30 and 2013-06-30 - Poverty Population

The "NBM Statistics, 2011-06-30 and 2013-06-30 - Poverty Population" Excel file contains county-level total population in poverty and population in poverty with availability according to the two different definitions of broadband (NOFA and NTIA) plus two adjusted versions (forward looking

and backward looking) for each in the June 30, 2011 and June 30, 2013 NBM releases. County-level total population in poverty was calculated by aggregating block-level populations in poverty. County-level populations in poverty with broadband availability were calculated by multiplying the block-level percentage of total population in poverty by the total population with availability and aggregating the result to the county level.

3.27.6 NBM Statistics, 2011-06-30 and 2013-06-30 - Rural Population

The "NBM Statistics, 2011-06-30 and 2013-06-30 - Rural Population" Excel file contains county-level total rural population and rural population with availability according to the two different definitions of broadband (NOFA and NTIA) plus two adjusted versions (forward looking and backward looking) for each in the June 30, 2011 and June 30, 2013 NBM releases. County-level rural population and was calculated by aggregating block-level populations in poverty. These county-level totals were calculated by aggregating the block-level population statistics provided for only the blocks flagged as rural.

3.28 PCC and SBA Entrepreneurship Training

PCC and SBA grantees reported individual training activities across locations in quarterly PPRs. ASR queried all PPRs submitted by PCC and SBA grantees for the following variables: training program name, description, and the number of participants.[66] ASR subset the results to the last quarter each program was reported to present a cumulative result.[67] ASR then manually reviewed the title and descriptions of reported training activities to identify those programs specific to entrepreneurship. The "PCC and SBA Entrepreneurship Training" Excel file contains all identified entrepreneurship activities and the number of participants.

3.29 PCC and SBA Job Search Training

PCC and SBA grantees reported individual training activities across locations in quarterly PPRs. ASR queried all PPRs submitted by PCC and SBA grantees for the following variables: training program name, description, and the number of participants.[68] ASR subset the results to the last quarter each program was reported to present a cumulative result.[69] ASR then manually reviewed the title and descriptions of reported training activities to identify those programs specific to job search. The "PCC and SBA Job Search Training" Excel file contains all identified job search activities and the number of participants.

3.30 PCC and SBA Service Areas - Evaluation Study Sample

During the first round of case study site visits, ASR identified a service area for each of the PCC and SBA grants included in the evaluation study sample based on the geographic area described by the grantee in its grant application, fact sheet, and annual and quarterly PPRs.[70] In cases where a service area was unavailable, ASR conferred with grantees to determine each grant's service area. Grant representatives and NTIA program staff provided comments on the service areas during the case study report review process. Some service areas were defined at the county level, while others were defined at the census tract level. ASR used these grantee-approved service areas, defined in *Interim Report 1*, for all population and household statistics derived for the *Final Report*.[71] ASR has made no changes to service areas since *Interim Report 1*. The "PCC and SBA Service Areas - Evaluation Study Sample" Excel file lists the PCC and SBA service areas.

3.31 PCC and SBA Service Areas - Rest of BTOP

For grantees not included in the evaluation study sample, ASR created service areas using CAC data.[72] Every location reported in the CAC (2011 and 2012) was geocoded using ArcGIS to determine the county in which it is located. A grant's service area is composed of all counties where a location was reported in CAC. Several PCC grantees did not have any locations listed in the CAC. For each of these grantees, ASR reviewed publicly available project applications to determine county-level service areas.[73] The "PCC and SBA Service Areas - Rest of BTOP" Excel file contains all reported grant locations and service area counties.

3.32 PCC Annual PPR Data

ASR used the cumulative total established and cumulative improved PCCs by institution type from PCC APRs for further analysis in the *Final Report.*[74] APRs include the following institution types: Schools (K-12); Libraries; Community Colleges; Universities and Colleges; Medical or Healthcare Facilities; Public Safety Entities; Job Training and/or Economic Development Institutions; Other Community Support (Governmental); and Other Community Support (Non-Governmental). The "PCC Annual PPR Data" Excel file contains these data.

3.33 PCC Quarterly PPR Data

ASR used the following variables from PCC PPRs for further analysis in the *Final Report*: number of installed workstations, average weekly users, number of upgraded connections, number of newly established connections, and number of additional hours open.[75] The "PCC Quarterly PPR Data" Excel file contains these data.

3.34 PCC, SBA, and CCI Quarterly Budget Data

Grantees reported budget data quarterly in PPRs, including planned budget, actual expenditures from the current reporting period, and anticipated expenditures for the following quarter. Grantees reported these for both federal and matching funds, in addition to the total amounts. ASR retrieved these data for every grantee for every reported quarter.[76] The "PCC, SBA, and CCI Quarterly Budget Data" Excel file contains these data.

3.35 SBA Quarterly PPR Data

SBA grantees report the cumulative number of new household and business subscriptions due to grant activities in quarterly PPRs.[77] The "SBA Quarterly PPR Data" Excel file contain these data.

3.36 Service Areas of Grant Applications Not Funded by BTOP or BIP

The "Service Areas of Grant Applications Not Funded by BTOP or BIP" Excel file contains the counties that would have been impacted by last mile or middle mile BTOP or BIP grants had those grants been funded. For the purposes of this analysis, this includes any application that was either not funded or withdrawn.

NTIA's Broadband Application Database was the primary source of these data.[78] ASR used two types of attachments in the database:

1. Executive Summaries

2. Public Notice Responses

ASR first examined Executive Summaries to find clear delineations of proposed service areas. If necessary, the following helped to clarify the proposed funded service area:

1. Community anchor institutions
2. Tower site locations
3. Points of presence
4. Interconnection points
5. FTTP or FTTH locations
6. Communities served by proposed wireless broadband
7. Impacted tribal entities and state and federal American Indian reservations

When the funded service area could not be established from the Executive Summary, Public Notice Responses helped fill the gap. Public Notice Response information was included as part of the service area if a company claims that it already provides service in the proposed service area and the company provides a description of the territory served.

"Service Areas of Grant Applications Not Funded by BTOP or BIP" delineates service areas by county, but not all grantees defined their service area at the county level. ASR converted service areas originally defined at other levels to the county level based on the geographic locations of the stated service areas.

"Service Areas of Grant Applications Not Funded by BTOP or BIP" does not include unfunded applications in Alaska, Hawaii, and the outlying areas of the United States. Public safety network and satellite-based service grants are also excluded from the list. Grants without clearly specified service areas are included in "Service Areas of Grant Applications Not Funded by BTOP or BIP," but the service areas are not defined. Some applicants applied for more than one BTOP or BIP grant. Projects proposed by applicants with multiple grant applications not funded by BTOP or BIP were treated as separately. In these cases, each grant application was recorded separately even if the service area was identical to another application not funded by BTOP or BIP. Joint BIP/BTOP applicants were recorded only once and have "BIP/BTOP" in the "Application Type" column.

"Service Areas of Grant Applications Not Funded by BTOP or BIP" includes grant applications not funded by BTOP or BIP of applicants that received funding for a different project. Grant applications not funded by BTOP or BIP with slightly different service areas than their funded counterparts are also included. However, if the same grant application was not funded in Round One but funded in Round Two, then the grant is considered funded and is not included in "Service Areas of Grant Applications Not Funded by BTOP or BIP."

3.37 Training Activities - APR

Grantees reported cumulative training hours annually to NTIA in APRs. ASR retrieved all annually reported training data.[79] The results were subset to the last available APR submitted by each PCC and SBA grantee. Grantees reported training hours and participants for the following categories: Open Lab Access, Multimedia, Office Skills, ESL, GED, College Preparatory Training, Basic Internet and Computer Use, Certified Training Programs, and Other.

Grantees also reported descriptions for the Other category. ASR manually reviewed all Other descriptions and, where possible, mapped the reported hours and participants to one of the focus areas discussed in *Interim Report 1*, *Interim Report 2*, and the *Final Report*.[80] Two columns were added for each of the five focus areas: the first represents the percentage of Other hours falling into the focus area, while the second represents the percentage of Other participants falling into the

focus area. The columns were populated with percentages based on the descriptions provided in the APRs.[81]

The “Training Activities – APR” Excel file contains APR training data reported by grantees and the mapping percentages created by ASR.

3.38 Training Activities - PCC

PCC grantees reported individual training activities across locations in quarterly PPRs. ASR retrieved all available PCC quarterly training data.[82] ASR subset the results to the last available PPR submitted by each PCC grantee to present a cumulative result. ASR then manually reviewed all training data reported by grantees to determine if any of the training activities fell into the Healthcare or Quality of Life/Civic Engagement focus areas. ASR added a binary indicator variable for each of these focus areas, each denoting if a reported training activity falls into one of these focus areas. The “Training Activities - PCC" Excel file contains training program data, including hours, participants, and classification, for every training activity included in the last available PPR for every PCC grant.

3.39 Training Activities - SBA

SBA grantees reported individual training activities across locations in quarterly PPRs. ASR retrieved all available SBA quarterly training data.[83] ASR subset the results to the last available PPR submitted by each SBA grantee to present a cumulative result. ASR then manually reviewed all training data reported by grantees to determine if any of the training activities fell into the Healthcare or Quality of Life/Civic Engagement focus areas. ASR added a binary indicator variable for each of these focus areas, each denoting if a reported training activity falls into these focus areas. The “Training Activities - SBA" Excel file contains training program data, including hours, participants, and classification, for every training activity included in the last available PPR for every SBA grant.

3.40 Training Activities - SCTCS

Annual training data for the SC Reach for Success PCC grant were reported by category by the grantee, the South Carolina State Board for Technical and Comprehensive Education (SCTCS). ASR retrieved all quarterly training activities for the grant.[84] These data were combined into a single file. ASR then manually reviewed every reported training activity, classifying each into one of the five focus areas discussed in *Interim Report 1*.[85] Not all activities could be classified. Classifications are stored in a single variable containing the focus area name.

The “Training Activities - SCTCS" Excel file contains the categorized SCTCS data. ASR created a pivot table in this file to calculate the percentage of cumulative training hours and participants falling into each focus area. These percentages were inserted into the mapping fields in the “Training Activities - APR" Excel file, described in Section 3.37. The “Training Activities - SCTCS" Excel file includes the data necessary to derive the percentages shown in “Training Activities - APR.”

Section 4. Prepared Data

Prepared data files are generated from one or more input files. Every prepared data step includes the following:

- One R script to read one or more input files and perform all necessary manipulations. R scripts are sets of instructions written in the R statistical programming language than can be executed in the R statistical package.
- One Excel file with a worksheet for every table output by the R script. Excel files allow all users to inspect visually the generated tables and non-R users to access prepared data in their software of choice.
- One RDA file containing every table output by the R script. RDA files maintain the attributes of the output tables and can be quickly loaded when required.

Every prepared data R script has a variable that defines the parent directory and calls the shared source code scripts described in Section 2. Users are required to redefine the parent directory to the correct location on their computer to reproduce the manipulation steps and output. This can be done in batch using the "Set Parent Directory in All Script Files" R script described in Section 1.3.

The following subsections describe the R scripts, in alphabetical order, used to transform the input data described in the previous section. Lists of the variables in the prepared data produced in each of these steps are provided in the "File List and Descriptions" Excel file.

4.1 ACS 2010

The "ACS 2010" R script prepares ACS (2006-2010) data (ACS) published by the Census Bureau, described in Section 3.19, for further analysis.[86] The script performs the following operations:

- ACS data are read from the "Census Bureau ACS (2006-2010) Data" Excel file as a table

The above results in a county-level table named "acs10." The table is saved in the "ACS 2010" RDA file and "ACS 2010" Excel file.

4.2 Budgets

The "Budgets" R script prepares grantee-reported budget data, described in Section 3.34, for further analysis. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

The script performs the following operations:

- The "Grants" RDA file is loaded
- Budget data are read from the "PCC, SBA, and CCI Quarterly Budget Data" Excel file
- Total federal, matching, and combined budget variables are extracted
- NTIA assigned three grants new award numbers during the award period. New award numbers are reverted to maintain consistency:

- Centennial Board of Cooperative Educational Services, Colorado Community Anchors, Broadband Consortium-Connecting Colorado's Middle Mile: NT10BIX5570156 reverts to NT11BIX5570001
- Zayo Bandwidth, LLC, Indiana Middle Mile Fiber for Schools, Communities, and Anchor Institutions: NT12BIX5570001 reverts to NT10BIX5570025
- Zayo Bandwidth, LLC, Connect Anoka County Community Broadband Network: NT12BIX5570002 reverts to NT10BIX5570071

- Grantees stop submitting PPRs as their award periods end. The last available PPR is determined and budget data are filtered to the last available PPR.

The above results in a table named "budget" saved to the "Budgets" RDA file and "Budgets" Excel file. The table is joined with the "Grants" table by award number to include grant information during the Excel write process.

4.3 CCI CAIs by Category

The "CCI CAIs by Category" R script summarizes the results of ASR's CAI classification process as a formatted table. This classification process is described in Section 3.12. The script uses the following:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

The script performs the following operations:

- The "Grants" RDA file is loaded
- Classification results are read from the "CCI Categorized Connected Locations" Excel file
- The categories variable is formatted for consistency with other prepared data
- The results are aggregated by counting the number of institution types for each grant

The above results in a table named **"**connected_by_category.**"** The table is saved in the "CCI CAIs by Category" RDA file and the "CCI CAIs by Category" Excel file. The table is joined with the "Grants" table by award number to include grant information during the Excel write process.

4.4 CCI CAIs Passed

The "CCI CAIs Passed" R script estimates the cumulative number of CAIs connected or passed by CCI grantees by grant, year, and CAI category. Estimates are derived from APR data, described in Section 3.12.[87] The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

The script performs the following operations:

- The "Grants" RDA file is loaded
- APR data are read from the "CCI Annual PPR Data" Excel file as a table
- APR data are reshaped so that each row represents one CAI category for one year for one grant
- Institution type is formatted for consistency with other prepared data

- Grantees stop submitting APRs as their award periods end. Data for concluded grants must be carried over into later years to more accurately represent cumulative BTOP activities. The script uses the following process to carry over values from the last reported APR to 2013, when necessary:
  - Year is formatted as a date
  - The first year with reported values (nonzero) is determined; if the grantee never reported values greater than zero, the year of the first available APR is used
  - The last available APR is determined
  - The Cartesian product of all CCI award numbers, all CAI categories, and all years with at least one APR is created
  - A rolled join is performed; APR data and the Cartesian product table are joined by award number, CAI category, and date, and the last available value for every other variable is carried forward for each award number and CAI category combination when a date is not present in the APR data
  - A Boolean indicator identifying years with carried-over values is created
  - Observations in the rolled table for years earlier than the first reported year or later than 2013 are filtered out

The above results in a table named "pass." The table is saved in the "CCI CAIs Passed" RDA file and "CCI CAIs Passed" Excel file. The table is joined with the "Grants" table by award number to include grant information during the Excel write process.

4.5 CCI Evaluation Study Sample CAIs

The "CCI Evaluation Study Sample CAIs" R script prepares data on the total number of connected CAIs for evaluation study grantees, described in Section 3.14, for further use in the intermediate-term analysis. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

The script performs the following operations:

- The "Grants" RDA file is loaded
- Total connected CAIs included in the case study reports data are read from the "CCI Connected CAIs in Case Study Reports" Excel file as a table
- Year and quarter number values are used to create a date format variable

The above results in a table named "conn_cais." The table is saved in the "CCI Case Study CAIs" RDA file and the "CCI Case Study CAIs" Excel file. The table is joined with the "Grants" table to include grant information during the Excel write process.

4.6 CCI Progress

The "CCI Progress" R script estimates quarterly network, quarterly CAI, total network, and total CAI statistics for BTOP CCI grants. These estimates are derived from PPRs, described in Section 3.16, and findings from CCI case study reports, described in Section 3.14.[88] The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

- "CCI Evaluation Study Sample CAIs" contains a table of total connected CAIs included in case study reports for the evaluation study grants. Section 4.4 provides a full description of "CCI Case Study CAIs."

The script performs the following operations:

- The "Grants" and "CCI Evaluation Study Sample CAIs" RDA files are loaded
- PPR data are read from the "CCI Quarterly PPR Data" Excel file as a table
- NTIA assigned three grants new award numbers during the award period. New award numbers are reverted to maintain consistency:
  - Centennial Board of Cooperative Educational Services, Colorado Community Anchors, Broadband Consortium-Connecting Colorado's Middle Mile: NT10BIX5570156 reverts to NT11BIX5570001
  - Zayo Bandwidth, LLC, Indiana Middle Mile Fiber for Schools, Communities, and Anchor Institutions: NT12BIX5570001 reverts to NT10BIX5570025
  - Zayo Bandwidth, LLC, Connect Anoka County Community Broadband Network: NT12BIX5570002 reverts to NT10BIX5570071
- Grantees stop submitting PPRs as their award periods end. Data for concluded grants must be carried over into later quarters to more accurately represent cumulative BTOP activities. The script uses the following process to carry over values from the last reported PPR to the end of 2013, when necessary:
  - Year and quarter number are formatted as a date
  - The first quarter with reported values (nonzero) is determined; if the grantee never reported values greater than zero, the quarter of the first available PPR is used
  - The last available PPR is determined
  - The Cartesian product of all CCI award numbers and all dates with at least one PPR is created
  - A rolled join is performed; PPR data and the Cartesian product table are joined by award number and date, and the last available value for every other variable is carried forward for each award number when a date is not present in the PPR data
  - A Boolean indicator identifying quarters with carried-over values is created
  - Observations in the rolled table for quarters earlier than the first reported quarter are filtered out
  - Data from NTIA includes one 2014 Q1 PPR. These data should not be used; all values for 2014 Q1 are filtered out.
- CAI data are extracted from the PPR data. Network data are extracted to a separate table.
- In the CAI data, CAI totals for the last quarter included in case study reports are updated to use the totals from the case studies. Data for later quarters are differenced, then added to the case study report totals.
- Overall totals (cumulative totals at the end of 2013) are extracted from the CAI data and the network data to create two totals tables

The above results in four tables:

1. Network Progress ("network_progress"): quarterly cumulative network progress (new miles, upgraded miles, leased new miles, leased existing miles, interconnection points, signed agreements)
2. CAI Progress ("cai_progress"): quarterly cumulative CAIs connected
3. Network Totals ("network_totals"): total network activity (new miles, upgraded miles, leased new miles, leased existing miles, interconnection points, signed agreements)
4. CAI Totals ("cai_totals"): total CAIs connected

These tables are saved in the "CCI Progress" RDA file, with object names corresponding to those given in parenthesis above. The tables are also written, with sheet names corresponding the descriptive names above, to the "CCI Progress" Excel file as separate sheets. Tables are joined with the grants table by award number to include grant information during the Excel write process.

4.7 CCI Speed and Pricing

The "CCI Speed and Pricing" R script prepares data on connected CAI subscription speeds and costs collected by ASR, described in Section 3.15, for further analysis. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

The script performs the following operations:

- The "Grants" RDA file is loaded
- Subscription speed and pricing data are read from the "CCI Case Study Connected CAI Collected Speed and Pricing Data" Excel file as a table
- Variables are given shorter names consistent with the naming conventions used in other data
- Undefined values are replaced with missing values. The values were not truly undefined in the "CCI Case Study Connected CAI Collected Speed and Pricing Data" Excel file, only missing.
- The CAI category variable is formatted for consistency with other prepared data

The above results in a table named "price_speed." The table is saved in the "CCI Speed and Pricing" RDA file and the "CCI Speed and Pricing" Excel file. The table is joined with the "Grants" table by award number to include grant information during the Excel write process.

4.8 Grants

The "Grants" R script prepares the table of BTOP grants extracted from PPRs, described in Section 3.25, for use in intermediate-term analysis.[89] The script performs the following operations:

- Data are read from the "List of Awarded BTOP Grants" Excel file as a table
- Variables are formatted to improve the table's usability

The resulting table "grants" is saved in the "Grants" RDA file and "Grants" Excel file.

4.9 Gross County Product

The "Gross County Product" R script estimates county-level GDP, or Gross County Product (GCP), using county personal income (LAPI), state-level GDP (GSP), county business patterns (CBP), and state business patterns (SBP) data described in Sections 3.3, 3.5, 3.20, and 3.22, respectively.[90]

The state is the lowest geographic level for which comprehensive GDP data are available.[91] ASR required county-level approximations of GDP to extrapolate benefits for counties served by BTOP CCI grants. To account for expected economic differences between counties served by BTOP and unserved areas of states, ASR decomposed state-level GDP using publicly available county-level economic statistics. Bauer and Lee (2006) and Barreca, Fannin, and Detre (2012) offer methods of estimating GDP at lower-than-published levels.[92] ASR followed these methods by apportioning state-level GDP to counties according to the distribution of personal income. BEA published personal income statistics for counties across the United States. For counties that were combined

in the personal income data, ASR used the distribution of private nonfarm employment according to CBP to apportion personal income across the combined counties.

The "Gross County Product" R script performs the following operations:

- GSP data are read from the "BEA Gross State Product - 2011" Excel file as a table
- SBP data are read from the "Census Bureau State Business Patterns - 2011" Excel file as a table
- CBP data are read from the "Census Bureau County Business Patterns - 2011" Excel file as a table
- In the CBP data, unreported employment figures are interpolated:
  - CBP employment figures are aggregated to the state level
  - SBP employment figures are joined to the CBP data by state FIPS code
  - The difference between the aggregated CBP figures and the SBP figures is the number of jobs that can be allocated across unreported counties in a state
  - CBP establishment figures are aggregated across unreported counties in a state. The percentage of establishments in each unreported county is then calculated. This is used as the distribution of unreported employment figures.
  - The percentage of establishments in an unreported county is multiplied by the number of jobs that can be allocated across unreported counties in a state
- LAPI data are read from the "BEA County Personal Income - 2011" Excel file as a table
- In the LAPI data, reported figures for combined areas are split into individual counties and equivalents:
  - Combined area FIPS codes in the LAPI data are mapped to county FIPS codes
  - For the counties and equivalents in combined areas, employment figures are joined on CBP data by county FIPS code and the distribution of employment across the combined area is calculated
  - LAPI figures are apportioned to individual counties and equivalents in combined areas according to the distribution of employment
  - Combined areas are filtered out of LAPI data
- In the LAPI data, the expected distribution of GDP in a state is calculated as county personal income divided by the sum of personal income for all counties in the state
- LAPI data and GSP data are joined by state FIPS code
- Gross county product is estimated as the percentage of personal income in a county times state GDP

The above results in a county-level table named "gcp." The table is saved in the "Gross County Product" RDA file and "Gross County Product" Excel file.

4.10 Household Statistics

The "Household Statistics" R script combines ACS (2006-2010) household income data (ACS) and NBM total households data (NBM) described in Sections 3.19 and 3.27, respectively.[93] The script uses the following prepared data:

- "ACS 2010" contains a table of county-level statistics from ACS (2006-2010), including average household income. Section 4.1 provides a full description of "ACS 2010."

The script performs the following operations:

- The “ACS 2010” RDA file is loaded
- NBM data are read from the “NBM Statistics, 2011-06-30 and 2013-06-30 - Housing Units” Excel file as a table and filtered to the June 30, 2011
- NBM and ACS data are joined by county FIPS code
- In the joined table, areas outside the fifty states plus Washington, DC and Puerto Rico, such as the Virgin Islands, are missing household income data. All missing values of household income are replaced by the median household income for all non-missing values.

The above results in a county-level table named “households.” The table is saved in the “Household Statistics” RDA file and “Household Statistics” Excel file.

4.11 Labor Force Statistics

The “Labor Force Statistics” R script combines BLS QCEW and LAUS data described in Sections 3.6 and 3.7, respectively.[94] The script performs the following operations:

- LAUS data are read from the “BLS Local Area Unemployment Statistics - 2011” Excel file as a table
- QCEW data are read from the “BLS Census of Employment and Wages - 2011” Excel file as a table
- LAUS and QCEW data are joined by county FIPS code

The above results in a county-level table named “labor_force.” The table is saved in the “Labor Force Statistics” RDA file and “Labor Force Statistics” Excel file.

4.12 NBM Population Statistics

The “NBM Population Statistics” R script manipulates NBM population, demographic, and broadband availability data. The NBM data represent county-level total population and county-level minority, over-sixty, in-poverty, and rural populations, all described in Section 3.27.[95] The script performs the following operations:

- Total population data are read from the “NBM Statistics, 2011-06-30 and 2013-06-30 - Population” Excel file as a table
- In the total population data, missing values for the adjusted versions of broadband availability are replaced with the corresponding values of unadjusted broadband availability. Missing values are present in the input data because each adjustment (forward looking and backward looking) applies to only one release. Therefore, adjusted data must be related to unadjusted data to examine changes over time. In practice, this is equivalent to replacing missing values of adjusted data with unadjusted values. This method also allows the procedures used on unadjusted data to be used, without any modifications, on adjusted data.
- Rural (“NBM Statistics, 2011-06-30 and 2013-06-30 - Rural Population”), minority (“NBM Statistics, 2011-06-30 and 2013-06-30 - Minority Population”), over-sixty (“NBM Statistics, 2011-06-30 and 2013-06-30 - Over 60 Population”), and in-poverty (“NBM Statistics, 2011-06-30 and 2013-06-30 - Poverty Population”) data are read from their respective Excel files into a list of tables
- Populations are extracted from each of the tables in the list of rural, minority, over-sixty, and in-poverty tables. The extracted populations are joined by county FIPS code, release year, and release month. This results in a single vulnerable populations table.

- A copy of the total population table is created. In this table, populations with broadband availability are converted to availability rates by dividing the availability populations by total population.
- A copy of the vulnerable populations table is created. Total population is then joined to this table by county FIPS code, release year, and release month. In this table, vulnerable populations are converted to vulnerable percentages of population by dividing the vulnerable populations by total population.
- The total population, availability rate, vulnerable population, and vulnerable percentage of population tables are reshaped from wide and short to narrow and long and concatenated. The concatenated table is then reshaped from narrow and long to wide and short so that each row represents a unique county and each column represents the value of one variable in one release (e.g., total population in the June 30, 2011 release, total population in the June 30, 2013 release, and so on).

The above results in a county-level table named "availability." The table is saved in the "NBM Population Statistics" RDA file and "NBM Population Statistics" Excel file.

4.13 PCC and SBA Average Household Size

The "PCC and SBA Average Household Size" R script estimates the average household size in the service area of each PCC and SBA grant. Estimates are derived from data published by the Census Bureau (ACS), described in Section 3.18.[96] The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "PCC and SBA Service Areas" contains tables that list grant service areas. One table contains service areas defined at the county level. A second table contains service areas defined at the census-tract level. Section 4.17 provides a full description of "PCC and SBA Service Areas."

The script performs the following operations:

- The "Grants" and "PCC and SBA Service Areas" RDA files are loaded
- ACS data are read from the "Census Bureau ACS (2005-2009) Data" CSV file as a table
- ACS data are joined to census tract service areas by census tract FIPS code, and aggregated to the county level and joined to county service areas by county FIPS code
- The joined census tract service areas and county service areas tables are concatenated
- Population and occupied housing units are aggregated to the grant level
- Average household size in the grant service area is estimated as total population divided by total occupied housing units
- ACS data are not available for the Virgin Islands. The national average household size is used for the average household size in the service area for grants that serve the Virgin Islands.[97]

The above results in a table named "avg_hh_size." The table is saved in the "PCC and SBA Average Household Size" RDA file and "PCC and SBA Average Household Size" Excel file. The table is joined with the "Grants" table by award number to include grant information during the Excel write process.

4.14 PCC and SBA Entrepreneurship Training Summary

The "PCC and SBA Entrepreneurship Training Summary" R script summarizes the number of participants in grantee-provided entrepreneurship training programs, described in Section 3.28, for further analysis.[98] The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

The script performs the following operations:

- The "Grants" RDA file is loaded
- Data are read from the "PCC and SBA Entrepreneurship Training" Excel file as a table
- Reported participants area aggregated for each grant

The above results in a table named "entrepreneurship." The table is saved in the "PCC and SBA Entrepreneurship Training Summary" RDA file and "PCC and SBA Entrepreneurship Training Summary" Excel file. The table is joined with the "Grants" table by award number to include grant information during the Excel write process.

4.15 PCC and SBA Job Search Training Summary

The "PCC and SBA Job Search Training Summary" R script summarizes the number of participants in grantee-provided job search training programs, described in Section 3.29, for further analysis.[99] The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

The script performs the following operations:

- The "Grants" RDA file is loaded
- Data are read from the "PCC and SBA Job Search Training" Excel file as a table
- Reported participants area aggregated for each grant

The above results in a table named "job_search." The table is saved in the "PCC and SBA Job Search Training Summary" RDA file and "PCC and SBA Job Search Training Summary" Excel file. The table is joined with the "Grants" table by award number to include grant information during the Excel write process.

4.16 PCC and SBA Labor Force Statistics

The "PCC and SBA Labor Force Statistics" R script estimates certain labor force statistics for the service area of each PCC and SBA grant. Estimates are derived from data published by BLS (LAUS), the Census Bureau (ACS), and DOL (Minimum Wage), described in Sections, 3.7, 3.18, and 3.23, respectively, and data jointly published by DOL and BLS (Duration), described in Section 3.8.[100] The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "PCC and SBA Service Areas" contains tables that list grant service areas. One table contains service areas defined at the county level. A second table contains service areas defined at the

census-tract level. A third table contains the service area for all grants, but redefines any census-tract-level service areas at the county level. A fourth table contains the service area for all grants, but redefines all counties and census tracts to the state level. Section 4.17 provides a full description of "PCC and SBA Service Areas."

The script performs the following operations:

- The "Grants" and "PCC and SBA Service Areas" RDA files are loaded
- ACS data are read from the "Census Bureau ACS (2005-2009) Data" CSV file as a table
- LAUS data are read from the "BLS Local Area Unemployment Statistics - 2009" Excel file as a table
- Minimum Wage data are read from the "DOL National and State Minimum Wages - 2009" Excel file as a table. The national minimum wage is extracted to a new variable. Any state minimum wages lower than the national minimum wage are replaced by the national minimum wage.
- Duration data are read from the "BLS-DOL Unemployment Duration - 2009" Excel file as a table
- ACS population data are aggregated to the county level and joined to the LAUS table by county FIPS code
- LAUS data are joined to defined county-level service areas by county FIPS code
- ACS data are joined to defined census-tract-level service areas by census tract FIPS code
- LAUS and ACS data are concatenated
- County weights are calculated as the size of the labor force in each county divided by the total labor force in the service area for each grant using the concatenated LAUS and ACS data
- State weights are calculated as the size of the labor force in each state divided by the total labor force in the service area for each grant using the concatenated LAUS and ACS data
- Minimum Wage data are joined with state-level service areas for all grants by state FIPS code
- Duration data are joined with state-level service areas for all grants by state FIPS code
- Data are aggregated to the grant level:
  - LAUS and ACS data are counts of individuals, such as population and employed persons. These data are summed by grant.
  - Minimum Wage data are averages. State weights are joined with Minimum Wage data by state FIPS code and weighted averages are calculated by grant.[101]
  - Duration data are averages. State weights are joined with Duration data by state FIPS code and weighted averages are calculated by grant.[102]
- Grant-level LAUS and ACS, Minimum Wage, and Duration data are joined by award number
- Rates needed in further analysis are added to the table:
  - labor force participation rate = labor force ÷ population
  - unemployment rate = total unemployment ÷ total labor force
- LAUS/ACS and Duration data are not available for the Virgin Islands. National labor force participation rate, unemployment rate, and average unemployment duration are used for the service area figures for grants that serve the Virgin Islands.[103]

The above results in a table named "labor_force." The table is saved in the "PCC and SBA Labor Force Statistics" RDA file and "PCC and SBA Labor Force Statistics" Excel file. The table is joined with the "Grants" table by award number to include grant information during the Excel write process.

4.17 PCC and SBA Service Areas

The "PCC and SBA Service Areas" R script combines service areas described in the Round 1 case study reports, described in Section 3.30, with service areas determined by geolocating CAC locations for PCC and SBA grants not in the evaluation study sample, described in Section 3.31.[104] The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

The script performs the following operations:

- The "Grants" RDA file is loaded
- Evaluation study sample data are read from the "PCC and SBA Service Areas - Evaluation Study Sample" Excel file
- Data for grants outside of the evaluation study sample are read from the "PCC and SBA Service Areas - Rest of BTOP" Excel file
- In-sample and out-of-sample data are concatenated
- Four additional tables containing service areas at different geographic levels are created for convenience: county-level (census tracts are converted to unique counties); state-level (counties and census tracts are converted to unique states); county level only (census tracts are dropped); and census-tract level only (counties are dropped).

The above results in five tables:

1. All Areas ("service.areas"): service areas for all grants, defined at either the county or census-tract level
2. All at County Level ("sa.county.level"): service areas for all grants, redefined at the county level
3. All at State Level ("sa.state.level"): service areas for all grants, redefined at the state level
4. County Areas ("county.areas"): service areas for only the grants explicitly defined at the county level
5. Tract Areas ("tract.areas"): service areas for only the grants explicitly defined at the census-tract level

These tables are saved in the "PCC and SBA Service Areas" RDA file, with table names corresponding to those given in parenthesis above. The tables are also written, with sheet names corresponding the descriptive names above, to the "PCC and SBA Service Areas" Excel file as separate sheets. Tables are joined with the grants table by award number to include grant information during the Excel write process.

4.18 PCC and SBA Training Hours and Participants

The "PCC and SBA Training Hours and Participants" R script summarizes the number of participants and training hours for each grant for each focus area. The script uses data reported in grant APRs and PPRs, described in Sections 3.37, 3.38, and 3.39.[105]

In *Interim Report 1*, ASR defined focus areas to describe grantee activities, outcomes, and impacts.[106] ASR mapped the annual training data to focus areas in order to estimate the number of training hours that fall into each focus area. Table A summarizes the reporting categories available to grantees and the focus areas to which they belong.

Table A. Training Hour Categorization

Annual Training Data	Focus Area
Basic Internet and Computer Use	Digital Literacy
Certification Programs	Education and Training
College Preparatory Training	Education and Training
ESL	Education and Training
GED	Education and Training
Multimedia	Digital Literacy
Office Skills	Workforce and Economic Development

As described in Sections 3.37, ASR reviewed all annual training data reported by grantees in the Other category to determine if any of the reported hours and participants could be mapped into one of the five focus areas. As described in Sections 3.38 and 3.39, ASR also reviewed the training activities reported in the last available PPR for each grantee and determined if any of these activities could be classified as Healthcare or Quality of Life/Civic Engagement. The process below is identical to those used for *Interim Report 1* and Round 2 PCC and SBA case study reports.[107]

The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

The "PCC and SBA Training Hours and Participants" R script performs the following operations:

- The "Grants" RDA file is loaded
- APR data are read from the "Training Activities - APR" Excel file as a table
- PCC PPR data are read from the "Training Activities - PCC" Excel file as a table
- SBA PPR data are read from the "Training Activities - SBA" Excel file as a table
- In the APR data, Other training hours and participants are multiplied by the focus area percentages determined during manual review. The results are added to focus area columns and subtracted from the Other columns.
- APR data are reshaped and reporting categories are mapped to focus areas using the categorization shown in Table A
- In the SBA PPR data, training hours per activity are estimated as seven times the number of participants, the average number of training hours per participant for PCC grants at the time of *Interim Report 1*.
- PCC PPR and SBA PPR data are reshaped, concatenated, and aggregated
- APR and PPR data are merged
- Adjustments to APR and PPR data were made. ASR estimated focus area training hours without these adjustments and found several negative training totals. The underlying data were examined and several issues were identified. The following adjustments address the identified issues:
  - University of California, Davis, California Telehealth: all hours reported in the APR were mapped to healthcare; PPR totals are therefore set to zero
  - City and County of San Francisco, San Francisco Community Broadband Opportunities Program: website hits were reported; these are removed since they are not training activities

- Georgia Partnership for TeleHealth, Inc., TeleConnect: all hours reported in the APR were mapped to healthcare; PPR totals are therefore set to zero
- Coppin State University, Coppin Heights-Rosemont: training hours and activities were not reported correctly in the APR; APR totals are set to zero and all reported PPR activities are classified into one of the five focus areas
- Wildwood Programs, Inc., Broadband Video for Human: no hours are mapped to Digital Literacy; Healthcare hours and participants in excess of remaining Other hours and participants will be removed from Workforce and Economic Development totals instead of Digital Literacy in the step below

- APR and PPR combination rules are applied:
  - If APR totals (hours or participants) mapped to the Healthcare or Quality of Life/Civic Engagement focus areas are greater than PPR totals, APR totals are used
  - If APR totals mapped to the Healthcare or Quality of Life/Civic Engagement focus areas are lower than PPR totals, PPR totals are used. The differences in the PPR totals and APR totals are removed from remaining Other totals. If the remaining Other totals do not have sufficient hours or participants, the remainders are removed from Digital Literacy totals.
- Extraneous columns are dropped and data are aggregated to the grant by focus area level
- Open Lab Access hours and participants were not included in this process. They are added back to the data as a separate "focus area" by again reading the "Training Activities - APR" Excel file as a table and merging the Open Lab Access data to the focus area totals.

The above results in a table named "training.lab." This table is saved in the "PCC and SBA Training Hours and Participants" RDA file and "PCC and SBA Training Hours and Participants" Excel file. The table is joined with the "Grants" table by award number to include grant information during the Excel write process.

4.19 PCC Centers Established and Improved

The "PCC Centers Established and Improved" R script estimates the cumulative number of PCCs established and improved by grant, year, and type of PCC. Estimates are derived from APR data, described in Section 3.32.[108] The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

The script performs the following operations:

- The "Grants" RDA file is loaded
- APR data are read from the "PCC Annual PPR Data" Excel file as a table
- APR data are reshaped so that each row represents one institution type for one year for one grant
- Institution type is formatted for consistency with other prepared data
- Grantees stop submitting APRs as their award periods end. Data for concluded grants must be carried over into later years to more accurately represent cumulative BTOP activities. The script uses the following process to carry over values from the last reported APR to 2013, when necessary:
  - Year is formatted as a date
  - The first year with reported values (nonzero) is determined; if the grantee never reported values greater than zero, the year of the first available APR is used
  - The last available APR is determined

- The Cartesian product of all PCC award numbers, all institution types, and all years with at least one APR is created
- A rolled join is performed; APR data and the Cartesian product table are joined by award number, institution type, and date, and the last available value for every other variable is carried forward for each award number and institution type combination when a date is not present in the APR data
- A Boolean indicator identifying years with carried-over values is created
- Observations in the rolled table for years earlier than the first reported year are filtered out

The above results in a table named "est_impr." The table is saved in the "PCC Centers Established and Improved" RDA file and "PCC Centers Established and Improved" Excel file. The table is joined with the "Grants" table by award number to include grant information during the Excel write process.

4.20 PCC Equipment

The "PCC Equipment" R script estimates cumulative hardware installations and upgrades, and additional operating hours by quarter. Estimates are derived from PPR data, described in Section 3.33.[109] The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

The script performs the following operations:

- The "Grants" RDA file is loaded
- PPR data are read from the "PCC Quarterly PPR Data" Excel file as a table
- Some PPR combinations (grant, year, and quarter) appear more than once in the data. Data values are identical for all duplicated PPRs. Extraneous rows are removed from the PPR data.
- Grantees stop submitting PPRs as their award periods end. Data for concluded grants must be carried over into later quarters to more accurately represent cumulative BTOP activities. The script uses the following process to carry over values from the last reported PPR to the end of 2013, where necessary:
  - Year and quarter number are formatted as a date
  - The first quarter with reported values (nonzero) is determined; if the grantee never reported values greater than zero, the quarter of the first available PPR is used
  - The last available PPR is determined
  - The Cartesian product of all PCC award numbers and all quarters with at least one PPR is created
  - A rolled join is performed; PPR data and the Cartesian product table are joined by award number and date, and the last available value for every other variable is carried forward for each award number when a date is not present in the PPR data
  - A Boolean indicator identifying quarters with carried-over values is created
  - Observations in the rolled table for quarters earlier than the first reported quarter are filtered out

The above results in a table named "inst_upgr." The table is saved in the "PCC Equipment" RDA file and "PCC Equipment" Excel file. The table is joined with the "Grants" table to include grant information during the Excel write process.

4.21 PCC Sites

The "PCC Sites" R script prepares PCC location data published by NTIA in the CAC, described in Section 3.9, for further analysis.[110] The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

The script performs the following operations:

- The "Grants" RDA file is loaded
- CAC data are read from the "PCC Sites Data" sheet in the "BTOP Map Data - 2012" Excel file as a table
- The following variables are selected: award number, institution type, PCC type, number of workstations, hours open during workweek, hours open during weekend, connection speed, and average weekly users
- The institution type variable is cleaned and formatted, reducing the number of unique values
- The PCC type variable is cleaned and formatted, reducing the number of unique values

The above results in a table named "pcc_sites." The table is saved in the "PCC Sites" RDA file and "PCC Sites" Excel file. The table is joined with the "Grants" table by award number to include grant information during the Excel write process.

4.22 PCC Unique Users

The "PCC Unique Users" R script estimates the total participants and unique users of PCCs for the award period of PCC grants. Estimates are derived from the CAC, described in Section 3.9; PPRs, described in Section 3.33; and public computer usage rates for public libraries in the United States.[111]

Grantees reported average weekly users across all locations in PPRs.[112] In the *Final Report*, ASR analyzed the benefits to individuals in the labor force who used BTOP-funded PCCs. Users in the labor force are only expected at certain location types. To determine the percentage of average weekly users that visited these locations, ASR used CAC data to determine the grant-wide percentage of users that visit the following types of locations:[113]

- Community Based Organization
- Library
- Non-Profit Organization
- Public Housing
- Tribal

ASR assumes that usage rates among these location types are consistent with usage rates for public library public computers. ASR uses the grant-wide percentage of users of the applicable location types listed above as the percentage of PPR average weekly users in the labor force.

Weekly user figures are not representative of unique users, as an individual may visit a PCC multiple times. A 2010 study on public computer use at United States public libraries surveyed users to determine the frequency at which they used library computers.[114] Becker et al. (2010) reported that 23 percent of users of library computers used a library computer "every day or most days," 24 percent used "at least once a week," 20 percent used "about one to three times a month," 20 percent used "several times a year," and 13 percent used "about once a year."[115] ASR assumed

that these user groups visited daily, weekly, monthly, quarterly, and yearly (five times per week, once per week, once per month, once per quarter, and once per year, respectively). Total users over the grant period is given by:

$$\text{total users} = \text{years} \times [(5 \times 52 \times \text{daily users}] + (52 \times \text{weekly users}) + (12 \times \text{monthly users}) + (4 \times \text{quarterly users}) + \text{yearly users}]$$

Under this framework, the number of unique users is the sum of daily, weekly, monthly, quarterly, and yearly users. Substituting user groups for the percentages listed above and solving for unique users yields:

$$\text{unique users} = \text{total users} \div (75.61 \times \text{years})$$

The "PCC Unique Users" R script applies the above methodology to the above data. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

The script performs the following operations:

- The "Grants" RDA file is loaded
- The 75.61 per-year factor, derived above, is calculated and stored
- CAC data are read from the "BTOP Map Data - 2011" and "BTOP Map Data - 2012" Excel files as tables and concatenated
- In the CAC data, the percentage of total reported users of applicable location types, described above, is calculated for each grant
- PPR data are read from the "PCC Quarterly PPR Data" Excel file as a table. The data contain duplicate quarters for some grants, but the reported user figures are the same. Duplicate PPRs are removed. Quarters with no reported users are filtered out.
- CAC percentages and PPR data are joined by award number. The average number of weekly users of the applicable location types is estimated as the CAC applicable location percentage times PPR average weekly users. When a grant does not have a CAC percentage, 100 percent is used.
- Average weekly users is converted to total quarterly users by multiplying the average by 13, the number of weeks in a quarter
- Quarter begin and end dates (in the form of YYYY-MM-DD) are determined from the year and quarter number
- User data are aggregated to the grant level by summing total quarterly users, finding the minimum (earliest) quarter begin date, and finding the maximum (latest) quarter end date
- In the aggregated data, the number of years is calculated as the number of years (number of days ÷ 365) from the minimum quarter begin date to the maximum quarter end date
- In the aggregated data, unique users is estimated as:

$$\text{unique users} = \text{total users} \div (\text{user factor} \times \text{calculated number of years})$$

The above results in a table named "users." The table is saved in the "PCC Unique Users" RDA file and "PCC Unique Users" Excel file. The table is joined with the grants table by award number to include grant information during the Excel write process.

4.23 Rates of Adoption by Households with Availability

The "Rates of Adoption by Households with Availability" R script estimates county-level rates of adoption by households with availability using FCC Form 477 household broadband adoption data (FCC) and NBM household broadband availability data (NBM) described in Sections 3.24 and 3.27, respectively.[116] The script performs the following operations:

- NBM data are read from the "NBM Statistics, 2011-06-30 and 2013-06-30 - Housing Units" Excel file as a table
- FCC data are read from the "FCC Form 477 County Household Adoption Levels - 2011-06-30" Excel file as a table
- NBM data are filtered to the June 30, 2011 release and the county FIPS codes, total households, and households with NOFA broadband availability, forward-looking availability, and backward-looking availability are selected
- NBM availability rates are calculated by dividing the broadband housing unit counts by the total number of housing units
- FCC data are filtered to the NOFA definition of broadband and the county FIPS codes and adoption levels are selected
- In the FCC data, the midpoints of the adoption levels are calculated. For example, the adoption level "0.8 – 1.0" has a midpoint of 0.9.
- NBM and FCC data are joined by county FIPS code
- County-level rates of adoption are estimated by dividing the midpoint of the FCC 477 household adoption level by the household availability, forward-looking availability, and backward-looking availability rates. Any rates of adoption greater than one are set to one.

The above results in a county-level table named "adoption." The table is saved in the "Rates of Adoption by Households with Availability" RDA file and "Rates of Adoption by Households with Availability" Excel file.

4.24 SBA Locations

The "SBA Locations" R script prepares SBA program data published by NTIA in the CAC, described in Section 3.9, for further analysis.[117] The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

The script performs the following operations:

- The "Grants" RDA file is loaded
- CAC data are read from the "SBA Programs Data" sheet in the "BTOP Map Data - 2012" Excel file as a table
- The following variables are selected: award number, site name, institution type, and organization type
- Some institutions have the incorrect institution type value "CAI" in the published data. The correct institution type values are erroneously stored in the organization type field. These institution type values are corrected using the organization type values. The organization type field is dropped.
- The institution type variable is cleaned and formatted, reducing the number of unique values

The above results in a table named "sba_locs." The table is saved in the "SBA Locations" RDA file and "SBA Locations" Excel file. The table is joined with the "Grants" table by award number to include grant information during the Excel write process.

4.25 SBA Subscribers

The "SBA Subscribers" R script estimates quarterly, cumulative by quarter, and total new household, individual, and business broadband subscriptions due to SBA activity. Estimates are derived from PPR data, described in Section 3.35.[118] The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "PCC and SBA Average Household Size" contains a table with the estimated average household size across the service area of each grant. Section 4.13 provides a full description of "PCC and SBA Average Household Size."

The script performs the following operations:

- The "Grants" and "PCC and SBA Average Household Size" RDA files are loaded
- PPR data are read from the "SBA Quarterly PPR Data" Excel file as a table
- Thirty-nine PPR combinations (grant, year, and quarter) appear more than once in the data. For nineteen of these, at least one of the reported subscription variables is not consistent between duplicated entries. ASR used the PPRs published on NTIA's website to determine which duplicated entry was correct. This was done for the following grants and PPRs:
  - ZeroDivide Tribal, Tribal Digital Village Broadband Adoption Program: 2011 Q3
  - ZeroDivide, Generation ZD Digital Literacy Program: 2011-Q3
  - City and County of San Francisco, San Francisco Community Broadband Opportunities Program: 2013 Q1 and 2013 Q2
  - City of Chicago, SmartChicago Sustainable Broadband Adoption: 2012 Q1
  - Connected Nation, Inc., Public Adoption through Libraries (OPAL II): Every Community Online: 2011 Q4, 2012 Q1, and 2013 Q4[119]
  - Eastern Upper Peninsula Intermediate School District, Sparking Broadband Use in the Upper Peninsula of Michigan: 2012 Q1
  - Urban Affairs Coalition, Freedom Rings: Sustainable Broadband Adoption: 2011 Q4, 2012 Q1, 2012 Q2, 2012 Q3, 2012 Q4, 2013 Q1, and 2013 Q3
  - Connected Tennessee, LLC, Computers 4 Kids: Preparing Tennessee's Next Generation for Success: 2012 Q3
  - University of Wisconsin System, Building Community Capacity through Sustainable Broadband Adoption: 2012 Q2 and 2012 Q4
- Grantees stop submitting PPRs as their award periods end. Data for concluded grants must be carried over into later quarters to more accurately represent cumulative BTOP activities. The script uses the following process to carry over values from the last reported PPR to the end of 2013, where necessary:
  - Year and quarter number are formatted as a date
  - The first quarter with reported values (nonzero) is determined; if the grantee never reported values greater than zero, the quarter of the first available PPR is used
  - The last available PPR is determined
  - The Cartesian product of all SBA award numbers and all quarters with at least one PPR is created

- A rolled join is performed; PPR data and the Cartesian product table are joined by award number and date, and the last available value for every other variable is carried forward for each award number when a date is not present in the PPR data
- A Boolean indicator identifying quarters with carried-over values is created
- Observations in the rolled table for quarters earlier than the first reported quarter are filtered out

- Household and business subscription numbers reported in PPRs are cumulative. These numbers are differenced to determine the number of new subscribers in each quarter.
- Grantees report home broadband subscriptions as the number of households rather than the number of individuals. Subscription data and average household data are joined by award number and individual cumulative and quarterly subscriptions are estimated by multiplying the number of households by the average household size in the service area.
- The last quarter of data (2013 Q4) is extracted to a new table. Quarterly numbers are dropped. This table represents the total subscriptions (households, businesses, and individuals) due to each grant by the end of 2013.

The above results in two tables:

1. Total Subscribers ("sba_subscr_total"): the total subscriptions by households, businesses, and individuals due to each grant by the end of 2013
2. Quarterly Subscribers ("sba_subscribers"): cumulative and new subscriptions by households, businesses, and individuals due to each grant for each quarter from the grantee's first available PPR to the end of 2013

These tables are saved in the "SBA Subscribers" RDA file, with object names corresponding to those given in parenthesis above. The tables are also written, with sheet names corresponding the descriptive names above, to the "SBA Subscribers" Excel file as separate sheets. Tables are joined with the "Grants" table by award number to include grant information during the Excel write process.

4.26 Standard Deviation of Availability Growth

The "Standard Deviation of Availability Growth" R script manipulates NBM population and broadband availability data, described in Section 3.27.[120] The script performs the following operations:

- Total population data are read from the "NBM Statistics, 2011-06-30 and 2013-06-30 - Population" Excel file as a table
- In the total population data, missing values for the adjusted versions of broadband availability are replaced with the corresponding values of unadjusted broadband availability. Missing values are present in the input data because each adjustment (forward looking and backward looking) applies to only one release. Therefore, adjusted data must be related to unadjusted data to examine changes over time. In practice, this is equivalent to replacing missing values of adjusted data with unadjusted values. This method also allows the procedures used on unadjusted data to be used, without any modifications, on adjusted data.
- In the total population data, populations with broadband availability are converted to availability rates by dividing the availability populations by total population.
- In the total population data, availability rates are differenced at the county level. In other words, the June 30, 2011 availability rate is subtracted from the June 30, 2013 availability rate for every county included in NBM.
- In the total population data, the standard deviation of each series of differenced availability rates is calculated. This yields six values, one value for each version of the availability data (one

unadjusted and two adjusted versions for two definitions of broadband), and reduces the table from the county level to the national level.

- Standard deviations are reshaped into a table with one column representing the broadband definition, one column the adjustment, and one column with the standard deviation value.

The above results in a national-level table named "sd_diff_avail." The table is saved in the "Standard Deviation of Availability Growth" RDA file and "Standard Deviation of Availability Growth" Excel file.

4.27 Standard Deviation of Employment Growth

The "Standard Deviation of Employment Growth" R script manipulates BLS LAUS data for 2011 and 2012, described in Section 3.7.[121] The script performs the following operations:

- LAUS data for 2011 are read from the "BLS Local Area Unemployment Statistics - 2011" Excel file as a table
- LAUS data for 2012 are read from the "BLS Local Area Unemployment Statistics - 2012" Excel file as a table
- Employment size data from LAUS 2011 and LAUS 2012 data are joined by county FIPS code
- In the joined table, the log-difference (percentage change) of employment size is calculated
- The standard deviation of the log-difference is calculated. This is a single value for the United States.

The above results in a national-level table named "sd_diff_empl." The table is saved in the "Standard Deviation of Employment Growth" RDA file and "Standard Deviation of Employment Growth" Excel file.

4.28 Treatment and Potential Control Counties

The "Treatment and Potential Control Counties" R script determines which counties included in NBM data fall into the service area of one or more BTOP grants and which are potential control counties for matched pairs analysis. Potential control counties are those in the proposed service area of a BTOP application that was not funded and did not contain any BTOP or BIP activity.

The script makes use of "NBM Population Statistics" prepared data, described in Section 4.12. The script also uses the following input data:

- Awarded BIP grant stated service areas ("BIP"), described in Section 3.1
- Counties containing redacted CAI or POP locations ("Redacted"), described in Section 3.3
- Awarded BTOP grant stated service areas ("BTOP"), described in Section 3.2
- CAC CAI locations ("CAI"), described in Section 3.10
- CAC POP locations ("POP"), described in Section 3.11
- Evaluation study sample grant service areas ("Evaluation Study Sample"), described in Section 3.16
- Service areas of grant applications not funded by BTOP or BIP ("Unfunded"), described in Section 3.35

The "Treatment and Potential Control Counties" R script performs the following operations:

- The "NBM Population Statistics" RDA file is loaded

- "Evaluation Study Sample" data are read from the "CCI Case Study Service Areas" Excel file as a table
- County FIPS codes in the "CAI" data are read from the "BTOP Map Data - CAI Locations with Counties" Excel file as a character vector
- County FIPS codes in the "POP" data are read from the "BTOP Map Data - POP Locations with Counties" Excel file as a character vector
- County FIPS codes in the "BTOP" data are read from the "Awarded BTOP Service Areas" Excel file as a character vector
- County FIPS codes in the "BIP" data are read from the "Awarded BIP Service Areas" Excel file as a character vector
- "Unfunded" data are read from the "Service Areas of Grant Applications Not Funded by BTOP or BIP" Excel file as a table. The table is then filtered to BTOP grants and replaced by a character vector of the proposed service area county FIPS codes.
- County FIPS codes in the "Redacted" data are read from the "Awarded BTOP Service Areas of Redacted Locations" Excel file as a character vector
- A table of unique county FIPS codes in the "NBM Population Statistics" data is created. This represents every county in the United States. The status of every county (treatment county, potential control county, or neither) is determined through the following:
    - Treatment counties have FIPS codes contained in the "Evaluation Study Sample" table
    - Counties ineligible for matching have FIPS codes in the "CAI," "POP," "BTOP," "BIP," or "Redacted" character vectors of county FIPS codes
    - Possible potential control counties have FIPS codes in the "Unfunded" character vector of county FIPS codes
    - Counties outside of the contiguous United States are marked for removal from matched pair analysis
    - Potential control counties are those that were determined to be possible potential controls that are in the contiguous United States and not in ineligible areas ("CAI," "POP," "BTOP," "BIP," or "Redacted" counties)
    - Removed counties are those that are not treatment or potential control counties (i.e., those outside of the contiguous United States, outside of a proposed service area of a BTOP grant application that was not funded, in the service area of an awarded BTOP or BIP grant, or some combination of the three)

The above results in three county-level tables:

1. Matching Status ("treat.control.status"): identifies every county as a treatment, potential control, or removed county
2. Potential Control Status ("control.elimination"): determinations of whether a county meets the requirements for inclusion as a potential control county, for every county in the United States
3. All BTOP Counties ("btop"): every county in the United States in the service area of a BTOP grant and whether or not it is in the service area of a case study grant

The tables are saved in the "Treatment and Potential Control Counties" RDA file, with table names corresponding to those given in parenthesis above. The tables are also written, with sheet names corresponding the descriptive names above, to the "Treatment and Potential Control Counties" Excel file as separate sheets.

Section 5. Statistical Analysis

The long-term analysis portion of the *Final Report* includes a collection of statistical analyses. Statistical analysis files are sets of programs to conduct these analyses and files storing the results. Every statistical analysis includes the following:

- One R script to load prepared data and perform all necessary manipulations and estimations. R scripts are sets of instructions written in the R statistical programming language than can be executed in the R statistical package.
- One Excel file with a sheet for each results table generated by the R script. Excel files allow all users to inspect visually the generated tables and non-R users to access statistical analysis results in their software of choice.
- One RDA file containing all results tables generated by the R script. RDA files maintain the attributes of the results tables and can be quickly loaded when required.

Every prepared data R script has a header that defines the parent directory and calls the shared source code scripts described in Section 2. Users are required to redefine the parent directory to the correct location on their computer to reproduce the manipulation steps and output. This can be done in batch using the "Set Parent Directory in All Script Files" R script described in Section 1.3.

The following subsections describe the R scripts, in the order of execution, used to load prepared data, read input files, and perform statistical analysis. Lists of the variables in the statistical analysis results tables are provided in the "File List and Descriptions" Excel file.

5.1 Matching

As described in the *Study Design*, an effective and well-established way to develop estimates of the effects of programs such as BTOP is the use of matched pairs analysis.[122] The "Matching" R script contains the necessary steps to develop a county-level matching data set and find treatment-control matches for all counties in the service area of evaluation study sample CCI grants.

ASR uses the following prepared data in the matching process:

- "ACS 2010" is a county-level table containing selected ACS 2006-2010 statistics derived from ACS (2006-2010) data published by the Census Bureau.[123] Section 4.1 provides a full description of "ACS 2010." The following county-level variables are used in this analysis:
  - the percentage of population that speaks a language other than English in the home, also referred to as non-English percentage of population
- "NBM Population Statistics" is a county-level data containing population, demographic, and broadband availability statistics. "NBM Population Statistics" is derived from NBM data provided to ASR by NTIA.[124] Section 4.12 provides a full description of "NBM Population Statistics." The following county-level variables from the June 30, 2011 NBM release are used in this analysis:
  - broadband availability rate
  - forward-looking broadband availability rate
  - backward-looking broadband availability rate
  - total population
  - rural percentage of population
  - minority percentage of population

- percentage of population over sixty years of age
- percentage of population in poverty

- "Treatment and Potential Control Counties," in the context of the "Matching" analysis, is a county-level table classifying all counties in the United States as a treatment, potential control, or removed county.[125] "Treatment and Potential Control Counties" is derived from data published by the USDA, ProPublica, and NTIA, as well as data collected by ASR from individual grantees and other public sources.[126] Section 4.28 provides a full description of "Treatment and Potential Control Counties." The following county-level variables are used in this analysis:
  - treatment, potential control, or removed classification

The "Matching" R script loads the prepared data listed above. A matching data set is then built by joining the three prepared data tables by county FIPS code. The resulting table contains the county-level variables used in this analysis plus the additional variables available in "NBM Population Statistics." Two more variables used for matching are created in the combined matching data set: the log of total population is calculated; and an indicator identifying counties with 100 percent rural population percentage is created.

ASR uses the Matching package in R to find nearest-neighbor matched pairs.[127] Only treatment and potential control data are required in this package. Therefore, the matching data set is filtered to treatment and potential control counties. ASR also developed several criteria to restrict possible treatment and control matches:

- Availability rates must be within one percentage point. This ensures that the treatment and control counties in each match are highly similar, so that the difference-in-differences in availability is not dependent on the initial availability rates. For example, treatment county T has an initial availability rate of 90 percent, while control county C has an initial availability rate of 75 percent. Availability in county T increases to 100 percent over the BTOP period, while availability in county C increases to 90 percent. Even though the availability rate in county C only reached the initial rate of county T, the difference-in-differences estimate is (100 - 90) - (90 - 75) = -5 percent.[128]
- The natural logarithms of populations must be within half a standard deviation. A restriction on population ensures that aggregated groups of treatment and control counties are roughly similar in total availability. For example, consider the following treatment to control pairs:
  - Treatment county T1, population 50, is matched to control county C1, population 100. Both have an availability rate which increases from 25 percent to 100 percent.
  - Treatment county T2, population 75, is matched to control county C2, population 100. Both have an availability rate which increases from 75 percent to 100 percent.

  The two county-level difference-in-differences are equal to zero. However, aggregating to the total treatment and total control levels results in the following difference-in-differences: ((50 × 1 + 75 × 1) - (50 × 0.25 + 75 × 0.75)) ÷ (50 + 75) - ((100 × 1 + 100 × 1) - (100 × 0.25 + 100 × 0.75)) ÷ (100 + 100) = -5 percent. Restricting the difference in population between matches reduces the size of this bias.

  The natural logarithms of populations are used for two reasons. First, the distribution of the natural logarithm of population is much closer to the normal distribution than the distribution of population. Second, a constant restriction on the natural logarithm is equivalent to a variable restriction on population. For example, with a one-unit restriction on the log population, a treatment county with population 10,000 could be match to controls with populations approximately between 3,700 and 27,000, while a treatment county with population 500,000 could be match to controls with populations approximately between 184,000 and 1,360,000. This is desirable because of the wide range of populations in counties in the treatment and potential control groups.
- The rural percentage of population must be 100 percent in both counties or must be less than 100 percent in both counties.

To implement these restrictions, ASR developed a wrapper function around the Matching package. This function applies the restrictions above, iterating the matching process with progressively looser restrictions until matches are found for all treatment counties. The wrapper function performs the following actions:

- The function takes the following arguments: the name of the availability rate used for the match, a vector of the names of the covariates used for the match, the matching data set, and an index vector identifying the treatment counties in the matching data set. In this analysis, the last two arguments point to the same matching data set and index vector for every use of the wrapper function.
- The specified availability rate, the log of population, the specified covariates, and the rurality indicator are extracted from the matching data set and converted to a matrix
- ASR places a restriction on the difference between the broadband availability rate between a treatment and potential controls. In the first matching iteration, the desired maximum difference is 1 percentage point, so the initial restriction is set at 0.01 ÷ 1.1. The Matching package requires that restrictions are specified as standard deviations, called calipers, so the standard deviation of the specified availability rate is calculated.
- Until every treatment is matched to a control, the following steps are repeated:
  - The availability rate matching restriction in percentage points is multiplied by 1.1. In the first iteration, this results in a restriction of 0.01.
  - The availability rate matching restriction in standard deviations is calculated as the restriction in percentage points divided by the standard deviation of the availability rate. The log of population matching restriction is 0.5 standard deviations for every iteration. The matching restrictions for all other variables are set high enough (10 standard deviations) that matching is unrestricted on these variables.
  - Exact matches are required for the rurality indicator and not required for all other variables
  - Matches are identified using the Match function from the Matching package. Mahalanobis distance is used to determine the similarity between treatment and control counties.
  - Treatment and control county FIPS codes are extracted from the matching results for this iteration and saved in a list
- The list of matching results is combined. The matched pair for the most restrictive iteration with a successful match is selected for each treatment county.
- The selected matches are returned as a table. The specified availability rate and covariates are stored as attributes of the table.

The wrapper function allows ASR to repeat consistently the matching process using different availability rates and covariates. ASR performed matches for the following specifications of availability rate and covariates:

1. NTIA Primary (“m.ntia.unad.prim”): availability rate (NTIA broadband definition) and rural percentage of population
2. NTIA Sensitivity (“m.ntia.unad.sens”): availability rate (NTIA broadband definition) and rural, over-sixty, minority, in-poverty, and non-English percentages of population
3. NTIA Forward-Looking Primary (“m.ntia.forw.prim”): forward-looking availability rate (NTIA broadband definition) and rural percentage of population
4. NTIA Forward-Looking Sensitivity (“m.ntia.forw.sens”): forward-looking availability rate (NTIA broadband definition) and rural, over-sixty, minority, in-poverty, and non-English percentages of population
5. NTIA Backward-Looking Primary (“m.ntia.back.prim”): backward-looking availability rate (NTIA broadband definition) and rural percentage of population

6. NTIA Backward-Looking Sensitivity ("m.ntia.back.sens"): backward-looking availability rate (NTIA broadband definition) and rural, over-sixty, minority, in-poverty, and non-English percentages of population
7. NOFA Primary ("m.nofa.unad.prim"): availability rate (NOFA broadband definition) and rural percentage of population
8. NOFA Sensitivity ("m.nofa.unad.sens"): availability rate (NOFA broadband definition) and rural, over-sixty, minority, in-poverty, and non-English percentages of population
9. NOFA Forward-Looking Primary ("m.nofa.forw.prim"): forward-looking availability rate (NOFA broadband definition) and rural percentage of population
10. NOFA Forward-Looking Sensitivity ("m.nofa.forw.sens"): forward-looking availability rate (NOFA broadband definition) and rural, over-sixty, minority, in-poverty, and non-English percentages of population
11. NOFA Backward-Looking Primary ("m.nofa.back.prim"): backward-looking availability rate (NOFA broadband definition) and rural percentage of population
12. NOFA Backward-Looking Sensitivity ("m.nofa.back.sens"): backward-looking availability rate (NOFA broadband definition) and rural, over-sixty, minority, in-poverty, and non-English percentages of population

By definition, all of the above matches also include the log of total population, the restrictions of availability and log of total population, and exact matching on the rurality indicator in the matching specification.

The twelve matching specifications above correspond to twelve output tables. These tables, plus the matching data set, are saved in the "Matching" RDA file, with table names corresponding to those given in parenthesis above. The twelve results tables are also written, with sheet names corresponding the descriptive names above, to the "Matching" Excel file as separate sheets along with the matching data set.

5.2 Effect on Availability

The effect of BTOP on broadband availability in a treatment county is the difference between the increase in broadband availability in the treatment county and the increase in broadband availability in its matched control county. To estimate the programmatic effect of BTOP, ASR uses this difference-in-differences method on the entire 408 county evaluation study sample. Formally, this is given by:

Programmatic Effect on Availability =
(availability rate in evaluation study sample, June 30, 2013 –
availability rate in evaluation study sample, June 30, 2011) –
(availability rate in all matched counties, June 30, 2013 –
availability rate in all matched counties, June 30, 2011)

The "Effect on Availability" R script contains the necessary steps to estimate the effect of BTOP on availability at the program level. ASR uses the following statistical analysis results to estimate effects:

- "Matching" contains the results of twelve specifications of treatment-control matching using nearest neighbor matching with restrictions. The results are stored as tables with one column identifying the treatment county FIPS code and a second with the matched control county FIPS code. "Matching" also contains the matching data set, a county-level table of NBM population statistics, ACS (2006-2010) statistics, and county matching group classification. Section 5.1 provides a full description of "Matching." The following results and data are used in this analysis:
  - NTIA Primary, NTIA Sensitivity, NTIA Forward-Looking Primary, NTIA Forward-Looking Sensitivity, NTIA Backward-Looking Primary, NTIA Backward-Looking Sensitivity, NOFA

Primary, NOFA Sensitivity, NOFA Forward-Looking Primary, NOFA Forward-Looking Sensitivity, NOFA Backward-Looking Primary, and NOFA Backward-Looking Sensitivity matching results
    - NBM data from the June 30, 2011 and June 30, 2013 releases:
        - availability, forward-looking availability, and backward-looking availability
        - total population

ASR uses the following prepared data to estimate BTOP effects on availability:

- "Treatment and Potential Control Counties," in the context of the "Effect on Availability" analysis, is a county-level table exclusively containing all counties in the United States in an awarded BTOP CCI grant's service area and whether or not the county is in an evaluation study sample CCI grant's service area. "Treatment and Potential Control Counties" is derived from data published by NTIA and data collected by ASR from individual grantees and other public sources.[129] Section 4.28 provides a full description of "Treatment and Potential Control Counties."

The "Effect on Availability" R script loads the statistical analysis and prepared data RDA files listed above. After loading these files, the script defines several functions that reshape matching results, estimate effects, and format results. Functions are used in order to consistently repeat this analysis on the twelve different matching results:

- The function "reshapeMatched" combines matching results with broadband availability and population data from the matching data set:
    - The function accepts the following arguments: the names of the matching results object, the name of the 2011 availability rate used to identify the matches, the name of the corresponding 2013 availability rate, the names of the 2011 and 2013 total population variables, and the matching data set. In practice, the last two arguments are the same for every matching result table.
    - The specified availability rates are used to find the corresponding availability populations. Availability populations and total populations are extracted from the matching data set for the treatment counties into the treatment table. The same variables are extracted for the control counties into the control table.
    - Tagging variables differentiating treatment and control counties are added to the treatment and control tables
    - The treatment and control tables are concatenated and the availability and total populations are aggregated. The overall availability rates are calculated as the sum of the availability populations divided by the sum of the total populations.
    - The results are returned as a wide table with the following columns:
        - Release: the NBM release date, formatted as a date
        - Treatment: the overall broadband availability rate in the treatment counties
        - Control: the overall broadband availability rate in the control counties
        - matches: the name of the matching results table
- The function "tagResults" uses the "matches" column in the specified table to create a formatted version of the specified table. The function accepts one argument: a table with a column "matches" that contains the names of matching results objects (e.g., "m.ntia.unad.prim"). The function splits the information in the matching results name into three formatted categorical variables: definition ("NTIA" or "NOFA"), adjustment ("Availability," "Forward Looking," or "Backward Looking"), and match type ("Primary" or "Sensitivity").
- The function "did" calculates difference-in-differences for reshaped data:
    - The function accepts one argument: a reshaped and keyed table.[130] A reshaped table should have, at minimum, columns titled "Release," "Treatment," and "Control," containing release

date, treatment availability rate, and control availability rate, respectively, and should be keyed by, at minimum, "Release." In practice, the tables supplied "did" are always the results of the "reshapeMatched" function that have been formatted with the "tagResults" function.

- The availability rates are differenced. The difference-in-differences is calculated.
- A table similar to the specified table is returned with renamed and additional variables. The table will contain all the original table's keys, plus the columns "Treatment Rate," "Control Rate," "Treatment Difference," "Control Difference," and "Difference-in-Differences."

The "Effect on Availability" R script uses the functions above to estimate the effects of BTOP on the different broadband availability rates according to the different matching results. A list of matching result objects is defined, and all of the matching results are retrieved as a list of tables. The 2011 availability rates used in each matching result is retrieved; these were stored as attributes of the matching results tables by the "Matching" R script. The corresponding 2013 rates are then determined by substituting "2013" for "2011" in the 2011 list.

ASR then maps the "reshapeMatched" function to the list of matching results tables, 2011 availability rates, and 2013 availability rates. This reshapes all twelve matching results in one step and returns the reshaped results as a list of tables. This list of tables is then concatenated and formatted using the "tagResults" function. The resulting table is then keyed by definition, adjustment, and match type. The keyed table is then supplied to the "did" function to calculate the difference-in-differences estimates of the effect of BTOP for the twelve different matching scenarios.

The definition, adjustment, match type, and difference-in-differences columns for the 2013 release are extracted to a new table for later presentation. The "Treatment and Potential Control Counties" prepared data are then used to calculate the evaluation study sample, rest of BTOP, and all of BTOP total populations. The estimated effects of BTOP (the difference-in-differences) are then multiplied by these total populations to estimate the effect of BTOP in terms of persons, and the results are joined to the new table.

ASR investigated the robustness of the estimated effects of BTOP to individual treatment and control pairs using a resampling method. To repeat this for different matching results, ASR defined a wrapper function to bootstrapping procedures from the boot package in R.[131] The "resampleEffect" function is described below:

- The function accepts the following arguments: the names of the matching results object, the name of the 2011 availability rate used to identify the matches, the name of the corresponding 2013 availability rate, a seed value used for random number generation, and the matching data set. In practice, the last argument is the same for every matching result table.
- The function converts the 2011 and 2013 availability rates variable names to availability population variable names. Availability populations and total populations are then extracted from the matching data set for the treatment and control counties. The resulting table has one row for every treatment-control matched pair with columns for 2011 and 2013 availability populations and 2011 and 2013 total populations for both the treatment and control counties.
- The nested function "didfun" is defined inside "resampleEffect." This function is required by the bootstrapping procedure. The internal function takes the merged treatment and control availability and population table as its first argument and a vector of row numbers as its second. The function calculates and returns the difference-in-differences estimate for the rows in the supplied row number vector. The row number vector is generated internally by the bootstrapping procedure.
- The seed supplied to "resampleEffect" is set and the merged treatment and control availability and population table and internal function are supplied to the bootstrapping function ("boot"). The number of replications is set to 1,000.
- The results of the bootstrapping function are used to estimate a basic 95 percent confidence interval using the "boot.ci" function.

- The results are returned as a table with a column for the estimated effect (the estimated effect for all rows, identical to the effect calculated by the "did" function), the lower 95 percent confidence limit, the upper 95 percent confidence limit, the individual replications of the effect value, and the matching results name. The table will have 1,000 rows because there are 1,000 replications of the effect value; all other columns will have identical values for all rows.

To resample the effect for all matching results, ASR followed a similar process to that used to reshape matching results and calculate difference-in-differences. The matching results, 2011 availability rates, and 2013 availability rates lists are reused for this process. Twelve seed values are determined for use with the resampling function.

ASR maps the "resampleEffect" function to the lists of matching results tables, 2011 availability rates, 2013 availability rates, and seed values. This performs the resampling calculations for all twelve matching results in one step and returns the results as a list of tables. This list of tables is then concatenated and formatted using the "tagResults" function. The resulting table is then keyed by definition, adjustment, and match type. The unique values of the estimated effect, the lower 95 percent confidence limit, and the upper 95 percent confidence limit are extracted to a separate table for later presentation. The individual replication values for the base case used in the *Final Report* (the NTIA definition of broadband, with no adjustments to availability, using only the Primary matching variables) are also extracted to a separate table for later presentation.

The "Effect on Availability" analysis results in four tables that are saved for later use:

1. Effects ("effects"): the difference-in-differences estimate and the evaluation study sample, rest of BTOP, and all of BTOP affected populations for the twelve different matching scenarios
2. Effects Intervals ("effects.intervals"): the difference-in-differences estimate and the 95 percent lower and upper confidence limits for the twelve different matching scenarios
3. Base Case Resamples ("base.replications"): the difference-in-differences estimate, 95 percent lower and upper confidence limits, and individual replication values for the NTIA Primary matching scenario
4. Overall Estimates ("did.overall"): the treatment and control rates for the 2011 and 2013 releases, the treatment and control differences, and the difference-in-differences for the twelve different matching scenarios

The four tables above are saved in the "Effect on Availability" RDA file, with table names corresponding to those given in parenthesis above. The four results tables are also written, with sheet names corresponding the descriptive names above, to the "Effect on Availability" Excel file as separate sheets.

5.3 Incidence Analysis

To investigate the long-term effect of BTOP on broadband availability to vulnerable populations, the incidence analysis statistical analysis follows the same difference-in-differences methodology used to estimate the overall effect of BTOP, but analyzes the availability rates among vulnerable populations instead of total populations. The "Incidence Analysis" R script contains the necessary steps to estimate the effect of BTOP on availability among vulnerable populations at the program level. ASR uses the following statistical analysis results to estimate effects:

- "Matching" contains the results of twelve specifications of treatment-control matching using nearest neighbor matching with restrictions. The results are stored as tables with one column identifying the treatment county FIPS code and a second with the matched control county FIPS code. "Matching" also contains the matching data set, a county-level table of NBM population statistics, ACS (2006-2010) statistics, and county matching group classification. Section 5.1 provides a full description of "Matching." The following results and data are used in this analysis:

- NTIA Primary, NTIA Sensitivity, NTIA Forward-Looking Primary, NTIA Forward-Looking Sensitivity, NTIA Backward-Looking Primary, NTIA Backward-Looking Sensitivity, NOFA Primary, NOFA Sensitivity, NOFA Forward-Looking Primary, NOFA Forward-Looking Sensitivity, NOFA Backward-Looking Primary, and NOFA Backward-Looking Sensitivity matching results
- ACS (2006-2010) demographic data:
  - non-English percentage of population

ASR uses input files containing NBM population, demographic, and broadband availability data to estimate BTOP effects on availability.[132] The NBM data represent county-level total population and county-level minority, over-sixty, in-poverty, and rural populations, all described in Section 3.27. The "Incidence Analysis" R script performs the following procedures on these data:

- Total ("NBM Statistics, 2011-06-30 and 2013-06-30 - Population"), rural ("NBM Statistics, 2011-06-30 and 2013-06-30 - Rural Population"), minority ("NBM Statistics, 2011-06-30 and 2013-06-30 - Minority Population"), over-sixty ("NBM Statistics, 2011-06-30 and 2013-06-30 - Over 60 Population"), and in-poverty ("NBM Statistics, 2011-06-30 and 2013-06-30 - Poverty Population") population NBM data are read from their respective Excel files into a list of tables. The list is then concatenated into one NBM table.
- Total population data are extracted from the NBM data and joined by county FIPS code with non-English percentage of population data. In the joined table, the non-English population is estimated as the non-English percentage of population multiplied by total population. The non-English broadband availability populations are estimated as the total population availability rates multiplied by the non-English population.
- Non-English population and broadband availability data are appended to the NBM table
- In the NBM table, missing values for the adjusted versions of broadband availability are replaced with the corresponding values of unadjusted broadband availability. Missing values are present in the input data because each adjustment (forward looking and backward looking) applies to only one release. Therefore, adjusted data must be related to unadjusted data to examine changes over time. In practice, this is equivalent to replacing missing values of adjusted data with unadjusted values. This method also allows the procedures used on unadjusted data to be used, without any modifications, on adjusted data.
- The NBM table is reshaped to mimic the matching data set: each row represents a unique county and each column represents the value of one variable in one release (e.g., minority population with broadband availability according to the NOFA definition the June 30, 2011 release).

ASR estimates the effects of BTOP on availability in vulnerable populations and the robustness of the effect for the base case (NTIA Primary) in the same step in "Incidence Analysis." To repeat the effect and confidence interval estimation for each vulnerable population, ASR defined a wrapper function to bootstrapping procedures from the boot package in R.[133] The "incResampleEffect" function is described below:

- The function accepts the following arguments: the names of the matching results object, the name of the 2011 vulnerable population with availability, the name of the corresponding 2013 vulnerable population with availability, the names of the 2011 and 2013 vulnerable populations, a seed value used for random number generation, and the NBM table. In practice, the name of the matching results and the NBM table are the same for every vulnerable population.
- The function extracts vulnerable populations with availability and vulnerable populations from the NBM table for the treatment and control counties. The resulting table has one row for every treatment-control matched pair with columns for 2011 and 2013 vulnerable populations with availability and the 2011 and 2013 vulnerable populations for both the treatment and control counties.

- The nested function “didfun” is defined inside “incResampleEffect.” This function is required by the bootstrapping procedure. The internal function takes the merged treatment and control vulnerable availability and population table as its first argument and a vector of row numbers as its second. The function calculates and returns the difference-in-differences estimate for the rows in the supplied row number vector. The row number vector is generated internally by the bootstrapping procedure.
- The seed supplied to “incResampleEffect” is set and the merged treatment and control vulnerable availability and population table and internal function are supplied to the bootstrapping function (“boot”). The number of replications is set to 1,000.
- The results of the bootstrapping function are used to estimate a basic 95 percent confidence interval using the “boot.ci” function
- The results are returned as a table with a column for the estimated effect (the estimated difference-in-differences using all matches), the lower 95 percent confidence limit, the upper 95 percent confidence limit, and the matching results name

ASR estimates the effect of BTOP on availability for vulnerable populations for the base case (NTIA Primary) only. A list of vulnerable populations variable prefixes is created. This list is used to generate lists of 2011 vulnerable populations with availability, 2013 vulnerable populations with availability, and 2011 and 2013 vulnerable populations. A seed is also generated for each vulnerable population.

ASR maps the “incResampleEffect” function to the NTIA Primary matching result and the lists of 2011 vulnerable populations with availability, 2013 vulnerable populations with availability, and 2011 and 2013 vulnerable populations. This performs the resampling calculations for all vulnerable populations in one step and returns the results as a list of tables. This list of tables is then concatenated and a formatted column representing vulnerable populations is added. Since the same matching result is used for all rows in this table, the matching results name column is dropped.

“Incidence Analysis” results in one table that is saved for later use:

1. Incidence (“incidence”): the difference-in-differences estimate and 95 percent lower and upper confidence limits for each vulnerable population

The table above is saved in the “Incidence Analysis” RDA file and “Incidence Analysis” Excel file.

5.4 Extrapolation Table

ASR extrapolates benefits by applying the estimated programmatic effect of BTOP to county-level statistics of BTOP service area counties. The “Extrapolation Table” R script combines statistical analysis results with other prepared data to create a single table representing every estimated effect of BTOP for every county in a BTOP service area. ASR uses the following statistical analysis results to estimate effects:

- “Effect on Availability,” in the context of “Extrapolation Table,” is a program-level table containing the difference-in-differences estimate for the twelve different matching scenarios. Each row in the table represents the estimated effect of BTOP for a unique combination of the definition of broadband, any adjustments made to the data, and the matching group. Section 5.2 provides a full description of the analysis leading to these results. The following results are used in this analysis:
    - NTIA Primary, NTIA Sensitivity, NTIA Forward-Looking Primary, NTIA Forward-Looking Sensitivity, NTIA Backward-Looking Primary, NTIA Backward-Looking Sensitivity, NOFA Primary, NOFA Sensitivity, NOFA Forward-Looking Primary, NOFA Forward-Looking

Sensitivity, NOFA Backward-Looking Primary, and NOFA Backward-Looking Sensitivity difference-in-differences estimates of the effect of BTOP

ASR uses the following prepared data to estimate BTOP effects on availability:

- "Rates of Adoption by Households with Availability" is a county-level table containing the estimated broadband rates of adoption by households with availability for all available counties in the United States derived from data published by FCC and data provided to ASR by NTIA.[134] Each row in "Rates of Adoption by Households with Availability" represents an estimate of the rate of broadband adoption in a single county for a particular version of the data (i.e., availability, forward-looking availability, and backward-looking availability). Section 4.23 provides a full description of "Rates of Adoption by Households with Availability."
- "Treatment and Potential Control Counties," in the context of the "Effect on Availability" analysis, is a county-level table exclusively containing all counties in the United States in an awarded BTOP CCI grant's service area and whether or not the county is in an evaluation study sample CCI grant's service area. "Treatment and Potential Control Counties" is derived from data published by NTIA and data collected by ASR from individual grantees and other public sources.[135] Section 4.28 provides a full description of "Treatment and Potential Control Counties."

The "Extrapolation Table" R script loads the statistical analysis results and prepared data listed above. The objective of this script is to create a table with one record for every combination of county FIPS code, definition of broadband, adjustments made to the data, and match type, with values for the estimated effect, the rate of adoption by households with availability, and whether or not the county is in the service area of a grant included in the evaluation study sample. The script uses the following joins:

- A master table is created by taking the Cartesian product of all possible values of definition of broadband, adjustments made to the data, match type, and county FIPS code (for BTOP counties only, loaded from "Treatment and Potential Control Counties")
- The master table and the table of estimated effects loaded from "Effect on Availability" are joined by definition, adjustment, and match type. This results in an additional column of estimated effects on the master table.
- The master table and the estimated rates of adoption table are joined by county FIPS code and adjustment. This results in an additional column of rates of adoption on the master table.
- The master table and the table of all BTOP counties, loaded from "Treatment and Potential Control Counties," are joined by county FIPS code. This results in an additional column containing an evaluation study flag on the master table.

The end result of "Extrapolation Table" is one table saved for later use:

1. Extrapolate ("extrapolate"): a table with one record for every combination of county FIPS code, definition of broadband, adjustments made to the data, and match type, with values for the estimated effect, the rate of adoption by households with availability, and whether or not the county is in the service area of an evaluation study sample grant

The table above is saved in the "Extrapolation Table" RDA file and "Extrapolation Table" Excel file.

Section 6. Final Report Executive Summary

This section includes descriptions of all files, programs, and processes used to summarize the activities, effects, impacts, or benefits of BTOP for the *Final Report*. Summary tables and figures are created from prepared data and input files, statistical analysis results, and combinations thereof. Each summary tables and figure listed below includes the following:

- One R script to load one or more sets of prepared data, read input files, and load statistical analysis results. Loaded and imported data are prepared as necessary to produce the summary tables and figures included in the *Final Report*. R scripts are sets of instructions written in the R statistical programming language than can be executed in the R statistical package.
- One output file for every summary table and figure. Tables are saved as Excel files; tables were copied from the Excel files and pasted into the *Final Report*. Figures are saved as PNG files; figures were inserted as pictures into the *Final Report*.

Every script has a header that defines the parent directory and calls the shared source code scripts described in Section 2. Users are required to redefine the parent directory to the correct location on their computer to reproduce the manipulation steps and output. This can be done in batch using the "Set Parent Directory in All Script Files" R script described in Section 1.3.

The following subsections describe the R scripts, in the order of presentation, used to create summary tables and figures for the *Executive Summary* of the *Final Report.* All scripts and generated files discussed below are located in the "0. Executive Summary" folder.

6.1 Final Report Executive Summary: About BTOP

Final Report Executive Summary: About BTOP contains budget and CCI activity summary figures that are not provided by tables elsewhere in the *Final Report*. The "About BTOP" script derives these figures. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "Budgets" is a grant-level table containing the federal, match, and total funding amounts for all BTOP grants from PPRs.[136] Section 4.2 provides a full description of "Budgets."
- "CCI Progress" is a set of four distinct summary tables containing grant-level network and CAI data derived from PPRs.[137] The "Network Progress" and "CAI Progress" tables contain cumulative totals by quarter, and the "Network Totals" and "CAI Totals" tables contain cumulative totals as of December 31, 2013. Section 4.6 provides a complete description of "CCI Progress."

The prepared data above are loaded. The following steps are executed separately to summarize budgetary and CCI activity data:

- "Grants" and "Budgets" tables are joined by award number and budget values are converted from dollars to millions of dollars. Data are tabulated for each grant type and evaluation study sample status, calculating total federal, matching, and total budgets for three sets of grants: all grants less excluded (public safety) and defunded grants. Results are printed to the screen and are not saved to any files.
- "Grants," "Budgets," "Network Totals," and "CAI Totals" tables are joined by award number, and PCC, SBA, excluded, and defunded grants are dropped. Total new, new leased, upgraded, and

existing leased fiber miles, total connected CAIs, total interconnection points, and total signed agreements variables are extracted. Fiber miles totals are combined. Variables are aggregated for each grant type and evaluation study sample status. Results are printed to the screen and are not saved to any files.

6.2 Final Report Executive Summary: Social and Economic Impacts of BTOP Projects

6.2.1 Final Report Figure 1. Estimated Effect of BTOP on Broadband Availability in the CCI Evaluation Study Sample Service Area for Different Populations of Interest

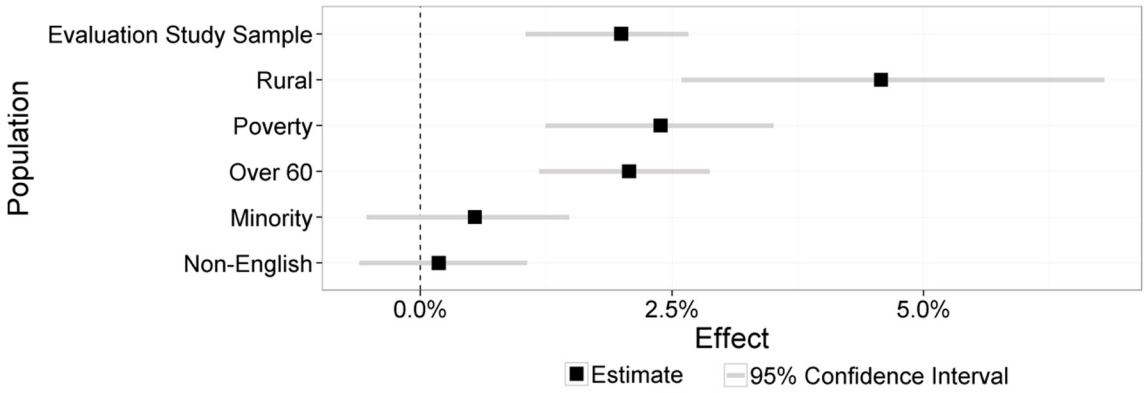
Final Report Figure 1 presents the results of ASR's analysis of the changes in broadband availability rates among vulnerable populations. The "Figure 1" R script reproduces this figure. ASR uses the following statistical analysis results to generate the figure:

- "Incidence Analysis" uses NBM data on population, demographics, and broadband availability; Census Bureau ACS (2006-2010) demographic data; and the matched pairs identified in ASR's "Matching" statistical analysis to estimate the effect of BTOP on broadband availability among vulnerable populations in the evaluation study sample service area. Section 5.3 provides a full description of "Incidence Analysis."
- "Effect on Availability" uses data on population, demographics, and broadband availability and the matched pairs identified in ASR's "Matching" statistical analysis to estimate the effect of BTOP on broadband availability in the evaluation study sample service area. Section 5.2 provides a full description of "Effect on Availability."

The script loads the statistical analysis results discussed above. Estimates and confidence intervals for the effect of BTOP on broadband availability among vulnerable populations are concatenated with estimates and confidence intervals for the effect on availability among the total population. Only estimates and confidence intervals for the primary set of matches using the unadjusted form of availability data based on the NTIA definition of broadband are used in this figure.

After the data are concatenated, a forest plot is created to present the estimates and confidence intervals for the different populations. This figure is then saved as the "Figure 1" PNG file. Figure 1 below presents the results of the R script. ASR used an image editor to change the orientation of the legend from vertical to horizontal.

Figure 1. Estimated Effect of BTOP on Broadband Availability in the CCI Evaluation Study Sample Service Area for Different Populations of Interest



Section 7. Final Report Section 1. Introduction

This section includes descriptions of all files, programs, and processes used to summarize the activities, effects, impacts, or benefits of BTOP for the *Final Report*. Summary tables and figures are created from prepared data and input files, statistical analysis results, and combinations thereof. Each summary tables and figure listed below includes the following:

- One R script to load one or more sets of prepared data, read input files, and load statistical analysis results. Loaded and imported data are prepared as necessary to produce the summary tables and figures included in the *Final Report*. R scripts are sets of instructions written in the R statistical programming language than can be executed in the R statistical package.
- One output file for every summary table and figure. Tables are saved as Excel files; tables were copied from the Excel files and pasted into the *Final Report*. Figures are saved as PNG files; figures were inserted as pictures into the *Final Report*.

Every script has a header that defines the parent directory and calls the shared source code scripts described in Section 2. Users are required to redefine the parent directory to the correct location on their computer to reproduce the manipulation steps and output. This can be done in batch using the "Set Parent Directory in All Script Files" R script described in Section 1.3.

The following subsections describe the R scripts, in the order of presentation, used to create summary tables and figures for *Section 1. Introduction* in the *Final Report.* All scripts and generated files discussed below are located in the "1. Introduction" folder.

7.1 Final Report Section 1.3 BTOP Grantee Descriptive Statistics

7.1.1 Final Report Table 2. Average BTOP Budgets (Million USD)

Final Report Table 2 summarizes the average and total budget size for PCC, SBA, and CCI grants in the evaluation study sample and in all of BTOP. The "Table 2" R script reproduces this table. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "Budgets" is a grant-level table containing the federal, match, and total funding amounts for all BTOP grants from PPRs.[138] Section 4.2 provides a full description of "Budgets."

The "Grants" and "Budgets" RDA files are loaded and the tables are joined by award number; excluded and defunded grants are dropped and budget values are converted from dollars to millions of dollars. Data are tabulated for each grant type and evaluation study sample status, calculating average and total federal, matching, and total budgets. This tabulation is written to the "Table 2" Excel file. Table 2 below presents the results of the R script, is presented below. *Final Report Table 2* presents the estimates for the study sample and all of BTOP.

Table 2. Average BTOP Budgets (Million USD)

Metric	Budget	PCC		SBA		CCI	
		Study Sample	All	Study Sample	All	Study Sample	All
Average	Federal	4.0	3.1	7.2	5.8	45.4	26.7
	Match	1.3	1.3	2.3	2.4	15.5	9.2
	Total	5.4	4.4	9.5	8.2	60.9	35.9
Total	Federal	32.2	199.2	50.1	250.3	545.0	2,905.7
	Match	10.8	84.9	16.2	101.3	185.6	1,005.4
	Total	42.9	284.1	66.3	350.6	730.6	3,911.2
Number of grants		8	65	7	43	12	109

7.1.2 Final Report Table 3. PCC and SBA Training

Final Report Table 3 combines and summarizes grant-level training and budget data. The table presents total and average training participants, training hours, and total budget and budget dollars per training hour for PCC and SBA grants in the evaluation study sample and in all of BTOP. The "Table 3" R script reproduces this table. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "Budgets" is a grant-level table containing the federal, match, and total funding amounts for all BTOP grants from PPRs.[139] Section 4.2 provides a full description of "Budgets."
- "PCC and SBA Training Hours and Participants" is a table containing a grant-level summary of participants and training hours for each focus area plus open lab access hours from PPRs.[140] Section 4.17 provides a full description of "PCC and SBA Training Hours and Participants."

The prepared data above are loaded. Open Lab is filtered out of the training hours and participants table; data are then aggregated to the grant level. "Grants," "Budgets," and aggregated "PCC and SBA Training Hours and Participants" are joined by award number, and CCI and defunded grants are dropped.[141]

Training participants, training hours, and total budgets are selected from the merged data. Four summary tables are created: aggregated totals by evaluation study sample status, all BTOP aggregated totals, averages by evaluation study sample status, and all BTOP averages. Aggregated totals summary tables are combined, and budget dollars per training hour is calculated. This table is then combined with the averages tables. Budget values are converted from dollars to millions of dollars.

The data are reshaped and columns and values are formatted. The data are then tabulated for each grant type and evaluation study sample status. This tabulation is written to the "Table 3" Excel file. Table 3 below presents the results of the R script. *Final Report Table 3* presents the estimates for the evaluation study sample and all of BTOP.

Table 3. PCC and SBA Training

Metric	PCC		SBA	
	Study Sample	All	Study Sample	All
Total training participants	483,751	2,473,818	819,421	1,956,807
Average number of training participants per grantee	60,469	38,059	117,060	45,507
Total training hours	4,241,155	11,533,901	1,734,035	9,299,758
Average number of training hours per grantee	530,144	177,445	247,719	216,273
Total budget (millions)	$43	$284	$66	$351
Average budget (millions) per grantee	$5.4	$4.4	$9.5	$8.2
Budget dollars per training hour	$10.12	$24.63	$38.26	$37.70

7.1.3 Final Report Table 4. PCCs and Lab Hours

Final Report Table 4 summarizes budget and computer center data for PCC grants in the evaluation study sample and in all of BTOP. The "Table 4" R script reproduces this table. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "Budgets" is a grant-level table containing the federal, match, and total funding amounts for all BTOP grants from PPRs.[142] Section 4.2 provides a full description of "Budgets."
- "PCC Equipment" is a table containing cumulative numbers of PCCs established and improved by grant, year, and category of PCC derived from PPRs.[143] Section 4.20 provides a full description of "PCC Equipment."
- "PCC Sites" is a table containing location data for all PCCs established or upgraded by grantees derived from the CAC.[144] For each location, data in the table include institution type, PCC type (new or improved), number of workstations, weekday and weekend hours, connection speed, and weekly users. Section 4.21 provides a full description of "PCC Sites."

The prepared data above are loaded. The number of new and improved PCCs is calculated and added together for each grant using "PCC Sites." Total weekday and weekend hours are calculated for each grant using "PCC Sites."

"PCC Equipment" is used to estimate the quarterly progress made from the beginning of each grant until the end of 2012. Quarterly progress will be applied to weekly hours of operation to estimate the total number of hours open. Quarterly progress is estimated using the number of new workstations, the number of workstations upgraded, and the number of wireless broadband connections established. For each of these variables, every quarter's value is divided by the 2012 Q4 value. This ratio is restricted to be between zero and one, or missing. The average of the three ratios, with missing values removed, is then calculated. This provides the estimated progress towards end of 2012 achievements for every quarter of activity.

"Grants," aggregated hours from "PCC Sites," and estimated quarterly progress from "PCC Equipment" are joined by award number and SBA, CCI, and defunded grants are dropped. Weekly

hours per quarter are estimated by multiplying aggregated hours by quarterly progress. Weekly hours are converted to quarterly by multiplying by thirteen. Hours are then aggregated to the grant level.

Grant-level estimated hours are joined with "Budgets" and summarized locations from "PCC Sites" by award number and SBA, CCI, and defunded grants are dropped. Improved PCCs, new PCCs, total PCCs, weekday hours, weekend hours, total hours, and total budgets are selected from the merged data. Two summary tables are created: aggregated totals by evaluation study sample status and all BTOP aggregated totals. Aggregated totals summary tables are combined, and budget dollars per total PCC and total hours are calculated. Budget values are converted from dollars to millions of dollars.

The data are reshaped and columns and values are formatted. The data are then tabulated for each evaluation study sample status and written to the "Table 4" Excel file. Table 4 below presents the results of the R script. *Final Report Table 4* presents the results for the evaluation study sample and all of BTOP.

Table 4. PCCs and Lab Hours

Activity	Evaluation Study Sample	All
Total budget (millions)	$43	$284
New PCCs established	93	656
Existing PCCs improved	359	2,471
Total PCCs	452	3,127
Budget per PCC	$94,955	$90,849
Total weekday lab hours (est.)	1,841,537	12,130,206
Total weekend lab hours (est.)	331,376	1,933,574
Total lab hours (est.)	2,172,913	14,063,780
Budget per lab hour (est.)	$19.75	$20.20

7.1.4 Final Report Table 5. SBA Household Connections

Final Report Table 5 combines and summarizes grant-level household subscriber and budget data for SBA grants in the evaluation study sample and all of BTOP. The "Table 5" R script reproduces this table. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "Budgets" is a grant-level table containing the federal, match, and total funding amounts for all BTOP grants from PPRs.[145] Section 4.2 provides a full description of "Budgets."
- "SBA Subscribers" contains a grant-level table of total new individual, household, and business subscribers due to SBA grant activity. The table relies on grantee-reported data from PPRs and publicly available data published by Census Bureau.[146] Section 4.25 provides a full description of "SBA Subscribers."

The prepared data above are loaded. "Grants," "Budgets," and "SBA Subscribers" are joined by award number and PCC, CCI, and defunded grants are dropped. Total household subscribers and total budgets are selected from the merged data. Two summary tables are created: aggregated totals by evaluation study sample status and all of BTOP aggregated totals. Aggregated totals

summary tables are combined, and budget dollars per household subscriber is calculated. Budget values are converted from dollars to millions of dollars. The data are then tabulated for each evaluation study sample status and written to the "Table 5" Excel file. Table 5 presents the results of the R script. *Final Report Table 5* presents the estimates for the study sample and all of BTOP.

Table 5. SBA Household Connections

Metric	Evaluation Study Sample	All
Household subscribers	334,440	736,812
Dollars spent per household subscriber	$198.38	$475.85

7.1.5 Final Report Table 6. CCI Projects vs. Sample

Final Report Table 6 summarizes grant-level budget data for CCI grants in the evaluation study sample and all of BTOP. The "Table 6" R script reproduces this table. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "Budgets" is a grant-level table containing the federal, match, and total funding amounts for all BTOP grants from PPRs.[147] Section 4.2 provides a full description of "Budgets."

The prepared data above are loaded. "Grants" and "Budgets" are joined by award number and PCC, SBA, excluded, and defunded grants are dropped. Budget values are converted from dollars to millions of dollars. Three tables summarizing values for evaluation study, other, and all CCI grants are created: average budget, count, and aggregated total budget. The summary tables are combined, and the percentages of total count and total budget are calculated. Values are formatted, and data are reshaped. The data are then tabulated for each evaluation study sample status and written to the "Table 6" Excel file. Table 6 below presents the results of the R script. *Final Report Table 6* presents estimates for the evaluation study sample and all of BTOP.

Table 6. CCI Projects vs. Sample

Metric	Evaluation Study Sample	All
Average project size (millions)	$61	$36
Number of projects	12	109
Total expenditures (millions)	$731	$3,911
Percent of CCI projects	11%	100%
Percent of CCI expenditures	19%	100%

7.1.6 Final Report Table 7. CCI Network Summary

Final Report Table 7 combines and summarizes grant-level network and budget data for CCI grants in the evaluation study sample and all of BTOP. The "Table 7" R script reproduces this table. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."

- "Budgets" is a grant-level table containing the federal, match, and total funding amounts for all BTOP grants from PPRs.[148] Section 4.2 provides a full description of "Budgets."
- "CCI Progress" is a set of four distinct summary tables containing grant-level network and CAI data derived from PPRs.[149] The "Network Progress" and "CAI Progress" tables contain cumulative totals by quarter, and the "Network Totals" and "CAI Totals" tables contain cumulative totals as of December 31, 2013. Section 4.6 provides a complete description of "CCI Progress."

The prepared data above are loaded. "Grants," "Budgets," "Network Totals," and "CAI Totals" are joined by award number, and PCC, SBA, excluded, and defunded grants are dropped. Total network miles, total connected CAIs, total signed agreements, and total budgets are selected from the merged data. Two summary tables are created: aggregated totals by evaluation study sample status and all BTOP aggregated totals. Aggregated totals summary tables are combined, and budget dollars per mile, per CAI, and per agreement are calculated. Budget values are converted from dollars to millions of dollars. Data are reshaped and formatted. The data are then tabulated for each evaluation study sample status. This tabulation is written to the "Table 7" Excel file. Table 7 below presents the results of the R script. *Final Report Table 7* presents the estimates for the evaluation study sample and all of BTOP.

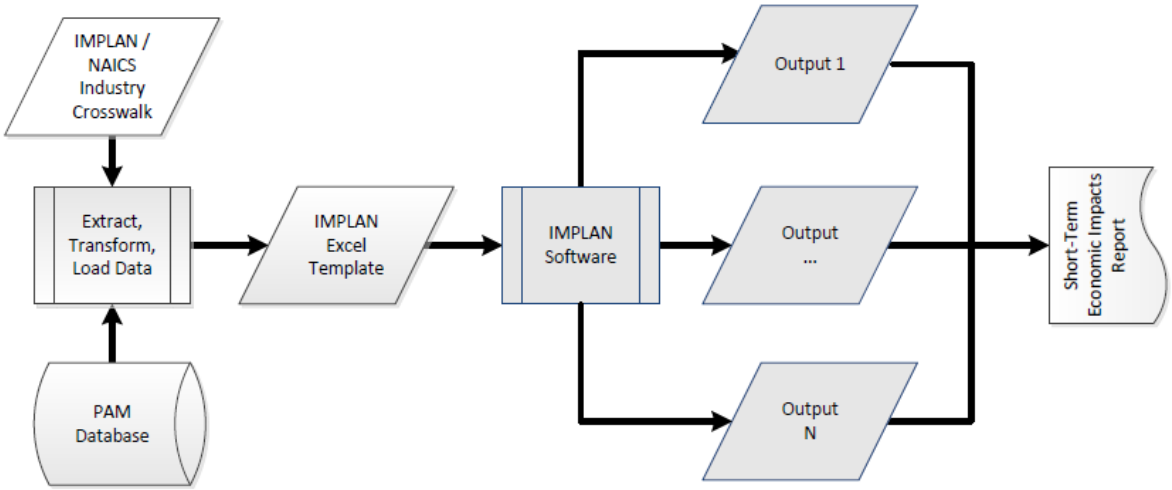
Table 7. CCI Network Summary

Metric	Evaluation Study Sample	All
Total budgeted (millions)	$731	$3,911
New network miles deployed	7,301	42,124
Budgeted dollars per mile	$100,074	$92,849
CAIs connected	5,193	21,240
Budgeted dollars per CAI	$140,698	$184,141
New agreements signed with last mile and wholesale providers	143	842
Budgeted dollars per agreement (millions)	$5.1	$4.6

Section 8. Final Report Section 2. Short-Term Economic Impacts

ASR estimated the short-term economic impacts of BTOP grants using publicly available grantee expenditure and budget figures. This section provides a full description of the methodology and data used to aid in the reproduction of this analysis. ASR used the IMPLAN software to estimate short-term economic impacts of BTOP. IMPLAN is an industry-standard tool for economic impact estimation. Figure A summarizes the analysis process. The following subsections describe each step included in the diagram.

Figure A. IMPLAN Analysis Process Flow



ASR completed the short-term economic impact analysis using a combination of Excel workbooks and SAS programs, in addition to the IMPLAN software. The "File List and Descriptions" spreadsheet describes each file used in the analysis. In the case of SAS programs, the preconditions needed for the program to run. Any resulting postconditions are also listed. This document references the files contained in the spreadsheet by name.

8.1 Prepare BTOP Grantee Budget Data

Grantee expenditure and budget data are publicly available on the BTOP website in quarterly Performance Progress Reports (PPRs).[150] Grantees submit cumulative expenditure data each quarter for a set of budget categories defined by NTIA. For simplicity, ASR created a data set containing this quarterly expenditure information for all BTOP projects from the Post-Award Monitoring (PAM) database provided by NTIA. As PAM contains unredacted data, this data set is not suitable for publication as provided by NTIA to ASR. The data set created by ASR will allow for the replication of this analysis without access to PAM. All of the data included in the data set are publicly available on the NTIA BTOP website. ASR included expenditure data for all BTOP grants, except seven public safety-focused CCI grants. ASR removed these grants from the evaluation study at the request of NTIA. The data set is available as the "BTOP_Expenditures_2013_06_11.xls" Excel file. All analysis of short-term economic impacts groups PCC and SBA grants together. CCI grants are treated separately.

Grantee budgets include both federal award dollars and non-federal match dollars. Grantees report budgets and expenditures quarterly, differentiating between federal and match dollars.[151] Using the "ExpenditureAnalysis_IMPLANprep" SAS program, ASR created a series of summary tables from the expenditure data set, including actual expenditure tables for summaries of spent funds and budget tables for projects' total budget figures. Summary tables are organized by grant type (PCC and SBA or CCI) and dollar type (federal or non-federal matching), and figures are grouped by budget categories defined by NTIA and used on quarterly PPRs.

The budget categories used to group expenditures and budget figures on the PPRs do not map directly to the industry sectors used in IMPLAN.[152] In order to run the IMPLAN input-output analysis, ASR mapped PPR budget categories to IMPLAN industry sectors. For example, the PPR budget category Personnel does not exist as an IMPLAN industry sector. ASR, in consultation with NTIA, assumed that this budget category mapped best to the Office Administrative Services industry sector. In some cases, ASR mapped PPR budget categories to several industry sectors to reflect the complexity of the funds spent under that PPR budget category. In these cases, NTIA provided ASR with additional budget detail to help determine the proportion of funds under a PPR budget category to allocate funds across multiple industry sectors. NTIA determined these proportions by analyzing detailed budgets for a sample of grants provided by NTIA. Grants were first broken out by type (PCC and SBA or CCI), and then by total budget size (small, medium, and large) with different dollar thresholds for the two grant types. Table B defines the total budget thresholds used to determine grant size.

Table B. Grant Size

Grant Type	Total Budget Dollar Ranges		
	Small	Medium	Large
PCC and SBA	Less than $5 million	At least $5 million and less than $10 million	At least $10 million
CCI	Less than $15 million	At least $15 million and less than $50 million	At least $50 million

Table C and Table D include the IMPLAN industry sectors used for each PPR budget category for PCC and SBA grants and CCI grants, respectively. ASR completed the mapping of grant dollars following the detailed mapping in these tables.

Table C. PPR Expenditure to Industry Sector Mapping for PCC and SBA Grants

PPR Field	IMPLAN Sector Code	Industry Sector Description	Allocation Percentage of PPR Budget Field		
			Small	Medium	Large
Other	384	Office Administrative Services	18.60%	27.94%	22.84%
	358	Insurance Agencies, Brokerages, and Related Activities	4.26%	6.44%	3.14%
	383	Travel Arrangement and Reservation Services	1.75%	0.63%	0.29%
	365	Commercial and Industrial Machinery and Equipment Rental and Leasing	1.27%	23.41%	14.71%
	313	Office Supplies (Except Paper) Manufacturing	2.43%	5.26%	0.39%
	367	Legal Services	0.00%	18.56%	9.75%
	373	Other Computer Related Services, Including Facilities Management	5.39%	2.04%	2.66%
	374	Management, Scientific, and Technical Consulting Services	21.29%	7.32%	31.29%
	393	Other Educational Activities	21.29%	1.36%	11.13%
	36	Construction of Other New Nonresidential Structures	0.00%	0.93%	0%
	360	Real Estate	11.86%	3.05%	1.90%
	377	Advertising and Related Services	11.86%	3.05%	1.90%
Personnel	384	Office Administrative Services	100%	100%	100%
Supplies	313	Office Supplies (Except Paper) Manufacturing	100%	100%	100%
Contractual	367	Legal Services	0.98%	0%	14.11%
	373	Other Computer Related Services, Including Facilities Management	13.65%	0%	23.76%
	374	Management, Scientific, and Technical Consulting Services	20.23%	100%	62.12%
	393	Other Educational Activities	65.14%	0%	0%
Equipment	365	Commercial and Industrial Machinery and Equipment Rental and Leasing	100%	100%	100%
Indirect Costs	358	Insurance Agencies, Brokerages, and Related Activities	100%	100%	100%
Fringe Benefits	358	Insurance Agencies, Brokerages, and Related Activities	100%	100%	100%
Construction	36	Construction of Other New Nonresidential Structures	100%	100%	100%
Travel	383	Travel Arrangement and Reservation Services	100%	100%	100%

Table D. PPR Expenditure to Industry Sector Mapping for CCI Grants

PPR Field	IMPLAN Sector Code	Industry Sector Description	Allocation Percentage of PPR Budget Field		
			Small	Medium	Large
Construction	36	Construction of Other New Nonresidential Structures	100%	100%	100%
Equipment	365	Commercial and Industrial Machinery and Equipment Rental and Leasing	100%	100%	100%
Land Structures, etc.	360	Real Estate	100%	100%	100%
Architectural & Engineering Fees	369	Architectural, Engineering, and Related Services	100%	100%	100%
Admin & Legal	367	Legal Services	50%	50%	50%
	384	Office Administrative Services	50%	50%	50%
Other	351	Telecommunications	9.98%	72.41%	6.68%
	384	Office Administrative Services	0.00%	5.09%	0.42%
	383	Travel Arrangement and Reservation services	0.00%	4.21%	0.00%
	365	Commercial and Industrial Machinery and Equipment Rental and Leasing	90.02%	18.29%	84.34%
	36	Construction of Other New Nonresidential Structures	0.00%	0.00%	8.56%
Site Work	39	Maintenance and Repair Construction of Nonresidential Maintenance and Repair	100%	100%	100%
Other Architectural & Engineering	369	Architectural, Engineering, and Related Services	100%	100%	100%
Project Inspection Fees	369	Architectural, Engineering, and Related Services	100%	100%	100%
Demolition & Removal	39	Maintenance and Repair Construction of Nonresidential Maintenance and Repair	100%	100%	100%
Relocation Expenses	335	Truck Transportation	100%	100%	100%

ASR used the "ExpenditureAnalysis_IMPLANprep" SAS program to map PPR dollars to IMPLAN sectors. After grouping, aggregating, and mapping dollars according to the logic contained in Table B, Table C, and Table D, the SAS program outputs twelve Excel spreadsheets formatted for use by IMPLAN. Each output Excel file includes a dollar summary of a unique combination of dollar type (federal budget, federal spent, total budget, or total spent) and grant type (PCC and SBA, CCI, or All Grants). For example, federal budgets for PCC and SBA projects.

8.2 Run the IMPLAN Analysis

After mapping expenditure and budget summary tables to industry sectors, ASR implemented the input-output analysis. ASR imported the twelve summary tables generated by the

"ExpenditureAnalysis_IMPLANprep" SAS program (described in the previous section) into the "BTOP IMPLAN Model" IMPLAN database file.

After importing each dollar type and grant type summary table into IMPLAN, ASR ran the IMPLAN analysis on each individual scenario to generate the short-term economic impact estimates attributed to the particular summary table. IMPLAN then calculates the estimates of direct, indirect, and induced impacts. IMPLAN analyzes summary tables individually, which allowed ASR to quantify the short-term economic impacts of particular scenarios, as well as BTOP as a whole. For example, by analyzing the summary table containing actual federal dollar expenditures for PCC and SBA grants through the first calendar quarter of 2013, ASR was able to estimate the short-term economic impacts of federal dollars spent by PCC and SBA grants through the end of that quarter.

8.3 Convert Employment Estimates

The IMPLAN input-output model calculates employment estimates that include all full-time, part-time, and temporary employment. Alone, these job estimates do not indicate the number of hours worked or the portion that represents full- or part-time employment. In order to standardize these figures, ASR converted all employment estimates to Full-Time Equivalents (FTE).

IMPLAN provides additional data that contains the FTE-to-Employment ratio for each industry sector. The "Convert_IMPLANemployment_to_FTE" spreadsheet contains these ratios. ASR used the "Convert_IMPLAN_to_FTE" SAS program to convert the seven IMPLAN employment estimate spreadsheets into seven FTE estimate spreadsheets based on the FTE-to-Employment ratios. Each IMPLAN employment estimate spreadsheet is prefixed with "IMPLANoutput," while each FTE estimate spreadsheet is prefixed with "FTE."

Section 9. Final Report Section 4. Long-Term Impacts

This section includes descriptions of all files, programs, and processes used to summarize the activities, effects, impacts, or benefits of BTOP for the *Final Report*. Summary tables and figures are created from prepared data and input files, statistical analysis results, and combinations thereof. Each summary tables and figure listed below includes the following:

- One R script to load one or more sets of prepared data, read input files, and load statistical analysis results. Loaded and imported data are prepared as necessary to produce the summary tables and figures included in the *Final Report*. R scripts are sets of instructions written in the R statistical programming language than can be executed in the R statistical package.
- One output file for every summary table and figure. Tables are saved as Excel files; tables were copied from the Excel files and pasted into the *Final Report*. Figures are saved as PNG files; figures were inserted as pictures into the *Final Report*.

Every script has a header that defines the parent directory and calls the shared source code scripts described in Section 2. Users are required to redefine the parent directory to the correct location on their computer to reproduce the manipulation steps and output. This can be done in batch using the "Set Parent Directory in All Script Files" R script described in Section 1.3.

The following subsections describe the R scripts, in the order of presentation, used to create summary tables and figures for *Section 4. Long-Term Impacts* in the *Final Report*. All scripts and generated files discussed below are located in the "4. Long-Term Impacts" folder.

9.1 Final Report Section 4.1 Summary of CCI Outcomes

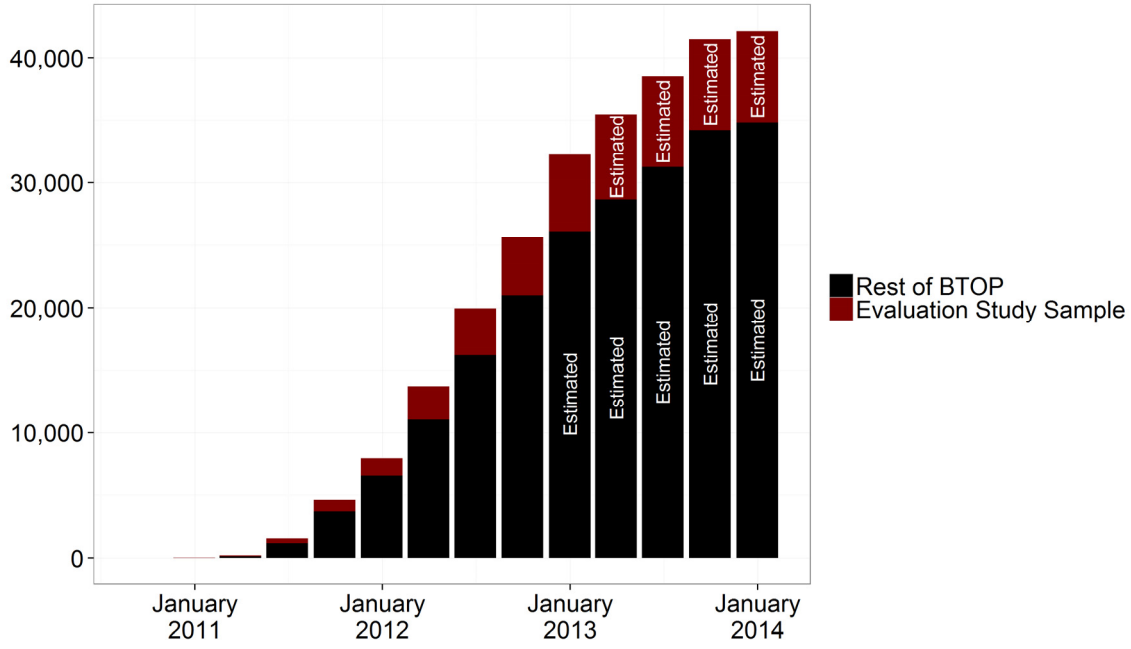
9.1.1 Final Report Figure 4. Cumulative New Fiber Miles Deployed by Quarter, all CCI Projects

Final Report Figure 4 presents cumulative total new fiber miles deployed by quarter for CCI grants in the evaluation study sample and the rest of BTOP. The "Figure 4" R script reproduces this figure. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "CCI Progress" is a set of four distinct summary tables containing grant-level network and CAI data. This script uses the "Network Progress" table only, derived from PPRs.[153] Section 4.6 provides a complete description of all four tables.

The prepared data above are loaded. "Grants" and "Network Progress" are joined by award number, and PCC, SBA, excluded, and defunded grants are dropped. New fiber miles deployed are aggregated by quarter and evaluation study sample status. A stacked bar chart, with fill colors representing the evaluation study sample and the rest of BTOP, is created from the aggregated data. Any aggregated groups that contain values carried over from previous quarters are labelled as "Estimated." The chart is written to the "Figure 4" PNG file. Figure 4 below displays the results of the R script.

Figure 4. Cumulative New Fiber Miles Deployed by Quarter, all CCI Projects



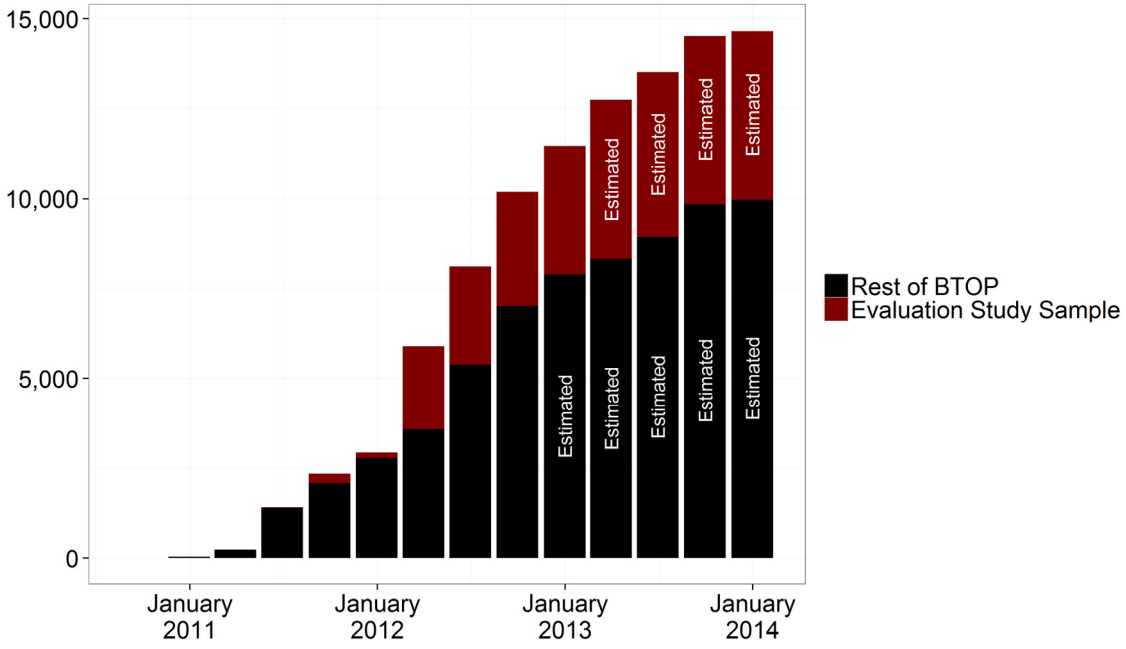
9.1.2 Final Report Figure 5. Interconnection Points by Quarter, all CCI Projects

Final Report Figure 5 summarizes cumulative interconnection points activated by quarter for CCI grants in the evaluation study sample and the rest of BTOP. The "Figure 5" R script reproduces this figure. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "CCI Progress" is a set of four distinct summary tables containing grant-level network and CAI data. This script uses the "Network Progress" table only, derived from PPRs.[154] Section 4.6 provides a complete description of all four tables.

The prepared data above are loaded. "Grants" and "Network Progress" are joined by award number, and PCC, SBA, excluded, and defunded grants are dropped. Interconnection points are aggregated by quarter and evaluation study sample status. A stacked bar chart, with fill colors representing the evaluation study sample and the rest of BTOP, is created from the aggregated data. Any aggregated groups that contain values carried over from previous quarters are labelled as "Estimated." The chart is written to the "Figure 5" PNG file. Figure 5 below presents the results of the R script.

Figure 5. Interconnection Points by Quarter, all CCI Projects



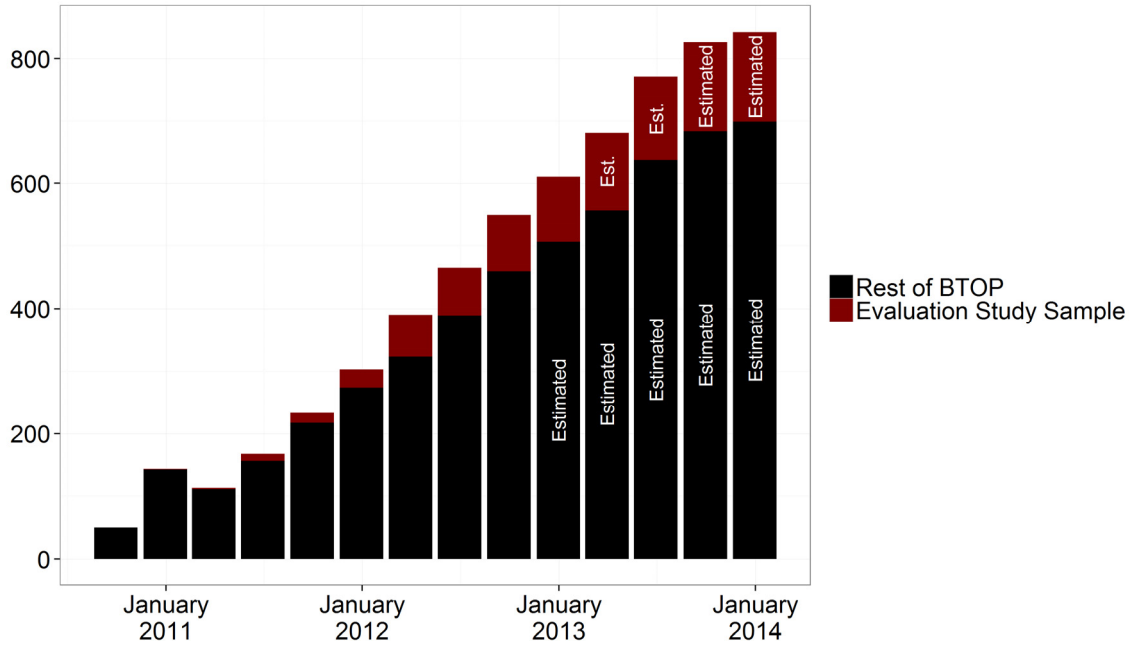
9.1.3 Final Report Figure 6. Cumulative Signed Third-Party Service Provider and Broadband Wholesaler Agreements, all CCI Projects

Final Report Figure 6 summarizes cumulative signed service agreements by quarter for CCI grants in the evaluation study sample and the rest of BTOP. The "Figure 6" R script reproduces this figure. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "CCI Progress" is a set of four distinct summary tables containing grant-level network and CAI data. This script uses the "Network Progress" table only, derived from PPRs.[155] Section 4.6 provides a complete description of all four tables.

The prepared data above are loaded. "Grants" and "Network Progress" are joined by award number, and PCC, SBA, excluded, and defunded grants are dropped. Signed agreements are aggregated by quarter and evaluation study sample status. A stacked bar chart, with fill colors representing the evaluation study sample and the rest of BTOP, is created from the aggregated data. Any aggregated groups that contain values carried over from previous quarters are labelled as "Estimated." The chart is written to the "Figure 6" PNG file. Figure 6 below presents the results of the R script.

Figure 6. Cumulative Signed Third-Party Service Provider and Broadband Wholesaler Agreements, all CCI Projects



9.2 Final Report Section 4.2 Effect of BTOP on Broadband Availability

9.2.1 Final Report Table 11. Potential Control Counties in the United States

Final Report Table 11 summarizes the process used by ASR to identify potential control counties in the United States. The "Table 11" R script reproduces this table. The script uses the following prepared data:

- "Treatment and Potential Control Counties" identifies every county included in NBM data according to the steps used to identify potential control counties. These data are derived from data published by the USDA, ProPublica, and NTIA, as well as data collected by ASR from individual grantees and other public sources.[156] Section 4.28 provides a full description of "Treatment and Potential Control Counties."

The script loads the prepared data discussed above. Variables representing "All counties and equivalents in the United States," "Less counties in Alaska, Hawaii, and outlying areas," "Less counties not in proposed BTOP service area," and "Less counties in awarded BTOP or BIP grant service area" are reshaped and aggregated. The results are then tabulated as saved to the "Table 11" Excel file. Table 11 below presents the results of the R script.

Table 11. Potential Control Counties in the United States

All counties and equivalents in the United States	**3,234**	
Less counties in Alaska, Hawaii, and outlying areas	3,109	
Less counties not in proposed BTOP service area	2,640	
Less counties in awarded BTOP or BIP grant service area	**884**	**Potential Controls**

9.2.2 Final Report Table 12. Estimated BTOP Effect on Broadband Availability

Final Report Table 12 summarizes the base case (NTIA definition of broadband) effect of BTOP on broadband availability. The table shows overall broadband availability rates in the treatment and selected control counties in 2011 and 2013, the differences between 2011 and 2013, and the difference-in-differences. The "Table 12" R script reproduces this table. The script uses the following statistical analysis results to generate the table:

- "Effect on Availability" uses data on population, demographics, and broadband availability and the matched pairs identified in ASR's "Matching" statistical analysis to estimate the effect of BTOP on broadband availability in the evaluation study sample service area. Section 5.2 provides a full description of "Effect on Availability."

The script loads the statistical analysis results discussed above. One of the tables included in these results presents the treatment and control rates and differences and difference-in-differences for all twelve matching scenarios. This table is subset to the base case (NTIA definition, no adjustments made to broadband data, and matching based on primary variables only), reshaped, and formatted. The results are then tabulated as saved to the "Table 12" Excel file. Table 12 below presents the results of the R script.

Table 12. Estimated BTOP Effect on Broadband Availability

Baseline Availability	Treatment Group	Control Group
June 30, 2011	91.53%	92.28%
June 30, 2013	94.40%	93.16%
Difference	2.87%	0.88%
Difference-in-differences	**2.00%**	

9.3 Final Report Section 4.4 Reduction in Broadband Costs

9.3.1 Final Report Table 13. Subscription Speed and Pricing Changes

Final Report Table 13 summarizes the monthly price per Mbps, connection speed, and total monthly cost of broadband subscriptions before BTOP and after BTOP. The "Table 13" R script reproduces this table. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "CCI CAIs by Category" summarizes the number of connected CAIs by type for all BTOP CCI grants derived from PPRs.[157] Section 4.3 provides a full description of "CCI CAIs by Category."
- "CCI Progress" is a set of four distinct summary tables containing grant-level network and CAI data. This script uses the "CAI Totals" table derived from PPRs.[158] Section 4.6 provides a complete description of "CAI Totals" and the remaining "CCI Progress" tables.
- "CCI Speed and Pricing" is a table containing CAI broadband subscription speeds and costs before and after BTOP. ASR collected these data during site visit interviews with grantees, CAIs, and service providers as part of the CCI case study report process. Section 4.7 provides a full description of "CCI Speed and Pricing."

The script loads the prepared data discussed above. The number of CAIs by institution type in the evaluation study sample is estimated using the "CAI Totals" table. This is accomplished through the following steps:

- "Grants" and "CAI Totals" are joined by award number and filtered to the twelve evaluation study sample CCI grants, and the total number of connected CAIs in the evaluation study sample is calculated
- "Grants" and "CCI CAIs by Category" are joined by award number and filtered to the twelve evaluation study sample CCI grants. The number of categorized connected CAIs is aggregated by institution type to obtain each institution type's total in the evaluation study sample categorization.
- In the aggregated by institution type table, the estimated total number connected by institution type in the evaluation study sample is estimated as the percentage of each institution type out of the total number categorized multiplied by the total number of connected CAIs in the evaluation study sample.

For collected speed and pricing data, only paired values are used in this table. For example, a CAI must have data for connection speed before BTOP and speed after BTOP to be included in the summary for either. This rule also applies to the monthly price per Mbps and the monthly total cost variables. Data are reshaped and label variables are formatted. The data are then summarized by finding the median monthly price per Mbps, median connection speed, and median total monthly cost for each CAI institution type and for all institutions. The percentage changes in median price per Mbps per month are calculated. The estimated number of CAIs connected by institution type is then joined to this table by institution type. The table is then formatted and written to the "Table 13" Excel file. Table 13 below presents the results of the R script.

Table 13. Subscription Speed and Pricing Changes

Institution Type	Connected in Evaluation Study Sample	Collected Speeds			Collected Prices					
		N	Median Speed (Mbps)		N	Median Total Monthly Cost		Median Price per Mbps per Month		
			Before	After		Before	After	Before	After	Pct. Change
Schools (K-12)	2,157	13	20	100	6	$1,150	$1,240	$293	$14	95%
University, College, or Other Postsecondary	347	47	45	1,000	25	$1,500	$1,500	$56	$2	96%
Medical/ Healthcare	930	7	6	100	6	$3,350	$900	$387	$16	96%
Library	603	7	3	20	5	$600	$300	$233	$15	94%
Other Community Support	578	8	26	550	5	$2,800	$2,500	$111	$5	95%
Public Safety	578	4	2	525	0	N/A	N/A	N/A	N/A	N/A

9.3.2 Final Report Table 14. Annual Extrapolated CAI Cost Differences, all CCI Projects

Final Report Table 14 summarizes the extrapolated cost differences to all CAIs connected by BTOP CCI grants due to improved connection speeds and reduced costs. The "Table 14" R script reproduces this analysis and the resulting table. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "CCI CAIs by Category" summarizes the number of connected CAIs by type for all BTOP CCI grants derived from PPRs.[159] Section 4.3 provides a full description of "CCI CAIs by Category."
- "CCI Progress" is a set of four distinct summary tables containing grant-level network and CAI data. This script uses the "CAI Totals" table derived from PPRs.[160] Section 4.6 provides a complete description of "CAI Totals" and the remaining "CCI Progress" tables.
- "CCI Speed and Pricing" is a table containing CAI broadband subscription speeds and costs before and after BTOP. ASR collected these data during site visit interviews with grantees, CAIs, and service providers as part of the CCI case study report process. Section 4.7 provides a full description of "CCI Speed and Pricing."

The script loads the prepared data discussed above. ASR uses the differences between the median values of connection speed and price per Mbps across institution types before and after BTOP at CAIs connected by CCI grants in the evaluation study sample. ASR assumes that differences between medians in the collected data are representative of the differences across all connected CAIs.

"CCI Speed and Pricing" data are used to estimate differences between before- and after-BTOP costs and speeds. Only paired values are used in these estimations. For example, a CAI must have data for connection speed before BTOP and speed after BTOP to be included in the summary for either. This rule also applies to the monthly price per Mbps variables. The median before and after connection speed and monthly price per Mbps are calculated by institution type.

ASR makes two adjustments. The first adjustment is made to two different institution types: the median speeds after BTOP for Other Community Support and Public Safety institutions are capped at 100 Mbps to account for the extremely high speeds sometimes reported for these institutions types during site visits. The median collected speeds for Other Community Support and Public Safety institutions led to extrapolated price differences that were much higher than expected given ASR's observation of institutions in these categories. ASR selected 100 Mbps because it is the median collected speed for all Schools (K-12), Medical/Healthcare, and Library institutions. The second adjustment is made because no Public Safety institutions connected by evaluation study grantees shared price data with ASR. Therefore, ASR uses the overall median price per Mbps for these institutions.

The differences between before- and after-BTOP speeds and costs are calculated as the differences between before- and after-BTOP medians, using the adjustments described above. For all categories of CAI except Public Safety and Other Community Support, the per-CAI cost difference is given by the formula below:

per-CAI cost difference =
(category median post-BTOP speed – category median pre-BTOP speed) ×
(category median pre-BTOP price per Mbps per month – category median post-BTOP price per Mbps per month) × 12

For Other Community Support CAIs, the per-CAI cost difference is given by the formula below:

per-CAI cost difference =
(100 Mbps – category median pre-BTOP speed) ×
(category median pre-BTOP price per Mbps per month – category median post-BTOP price per Mbps per month) × 12

For Public Safety CAIs, the per-CAI cost difference is given by the formula below:

per-CAI cost difference =
(100 Mbps – category median pre-BTOP speed) ×
(overall median pre-BTOP price per Mbps per month – overall median post-BTOP price per Mbps per month) × 12

ASR then extrapolated the total cost difference due to CCI grants by multiplying the per-CAI cost difference by the estimated number of CAIs, as given by:

total cost difference = per-CAI cost difference × number of CAIs in category

"Grants" and "CAI Totals" are joined by award number, and PCC, SBA, defunded, and excluded grants are dropped. The total number of connected institutions is calculated. "CCI CAI by Category" data are aggregated across all grants. The percentage of total connected CAIs is calculated by institution type. ASR assumes that, although the CAI data included in this table do not account for all connected CAIs reported, the distribution of CAIs by institution type in the table is representative of the distribution of all CAIs by institution type. The total number of connected institutions is then multiplied by this distribution to estimate the number of CAIs in each institution type.

Estimated CAIs and estimated annual cost savings per CAI are joined by institution type. The total cost difference to each institution type is extrapolated as:

total cost difference = per-CAI cost difference × number of CAIs in category The extrapolations are aggregated to provide estimates of the number of CAIs, total annual cost difference, and total annual cost difference per CAI across all institution types. The table of extrapolations is then formatted. The results are written to the "Table 14" Excel file. Table 14 below presents the results of the R script.

Table 14. Annual Extrapolated CAI Cost Differences, all CCI Projects

Institution Type	Percent of All CAIs	Number of CAIs	Total Cost Difference (Millions)	Per CAI Cost Difference
School (K-12)	36%	7,726	$2,072	$268,136
University, College, or Other Postsecondary	7%	1,498	$928	$619,477
Medical/Healthcare	12%	2,640	$1,104	$418,112
Library	7%	1,515	$67	$44,540
Other Community Support	23%	4,951	$471	$95,087
Public Safety	14%	2,910	$374	$128,574
All institutions	**100%**	**21,240**	**$5,016**	**$236,151**

Section 10.Final Report Section 6. Progress towards Recovery Act Goals

This section includes descriptions of all files, programs, and processes used to summarize the activities, effects, impacts, or benefits of BTOP for the *Final Report*. Summary tables and figures are created from prepared data and input files, statistical analysis results, and combinations thereof. Each summary tables and figure listed below includes the following:

- One R script to load one or more sets of prepared data, read input files, and load statistical analysis results. Loaded and imported data are prepared as necessary to produce the summary tables and figures included in the *Final Report*. R scripts are sets of instructions written in the R statistical programming language than can be executed in the R statistical package.
- One output file for every summary table and figure. Tables are saved as Excel files; tables were copied from the Excel files and pasted into the *Final Report*. Figures are saved as PNG files; figures were inserted as pictures into the *Final Report*.

Every script has a header that defines the parent directory and calls the shared source code scripts described in Section 2. Users are required to redefine the parent directory to the correct location on their computer to reproduce the manipulation steps and output. This can be done in batch using the "Set Parent Directory in All Script Files" R script described in Section 1.3.

The following subsections describe the R scripts, in the order of presentation, used to create summary tables and figures for *Section 6. Progress towards Recovery Act Goals* in the *Final Report.* All scripts and generated files discussed below are located in the "6. Progress towards Recovery Act Goals" folder.

10.1 Final Report Section 6.1 Improve Access to Unserved and Underserved Areas of the Country

10.1.1 Final Report Figure 7. Estimated Effect of BTOP on Broadband Availability in the CCI Evaluation Study Sample Service Area for Different Populations of Interest

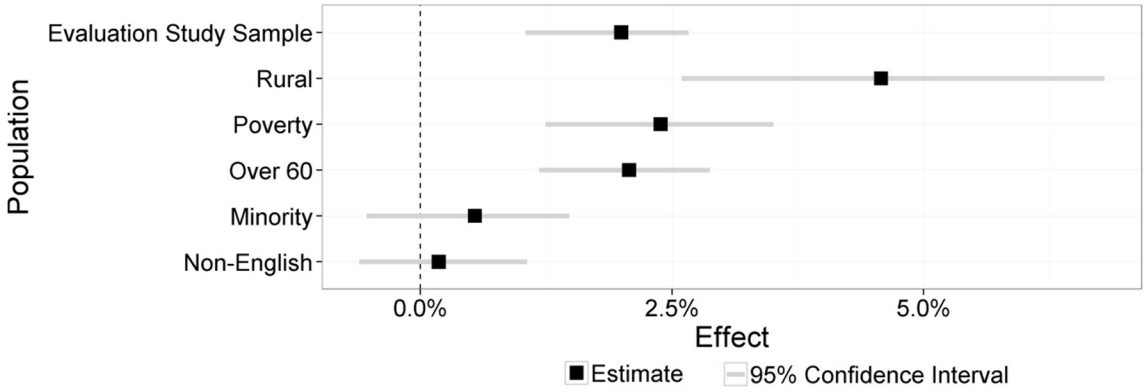
Final Report Figure 7 presents the results of ASR's analysis of the changes in broadband availability rates among vulnerable populations. The "Figure 7" R script reproduces this figure using the following statistical analysis results:

- "Incidence Analysis" uses NBM data on population, demographics, and broadband availability; Census Bureau ACS (2006-2010) demographic data; and the matched pairs identified in ASR's "Matching" statistical analysis to estimate the effect of BTOP on broadband availability among vulnerable populations in the evaluation study sample service area. Section 5.3 provides a full description of "Incidence Analysis."
- "Effect on Availability" uses data on population, demographics, and broadband availability and the matched pairs identified in ASR's "Matching" statistical analysis to estimate the effect of BTOP on broadband availability in the evaluation study sample service area. Section 5.2 provides a full description of "Effect on Availability."

The script loads the statistical analysis results discussed above. Estimates and confidence intervals for the effect of BTOP on broadband availability among vulnerable populations are concatenated with estimates and confidence intervals for the effect on availability among the total population. Only estimates and confidence intervals for the primary set of matches using the unadjusted form of availability data based on the NTIA definition of broadband are used in this figure.

After the data are concatenated, a forest plot is created to present the estimates and confidence intervals for the different populations. This figure is then saved as the "Figure 7" PNG file. Figure 7 below presents the results of the R script. ASR used an image editor to change the orientation of the legend from vertical to horizontal.

Figure 7. Estimated Effect of BTOP on Broadband Availability in the CCI Evaluation Study Sample Service Area for Different Populations of Interest



10.2 Final Report Section 6.2 Broadband Education, Awareness, Training, Access, Equipment, and Support

10.2.1 Final Report Table 15. New and Upgraded PCCs

Final Report Table 15 summarizes the PCCs established and upgraded for PCC grants in the evaluation study sample and all of BTOP. The "Table 15" R script reproduces this table. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "PCC Centers Established and Improved" is a table containing cumulative numbers of PCCs established and improved by grant, year, and category of PCC derived from PPRs.[161] Section 4.19 provides a full description of "PCC Centers Established and Improved."

The script loads the prepared data discussed above. "Grants" and "PCC Centers Established and Improved" are joined by award number and defunded grants are dropped. Established and improved PCC figures are combined. Data are tabulated for each institution type and evaluation study sample status, calculating total PCCs established and improved. This tabulation is written to the "Table 15" Excel file. Table 15 below presents the results of the R script. *Final Report Table 15* presents the estimates for the study sample and all of BTOP.

Table 15. New and Upgraded PCCs

Institution Type	Evaluation Study Sample	All
Schools (K-12)	1	118
Libraries	289	2,120
Community Colleges	50	120
Universities and Colleges	1	8
Medical or Healthcare Facilities	1	62
Public Safety Entities	0	4
Job Training and/or Economic Development Institutions	32	163
Other Community Support (Governmental)	74	410
Other Community Support (Non-Governmental)	92	300
Total	**540**	**3,305**

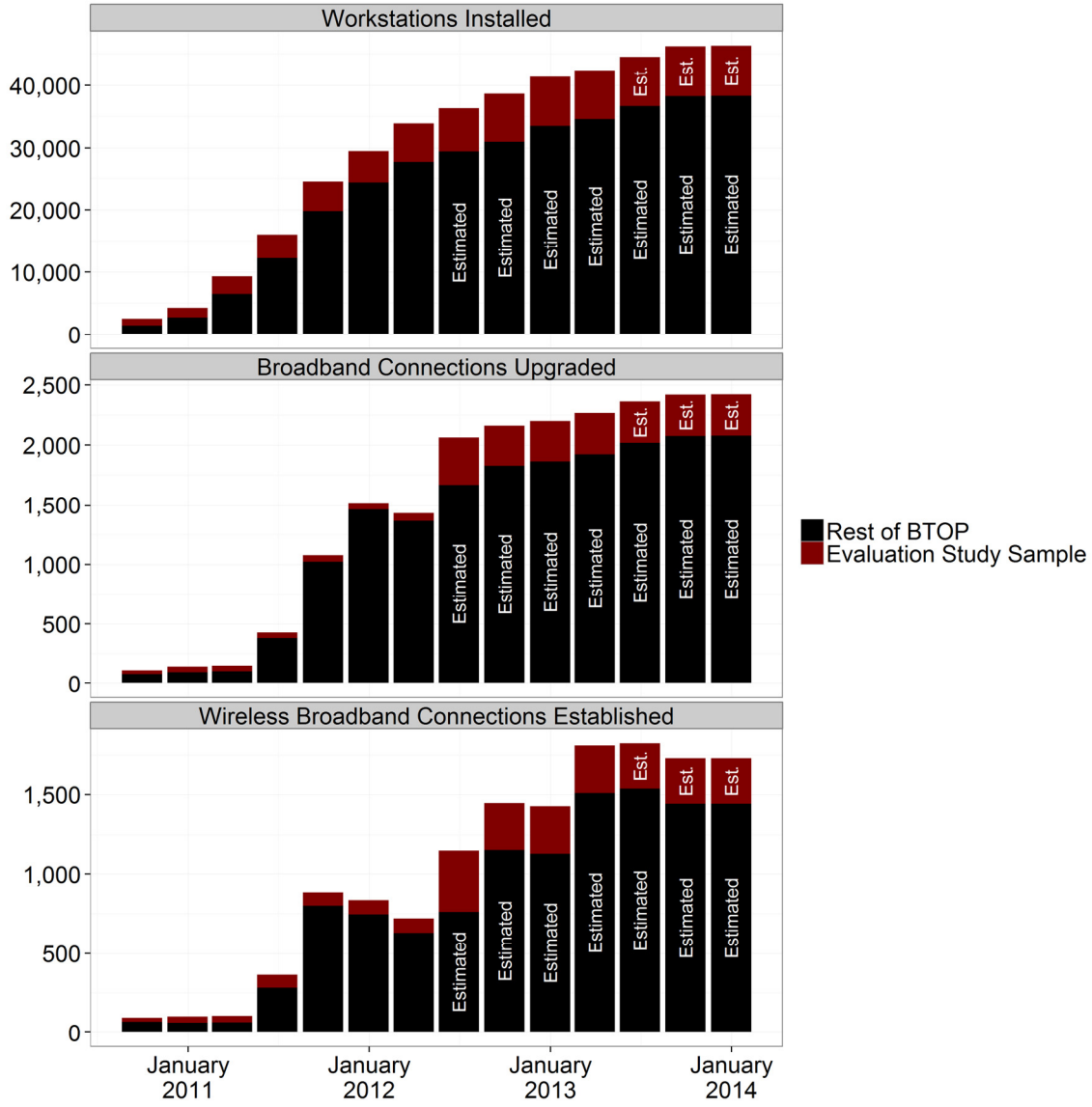
10.2.2 Final Report Figure 8. Cumulative PCC Hardware Installations and Upgrades

Final Report Figure 8 summarizes cumulative workstations installed, broadband connections upgraded, and wireless broadband connections established for PCC grants in the evaluation study sample and the rest of BTOP. The "Figure 8" R script reproduces this figure. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "PCC Equipment" is a table containing cumulative numbers of PCCs established and improved by grant, year, and category of PCC derived from PPRs.[162] Section 4.20 provides a full description of "PCC Equipment."

The script loads the prepared data discussed above. "Grants" and "PCC Equipment" are joined by award number, and SBA, CCI, and defunded grants are dropped. Cumulative workstations installed, broadband connections upgraded, and wireless broadband connections established are selected and aggregated by quarter and evaluation study sample status. Data are reshaped to allow a separate panel to be created for each variable. A stacked bar chart, with fill colors representing the evaluation study sample and the rest of BTOP, is created from the aggregated data with separate panels for cumulative workstations installed, broadband connections upgraded, and wireless broadband connections established. Any aggregated groups that contain values carried over from previous quarters are labelled as "Estimated." The chart is written to the "Figure 8" PNG file. Figure 8 below presents the results of the R script.

Figure8. Cumulative PCC Hardware Installations and Upgrades



10.2.3 Final Report Table 16. SBA Institutions

Final Report Table 16 summarizes the different institution types hosting programs for SBA grants in the evaluation study sample and all of BTOP. The "Table 16" R script reproduces this table. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "SBA Locations" is a table containing location data for all SBA program locations derived from the CAC.[163] Section 3.9 provides a full description of "SBA Locations."

The script loads the prepared data discussed above. "SBA Locations" data are used for this table. Data are tabulated for each institution type and evaluation study sample status, calculating total SBA locations. This tabulation is written to the "Table 16" Excel file. Table 16 below presents the

results of the R script, is presented below. *Final Report Table 16* presents the estimates for the study sample and all of BTOP.

Table 16. SBA Institutions

Institution Type	Evaluation Study Sample	All
Community-Based Organization	81	190
Community College	59	104
For-Profit Organization	14	25
Government Facility	83	134
Institution of Higher Education	17	92
Library	40	349
Medical or Healthcare Provider	6	370
Nonprofit Organization	256	854
Public Housing	9	134
Public Safety	50	56
School (K-12)	108	615
Tribal	1	339
Other	1	87
All institutions	**725**	**3,349**

10.2.4 Final Report Table 18. PCC and SBA Training Hours

Final Report Table 18 summarizes grant-level training data. The "Table 18" R script reproduces this table. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "PCC and SBA Training Hours and Participants" is a table containing a grant-level summary of participants and training hours for each focus area and open lab access hours derived from PPRs.[164] Section 4.17 provides a full description of "PCC and SBA Training Hours and Participants."

The script loads the prepared data discussed above. "Grants" and "PCC and SBA Training Hours and Participants" are joined by award number and CCI and defunded grants are dropped. Open Lab hours are extracted to a separate table. Training hours data are aggregated for all of BTOP by grant type and focus area, and by grant type for all focus areas. Open Lab hours are aggregated for all of BTOP by grant type. The aggregated tables are combined, and total hours (PCC plus SBA) are calculated. These results are written to the "Table 18" Excel file. Table 18 below presents the results of the R script.

Table 18. PCC and SBA Training Hours

Focus Area	PCC	SBA	Total
Workforce and Economic Development	2,067,847	362,593	2,430,440
Education and Training	2,798,246	701,815	3,500,061
Healthcare	713	2,226,264	2,226,977
Quality of Life/Civic Engagement	4,175	9,013	13,188
Digital Literacy	5,231,319	5,901,255	11,132,574
Other	1,431,601	98,818	1,530,419
Total	**11,533,901**	**9,299,758**	**20,833,659**
Open Lab Access	974,721	3,607	978,328

10.3 Final Report Section 6.3 Public Safety Agencies

Section 6.3 Public Safety Agencies of the *Final Report* contains CCI activity summary figures that are not provided in tables elsewhere in the *Final Report*. The "Section 6.3" script derives these figures. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "CCI CAIs Passed" is a table of the cumulative number of CAIs connected or passed by grantee, CAI category, and year derived from APRs.[165] Section 4.4 provides a complete description of "CCI CAIs Passed."

The prepared data above are loaded and joined by award number. The joined data are subset to the last available year, and defunded and excluded grants are filtered out. CAI counts are then aggregated by CAI category. The total number of CAIs and percentage by CAI category are calculated. The results are printed to the screen and are not saved to any files.

10.4 Final Report Section 6.4 Demand for Broadband, Economic Growth, and Job Creation

10.4.1 Final Report Section 6.4.1 Demand for Broadband

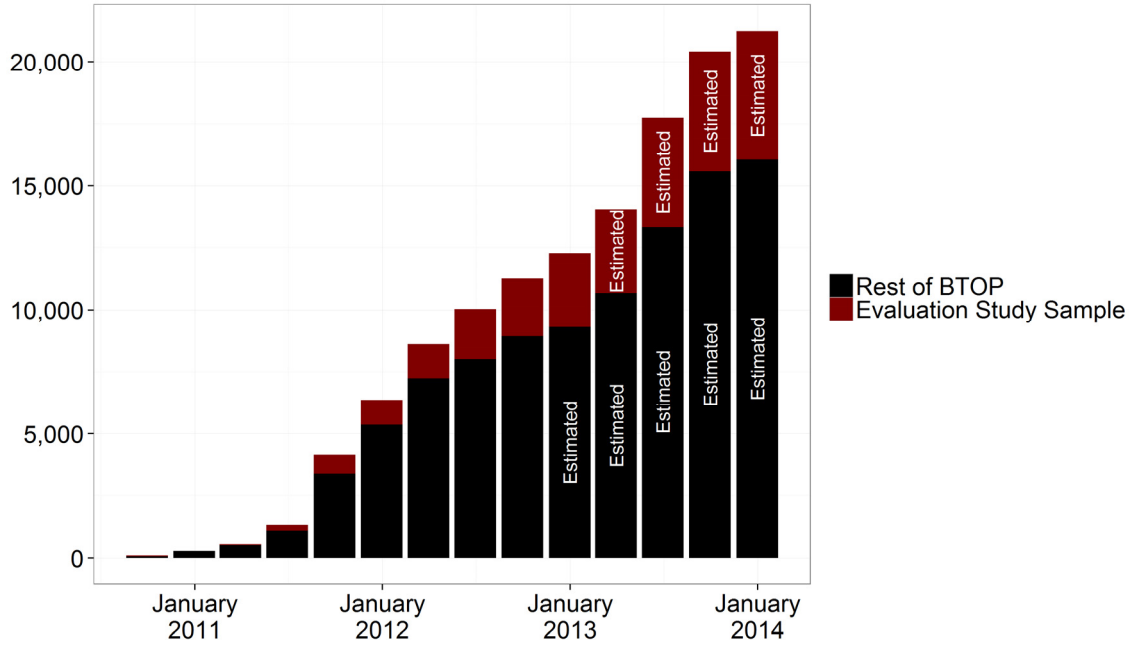
10.4.1.1 Final Report Figure 9. Cumulative CCI Community Anchor Institution Connections

Final Report Figure 9 summarizes the cumulative CAI connections made by CCI grants in the evaluation study sample and the rest of BTOP. The "Figure 9" R script reproduces this figure. The script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "CCI Progress" is a set of four distinct summary tables containing grant-level network and CAI data. This script uses the "CAI Progress" table only, derived from PPRs.[166] Section 4.6 provides a complete description of "CAI Progress."

The script loads the prepared data discussed above. "Grants" and "CCI Progress" are joined by award number, and PCC, SBA, excluded, and defunded grants are dropped. Cumulative CAI connections are aggregated by quarter and evaluation study sample status. A stacked bar chart, with fill colors representing the evaluation study sample and the rest of BTOP, is created from the aggregated data. Any aggregated groups that contain values carried over from previous quarters are labelled as "Estimated." The chart is written to the "Figure 9" PNG file. Figure 9 below presents the results of the R script.

Figure 9. Cumulative CCI Community Anchor Institution Connections



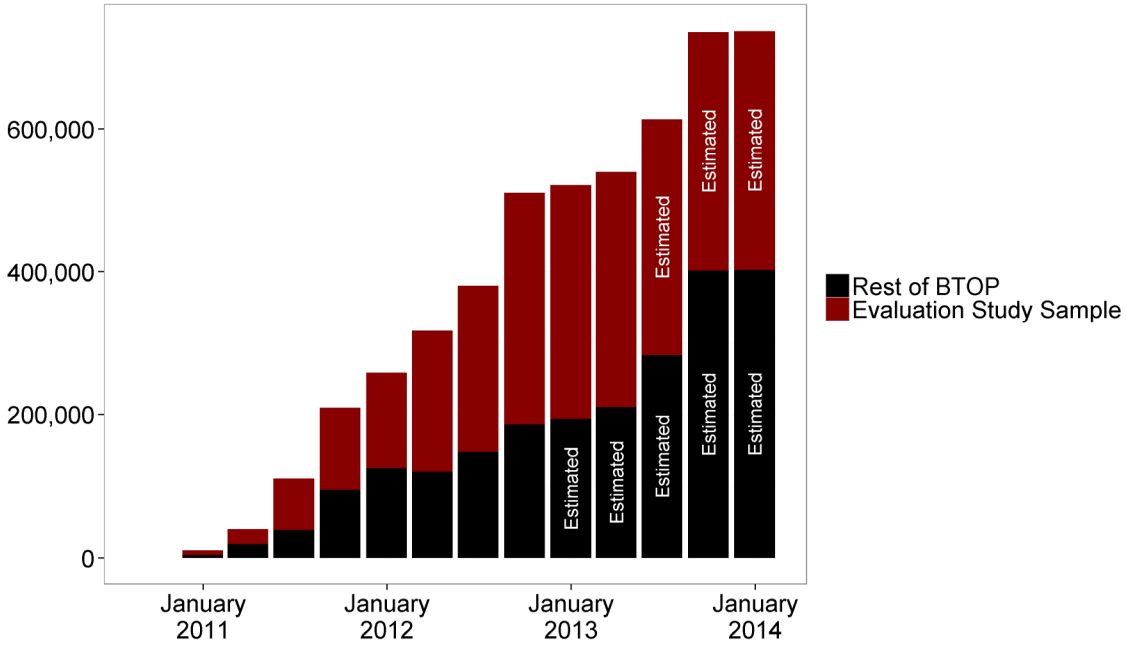
10.4.1.2 Final Report Figure 10. Cumulative New Household Subscribers (SBA Grantees)

Final Report Figure 10 summarizes the cumulative new household broadband subscribers due to SBA grants in the evaluation study sample and the rest of BTOP. The "Figure 10" R script reproduces the analysis and tables presented in the *Final Report*. The R script uses the following prepared data:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "SBA Subscribers" contains a grant-level table of cumulative new individual, household, and business subscribers by quarter due to SBA grant activity. The table relies on PPRs and publicly available data published by Census Bureau.[167] Section 4.25 provides a full description of "SBA Subscribers."

The script loads the prepared data discussed above. "Grants" and "SBA Subscribers" are joined by award number, and PCC, CCI, and defunded grants are dropped. Cumulative household subscribers are aggregated by quarter and evaluation study sample status. A stacked bar chart, with fill colors representing the evaluation study sample and the rest of BTOP, is created from the aggregated data. Any aggregated groups that contain values carried over from previous quarters are labelled as "Estimated." The chart is written to the "Figure 10" PNG file. Figure 10 below presents the results of the R script.

Figure 10. Cumulative New Household Subscribers (SBA Grantees)



Section 11.Final Report Appendix C. Quantitative Intermediate Impacts

This section includes descriptions of all files, programs, and processes used to estimate and summarize the quantitative intermediate impacts of BTOP for the *Final Report*. Impacts are estimated from prepared data and results found in the broadband literature and BTOP evaluation studies. Each impact listed below includes the following:

- One R script to load one or more sets of prepared data. These data are prepared as necessary, and sets of methodological steps are performed to estimate quantitative intermediate impacts. Manipulation steps follow to produce the results tables included in the *Final Report*. R scripts are sets of instructions written in the R statistical programming language than can be executed in the R statistical package.
- At least one results table, saved in Excel format, for each estimation script. Results tables were copied from the Excel files and pasted into the *Final Report*. Unlike the programs described in the earlier sections of this document, there can be multiple Excel output files associated with each script.

Every script has a header that defines the parent directory and calls the shared source code scripts described in Section 2. Users are required to redefine the parent directory to the correct location on their computer to reproduce the manipulation steps and output. This can be done in batch using the "Set Parent Directory in All Script Files" R script described in Section 1.3.

The following subsections describe the R scripts, in the order of presentation, used to estimate and summarize the quantitative intermediate impacts for *Appendix C. Quantitative Intermediate Impacts* in the *Final Report.* All scripts and generated files discussed below are located in the "C. Quantitative Intermediate Impacts" folder.

11.1 Final Report Appendix C.2 Obtained Employment (Impact ID: PCC.W.1)

11.1.1 Final Report Table 40. Total Estimated New Internet Job Searchers and Decreased Unemployment Duration (Weeks) Due to PCC Grants

Kuhn and Mansour (2013) estimated that unemployment durations are 24 percent shorter for Internet job searchers than offline job searchers.[168] ASR used grantee-reported data on PCC users and job-search training attendees to estimate the number of new Internet job searchers and decreased unemployment duration due to PCC grants.

The "PCC.W.1" R script uses the following prepared data to provide a quantitative estimate of PCC grants' impact on employment:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "PCC and SBA Job Search Training Summary" is a grant-level table summarizing the number of participants in grantee job-search-related training programs. "PCC and SBA Job Search Training Summary" relies on grantee-reported data from PPRs.[169] Section 4.15 provides a full

description of "PCC and SBA Job Search Training Summary." The following grant-level variables are used in these calculations:

  - job-search-related training participants

- "PCC and SBA Labor Force Statistics" is a grant-level table containing several labor force statistics. "PCC and SBA Labor Force Statistics" is derived from publicly available data published by BLS, Census Bureau, and DOL/BLS.[170] Section 4.16 provides a full description of "PCC and SBA Labor Force Statistics." The following grant-level variables are used in these calculations:
  - labor force participation rate
  - unemployment rate
  - average unemployment duration
  - minimum wage
- "PCC Unique Users" is a grant-level table containing the estimated number of unique users in the labor force over the course of the grant. "PCC Unique Users" relies on PPRs, the CAC, and survey results reported in Becker et al. (2010).[171] Section 4.22 provides a full description of "PCC Unique Users." The following grant-level variables are used in these calculations:
  - unique users

ASR used the R script "PCC.W.1" to load the prepared data discussed above and perform the calculations for this impact. The prepared data are joined by grant award number, resulting in a single table containing all grant-level statistics. This step also subsets the data to include only PCC grants.

ASR made the following assumptions to estimate this impact:

- The labor force participation rate and unemployment rate of a particular grant's PCC users equals that of the grant's service area population
- All PCC job-search-related training participants are unique and are new Internet job searchers.
- The average unemployment duration in a state uniformly describes unemployment periods of individuals in PCC service areas within that state

Based on the assumptions above, the following steps estimate benefits due to BTOP:

- New job searchers as a result of PCCs is estimated as unique users multiplied by the labor force participation rate and unemployment rate for each grant
- The number of job search training participants is added to PCC users searching for jobs to account for both types of grant activities that lead to users searching for jobs online (open PCC hours and job search training)
- Total reduction in unemployment duration is estimated as the number of new Internet job searchers due to PCC grants multiplied by the average unemployment duration in each service area and the study coefficient from Kuhn and Mansour (2013)
- Total economic benefit (in millions) from decreased unemployment duration due to PCC grants is estimated as the projected total decrease in unemployment duration multiplied by the minimum wage times forty (the minimum weekly wage for full-time work) in the grant's service area, then divided by one million

The estimates of new job searchers and total decrease in unemployment duration are aggregated for the evaluation study sample, PCC grants not in the evaluation study sample, and all PCC grants and written to the "PCC.W.1 - Table 40" Excel file. Table 40, as created by the R script, is presented below. Only the total estimates for all PCC projects were included in the *Final Report Table 40*. Other estimates were discussed in the text surrounding the table.

Table 40. Total Estimated New Internet Job Searchers and Decreased Unemployment Duration (Weeks) Due to PCC Grants

Measure	Estimate
New Internet job searchers	59,792
Decreased duration	331,796

The estimates of economic benefit are aggregated for the evaluation study sample, PCC grants not in the evaluation study sample, and all PCC grants and written to the "PCC.W.1 - Economic Benefit" Excel file. Table E, as created by the R script, is presented below. This table was not presented in the *Final Report,* but the estimates were discussed in the text relevant to the economic benefit of this impact.

Table E. Total Estimated Economic Benefit from Decreased Unemployment Duration Due to PCC Grants (Million USD, Annual)

Estimate	Evaluation Study Sample	Rest of BTOP	Total
Total Economic Benefit	17	77	94

11.2 Final Report Appendix C.3 Started or Grew Businesses (Impact ID: PCC.W.2)

11.2.1 Final Report Table 41. Total Estimated Number of Individuals Engaged in Any Entrepreneurial Activity through PCC Grants

Becker et al. (2010) investigated trends in public library patrons' use of free access to computers and the Internet.[172] The study found that that 7 percent of public-access computer users at public libraries across the United States engaged in "activities related to starting or managing a business of their own."[173] Additionally, 3.3 percent started a business using the library computers and 3.5 percent looked for new customers.[174] Nearly half of those who looked for new customers, 1.7 percent of all users, grew their business.[175] Additionally, PCC grantees noted that users and training participants engaged in entrepreneurial activities.[176] ASR used grantee-reported data (users and training attendees) to estimate the number of users engaging in entrepreneurial activities due to PCC grants.

The "PCC.W.2" R script uses the following prepared data to provide a quantitative estimate of PCC grants' impact on entrepreneurial activity:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "PCC and SBA Entrepreneurship Training Summary" is a grant-level table summarizing the number of participants in grantee-provided entrepreneurship training programs. The table relies on grantee-reported data from PPRs.[177] Section 3.26 provides a full description of "PCC and SBA Entrepreneurship Training Summary." The following grant-level variables are used in these calculations:
  - entrepreneurship-related training participants
- "PCC Unique Users" is a grant-level table containing the estimated number of unique users in the labor force over the course of the grant. The table relies on grantee-reported data from PPRs and CAC and on survey results reported in Becker et al. (2010).[178] Section 4.22 provides

a full description of "PCC Unique Users." The following grant-level variables are used in these calculations:

- unique users

The script loads the prepared data described above. The data are merged by award number, resulting in a single grant-level table containing all data. This step also subsets the data to include only PCC grants.

ASR made the following assumptions to estimate this impact:

- The entrepreneurship survey results reported in Becker et al. (2010) are representative of entrepreneurship usage at PCCs
- All PCC entrepreneurship training participants are unique and are engaged in entrepreneurial endeavors
- The entrepreneurship survey results reported in Becker et al. (2010) are representative of different entrepreneurial activities among PCC users and PCC training program participants

Based on the assumptions above, the following steps estimate benefits due to BTOP:

- The number of unique PCC users for each grant is multiplied by the study coefficient from Becker et al. (2010) representing the proportion of computer users at public libraries engaged in "activities related to starting or managing a business of their own."[179] This estimates the number of unique users engaged in entrepreneurial endeavors.
- The number of entrepreneurship training participants is added with the number of unique users engaged in entrepreneurial endeavors to determine the total number of individuals engaged in entrepreneurial activities.
- The estimated number of users engaged in entrepreneurial activities is multiplied by rates reported in Becker et al. (2010) for library users that started a business, looked for new customers, and grew their business. This estimates the number of individuals that started a business, looked for new customers, and grew their business due to PCC activities.

The estimates of unique users, entrepreneurship trainees, and individuals engaged in entrepreneurial activities are aggregated for the evaluation study sample, PCC grants not in the evaluation study sample, and all PCC grants and written to the "PCC.W.2 - Table 41" Excel file. Table 41 below presents the results of the R script. *Final Report Table 41* presents the total estimates for all of BTOP.

Table 41. Total Estimated Number of Individuals Engaged in Any Entrepreneurial Activity through PCC Grants

Estimate	Total
Unique users	746,698
Entrepreneurship trainees	1,186
Engaged in entrepreneurial activities	53,455

11.2.2 Final Report Table 42. Total Estimated Number Engaged in Different Entrepreneurial Activities Due to PCC Grants

The estimates of individuals starting a business, looking for new customers, and growing their business are aggregated for the evaluation study sample, PCC grants not in the evaluation study sample, and all PCC grants and written to the "PCC.W.2 - Table 42" Excel file. Table 42 below

presents the results of the R script. *Final Report Table 42* presents the total estimates for all of BTOP.

Table 42. Total Estimated Number Engaged in Different Entrepreneurial Activities Due to PCC Grants

Estimate	Total
Started a business	25,200
Looked for new customers	26,727
Grew their business	12,982

11.3 Final Report Appendix C.4 Obtained Employment (Impact ID: SBA.W.1)

11.3.1 Final Report Table 43. Total Estimated New Internet Job Searchers and Decreased Unemployment Duration (Weeks) Due to SBA Grants

Kuhn and Mansour (2013) estimated that unemployment durations are 24 percent shorter for Internet job searchers than offline job searchers.[180] ASR used grantee-reported data on SBA subscribers and job-search training attendees to estimate the number of new Internet job searchers and decreased unemployment duration due to SBA grants.

The "SBA.W.1" R script uses the following prepared data to provide a quantitative estimate of SBA grants' impact on obtaining employment:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "PCC and SBA Job Search Training Summary" is a grant-level table summarizing the number of participants in grantee job-search-related training programs. The table relies on grantee-reported data from PPRs.[181] Section 4.15 provides a full description of "PCC and SBA Job Search Training Summary." The following grant-level variables are used in these calculations:
  - job-search-related training participants
- "PCC and SBA Labor Force Statistics" is a grant-level table containing several labor force statistics. The table is derived from publicly available data published by BLS, Census Bureau, and DOL/BLS.[182] Section 4.16 provides a full description of "PCC and SBA Labor Force Statistics." The following grant-level variables are used in these calculations:
  - labor force participation rate
  - unemployment rate
  - average unemployment duration
  - minimum wage
- "SBA Subscribers" is a grant-level table containing the estimated number of new individual, household, and business subscribers due to SBA grant activity. The table relies on grantee-reported data from PPRs and publicly available data published by the Census Bureau.[183] Section 4.25 provides a full description of "SBA Subscribers." The following grant-level variables are used in these calculations:
  - total individual subscribers

The script loads the prepared data described above. The data are merged by award number, resulting in a single grant-level table containing all data. This step also subsets the data to include only SBA grants.

ASR made the following assumptions to estimate this impact:

- The labor force participation and unemployment rates for new subscribers are equal to those of the grant's service area population as a whole
- All SBA job-search-related training participants are unique and are new Internet job searchers
- The average unemployment duration in a state uniformly describes unemployment periods of individuals in SBA service areas within that state

Based on the assumptions above, the following steps estimate benefits due to BTOP:

- The number of new job searchers resulting from new subscribers is estimated by multiplying total individual subscribers by service area labor force participation and unemployment rates.
- The number of job search training participants is added to the number of new job searchers resulting from new subscribers to account for both types of grant activities that lead to users searching for jobs online (new household subscriptions and job search training).
- The number of new Internet job searchers resulting from SBA grants is multiplied by the average unemployment duration in each service area and the study coefficient from Kuhn and Mansour (2013). This creates a variable representing the estimated decrease in total unemployment duration for each SBA grant.
- To estimate the economic benefit (in millions) of decreased unemployment durations due to SBA grants, each grant's projected total decrease in unemployment duration is multiplied by the minimum wage times forty (the weekly minimum wage for full-time work) in the grant's service area, then divided by one million.

The estimates of new job searchers and total decrease in unemployment duration are aggregated for the evaluation study sample, SBA grants not in the evaluation study sample, and all SBA grants and written to the "SBA.W.1 - Table 43" Excel file. Table 43, as created by the R script, is presented below. Only the total estimates for all PCC projects were included in the *Final Report Table 43*. Other estimates were discussed in the text surrounding the table.

Table 43. Total Estimated New Internet Job Searchers and Decreased Unemployment Duration (Weeks) Due to SBA Grants

Measure	Estimate
New Internet job searchers	104,259
Decreased duration	626,980

The estimates of economic benefit are aggregated for the evaluation study sample, SBA grants not in the evaluation study sample, and all SBA grants and written to the "SBA.W.1 - Table F" Excel file. Table F below presents the results of the R script. This table was not presented in the *Final Report,* but the estimates were discussed in the text relevant to the economic benefit of this impact.

Table F. Total Estimated Economic Benefit from Decreased Unemployment Duration Due to SBA Grants (Million USD, Annual)

Estimate	Evaluation Study Sample	Rest of BTOP	Total
Total Economic Benefit	94	96	190

11.4 Final Report Appendix C.5 Started or Grew Businesses (Impact ID: SBA.W.2)

11.4.1 Final Report Table 44. Total Estimated Number Engaged in Different Entrepreneurial Activities Due to SBA Grants

Becker et al. (2010) investigated trends in public library patrons' use of free access to computers and the Internet.[184] The study found that that 7 percent of public-access computer users at public libraries across the United States engaged in "activities related to starting or managing a business of their own."[185] Additionally, 3.3 percent started a business using the library computers and 3.5 percent looked for new customers.[186] Nearly half of those who looked for new customers, 1.7 percent of all users, grew their business.[187] Additionally, SBA grantees noted that training participants engaged in entrepreneurial activities.[188] ASR used grantee-reported data on training attendees to estimate the number of users engaging in various entrepreneurial activities due to SBA grants.

The "SBA.W.2a" R script uses the following prepared data to provide a quantitative estimate of SBA grants' impact on entrepreneurial activity:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "PCC and SBA Entrepreneurship Training Summary" is a grant-level table summarizing the number of participants in grantee-provided entrepreneurship training programs. This table relies on grantee-reported data from PPRs.[189] Section 3.26 provides a full description of "PCC and SBA Entrepreneurship Training Summary." The following grant-level variables are used in these calculations:
  - entrepreneurship-related training participants

The script loads the prepared data described above. The data are merged by award number, resulting in a single grant-level table containing all data. This step also subsets the data to include only SBA grants.

ASR made the following assumptions to estimate this impact:

- The entrepreneurship survey results reported in Becker et al. (2010) are representative of SBA entrepreneurship training participants
- All SBA entrepreneurship training participants are unique and are engaged in entrepreneurial endeavors
- The entrepreneurship survey results reported in Becker et al. (2010) are representative of different entrepreneurial activities among SBA entrepreneurship training program participants

Based on the assumptions above, the following steps estimate benefits due to BTOP:

- The estimated number of users engaged in entrepreneurial activities (total entrepreneurship training participants) is multiplied by rates reported in Becker et al. (2010) for library users that started a business, looked for new customers, and grew their business. This estimates the number of individuals that started a business, looked for new customers, and grew their business due to SBA activities.

The estimates of entrepreneurship trainees, individuals starting a business, looking for new customers, and growing their business are aggregated for the evaluation study sample, SBA grants not in the evaluation study sample, and all SBA grants and written to the "SBA.W.2a - Table 44"

Excel file. Table 44 below presents the results of the R script. *Final Report Table 44* includes the total estimates for all of BTOP.

Table 44. Total Estimated Number Engaged in Different Entrepreneurial Activities Due to SBA Grants

Estimate	Total
Engaged in entrepreneurial activities (entrepreneurship training attendees)	1,342
Started a business	633
Looked for new customers	671
Grew their business	326

11.4.2 Final Report Appendix C.5.1 Business Productivity Gains (Impact ID: SBA.W.2)

11.4.2.1 Final Report Table 45. Total New Business Subscribers and Years of Business Subscriptions Due to SBA Grants

Grimes, Ren, and Stevens (2009) estimated that broadband adoption by a firm increases the firm's productivity by 7.6 percent to 13 percent.[190] ASR used grantee-reported data on new business subscribers to estimate the economic benefit of increased productivity due to SBA grant activity.

The "SBA.W.2b" R script uses the following prepared data to provide a quantitative estimate of SBA grants' impact on businesses' productivity:

- "Grants" contains a table describing every awarded BTOP grant. Section 4.8 provides a full description of "Grants."
- "SBA Subscribers" is a grant-level table containing the estimated number of new individual, household, and business subscribers due to SBA grant activity. The table relies on grantee-reported data from PPRs and publicly available data published by Census Bureau.[191] Section 4.25 provides a full description of "SBA Subscribers." The following grant-level variables are used in these calculations:
  - quarterly business subscriptions

The script loads the prepared data described above. The data are merged by award number, resulting in a single grant-level table containing all data. This step also subsets the data to include only SBA grants.

ASR made the following assumptions to estimate this impact:

- Businesses with new broadband subscriptions due to SBA grants would not subscribe otherwise
- All new business subscribers maintained service through the end of 2013
- On average, new business subscribers due to SBA grants are equal in size to the average nonfarm sole proprietorships. A report written for the SBA Office of Advocacy notes that the majority of small business entities in the United States are nonfarm sole proprietorships with average annual tax receipts of $56,416.[192]

Based on the assumptions above, the following steps estimate benefits due to BTOP:

- The years of business subscriptions due to SBA grants is estimated as the number of new business subscriptions in a quarter times the number of years from the end of the quarter until the end of 2013 (the number of days from the end of the quarter until the end of 2013 divided by 365).
- The economic impact of increased productivity due to new business subscriptions is estimated by multiplying the years of business subscriptions due to SBA grants by $56,416 and the Grimes, Ren, and Stevens (2009) study coefficient.

The estimates of new business subscribers and years of business subscriptions are aggregated for the evaluation study sample, SBA grants not in the evaluation study sample, and all SBA grants and written to the "SBA.W.2b - Table 45" Excel file. Table 45 below presents the results of the R script. *Final Report Table 45* presents the total estimates for all of BTOP.

Table 45. Total New Business Subscribers and Years of Business Subscriptions Due to SBA Grants

Measure	Total
Business subscribers	6,484
Years of business subscriptions	14,714

The estimates of economic benefit are aggregated for the evaluation study sample, SBA grants not in the evaluation study sample, and all SBA grants and written to the "SBA.W.2b - Economic Benefit" Excel file. Table G below presents the results of the R script.

Table G. Total Estimated Economic Benefit of Increased Productivity at Businesses with New Broadband Subscriptions Due to SBA Grants (Million USD)

Estimate	Evaluation Study Sample	Rest of BTOP	Total
Total Economic Benefit	13	50	63

Section 12.Final Report Appendix D. Long-Term Quantitative Analysis Methodology and Data

This section includes descriptions of all files, programs, and processes used for one of two purposes: to summarize the activities, effects, or characteristics of BTOP; or to extrapolate and summarize the quantitative long-term benefits of BTOP. Summary tables and figures are created from prepared data and input files, statistical analysis results, and combinations thereof. Benefits are extrapolated from prepared data, statistical analyses, and results found in the broadband literature.

Each subsection below includes the following:

- One R script to load one or more sets of prepared data, read input files, and load statistical analysis results. For summary tables and figures, loaded and imported data are prepared as necessary to produce the summary tables and figures included in the *Final Report.* For benefits extrapolations, data are prepared as necessary and sets of methodological steps are performed to extrapolate long-term benefits. R scripts are sets of instructions written in the R statistical programming language than can be executed in the R statistical package.
- At least one results table, saved in Excel format, for each subsection. Results tables were copied from the Excel files and pasted into the *Final Report*. Unlike the programs described in the earlier sections of this document, there can be multiple Excel output files associated with each script.

Every script has a header that defines the parent directory and calls the shared source code scripts described in Section 2. Users are required to redefine the parent directory to the correct location on their computer to reproduce the manipulation steps and output. This can be done in batch using the "Set Parent Directory in All Script Files" R script described in Section 1.3.

The following subsections describe the R scripts, in the order of presentation, used to estimate and summarize the quantitative intermediate impacts for *Appendix D. Long-Term Quantitative Analysis and Data* in the *Final Report.* All scripts and generated files discussed below are located in the "D. Long-Term Quantitative Analysis and Data" folder.

12.1 Final Report Appendix D.1 Identifying Counties Receiving Broadband due to BTOP CCI Projects

12.1.1 Final Report Table 46. Characteristics of Selected BTOP CCI Grants and Service Area Counties

Final Report Table 46 summarizes the service areas of each of the twelve evaluation study CCI grants. Included in the table are the counties included in each grant's service area (including counties added due to the presence of CAIs), the total number of counties in each grant's service area, and the total land area, population, and population without broadband availability of each grant's service area. Sample-wide totals are also included. The "Table 46" R script reproduces this table using the following prepared data:

- "NBM Population Statistics" is a county-level data containing population, demographic, and broadband availability statistics derived from NBM data provided to ASR by NTIA.[193] Section 4.12 provides a full description of "NBM Population Statistics." The script uses the following county-level attributes from the June 30, 2011 NBM release:
  - total population
  - population with broadband availability

ASR uses the following input files to generate the table:

- "CCI Case Study Service Areas," described in Section 3.16
- "Census Bureau County Shapefiles Selected Attributes," described in Section 3.21. The following county-level attributes are used:
  - official county name
  - land area

The script loads the prepared data and reads the input files discussed above.

Land area is converted from square meters to square miles. The string "County" is stripped from official county names; the string "city" is not stripped in order to differentiate between counties and cities that share names. Grantee award numbers are replaced with grantee names and states. The list of grantees and counties is merged with county attributes to retrieve official names and county size, and merged with NBM statistics to retrieve population and population without broadband availability (population minus population with broadband availability). Individual rows representing counties are aggregated to the grant level by concatenating counties into a list, counting the number of counties, and summing area, population, and population without availability. This is done in two steps to differentiate counties added to service areas because of a connected CAI. Totals are calculated for the entire evaluation study sample as well. The grant- and sample-level aggregations are combined and written to the "Table 46" Excel file. Table 46 below presents the results of the R script.

Table 46. Characteristics of Selected BTOP CCI Grants and Service Area Counties

Grant (State)	Counties	Number of Counties	Land Area (mi^2)	Population (June 30, 2011)	Population without Availability (June 30, 2011)
Clearwave Communications (Illinois)	Alexander, Clay, Clinton, Edwards, Franklin, Hamilton, Jackson, Jefferson, Johnson, Marion, Massac, Perry, Pulaski, Randolph, Richland, Saline, St. Clair, Union, Wabash, Washington, Wayne, White, Williamson	23	9,995	795,832	138,121
Executive Office of the State of West Virginia (West Virginia)	Barbour, Berkeley, Boone, Braxton, Brooke, Cabell, Calhoun, Clay, Doddridge, Fayette, Gilmer, Grant, Greenbrier, Hampshire, Hancock, Hardy, Harrison, Jackson, Jefferson, Kanawha, Lewis, Lincoln, Logan, Marion, Marshall, Mason, McDowell, Mercer, Mineral, Mingo, Monongalia, Monroe,	55	24,038	1,858,030	853,734

Grant (State)	Counties	Number of Counties	Land Area (mi^2)	Population (June 30, 2011)	Population without Availability (June 30, 2011)
	Morgan, Nicholas, Ohio, Pendleton, Pleasants, Pocahontas, Preston, Putnam, Raleigh, Randolph, Ritchie, Roane, Summers, Taylor, Tucker, Tyler, Upshur, Wayne, Webster, Wetzel, Wirt, Wood, Wyoming				
Lane Council of Governments (Oregon)	Douglas, Klamath, Lane	3	15,530	529,963	36,425
Massachusetts Technology Park (Massachusetts) [194]	Berkshire, Franklin, Hampden, Hampshire, Middlesex, Worcester	6	5,099	3,131,209	32,476
MCNC (North Carolina)	Alleghany, Anson, Ashe, Avery, Beaufort, Bertie, Brunswick, Buncombe, Cabarrus, Caldwell, Camden, Carteret, Caswell, Chatham, Chowan, Cleveland, Columbus, Craven, Cumberland, Currituck, Dare, Edgecombe, Franklin, Gaston, Gates, Graham, Granville, Halifax, Harnett, Haywood, Henderson, Hertford, Hyde, Jackson, Lee, Lincoln, Madison, Martin, McDowell, Mecklenburg, Mitchell, Moore, Nash, New Hanover, Northampton, Onslow, Pasquotank, Perquimans, Person, Pitt, Polk, Richmond, Robeson, Rockingham, Rutherford, Scotland, Stokes, Surry, Swain, Transylvania, Tyrrell, Union, Vance, Wake, Warren, Washington, Watauga, Wilson, Yancey	69	32,730	6,434,948	435,267
Merit Network, Inc. (Michigan)	Allegan, Antrim, Arenac, Bay, Benzie, Berrien, Branch, Cass, Charlevoix, Clare, Crawford, Emmet, Gladwin, Grand Traverse, Hillsdale, Iosco, Isabella, Kalkaska, Lake, Lenawee, Manistee, Mason, Midland, Monroe, Montmorency, Muskegon, Oceana, Otsego, Ottawa, Roscommon, St. Joseph, Van Buren *Added due to connected CAI:*	33	17,640	1,986,258	208,060

Grant (State)	Counties	Number of Counties	Land Area (mi^2)	Population (June 30, 2011)	Population without Availability (June 30, 2011)
	Cheboygan				
Mid-Atlantic Broadband Communities Corporation (Virginia)	Amelia, Bedford, Bedford city, Buckingham, Campbell, Charlotte, Chesterfield, Cumberland, Dinwiddie, Emporia city, Franklin, Greensville, Halifax, Henry, Lunenburg, Lynchburg city, Martinsville city, Petersburg city, Pittsylvania, Powhatan, Prince George, Sussex	22	8,590	975,845	207,605
OneCommunity (Ohio)[195]	Ashland, Ashtabula, Champaign, Clermont, Columbiana, Coshocton, Crawford, Cuyahoga, Erie, Franklin, Geauga, Holmes, Huron, Lake, Lorain, Lucas, Mahoning, Marion, Medina, Montgomery, Morrow, Ottawa, Portage, Richland, Sandusky, Seneca, Stark, Summit, Trumbull, Tuscarawas, Washington, Wayne, Wood	33	15,397	7,259,807	142,610
OSHEAN (Rhode Island)	Bristol, Kent, Newport, Providence, Washington; and Bristol, Massachusetts	6	1,587	1,592,160	4,035
South Dakota Network (South Dakota)	Beadle, Brookings, Brown, Butte, Clark, Codington, Deuel, Grant, Hamlin, Hand, Hughes, Hyde, Kingsbury, Lake, Lawrence, Lincoln, Marshall, McCook, Meade, Minnehaha, Pennington, Spink, Walworth *Added due to connected CAI*: Aurora, Bennett, Bon Homme, Brule, Campbell, Charles Mix, Clay, Custer, Davison, Day, Dewey, Douglas, Edmunds, Fall River, Faulk, Gregory, Haakon, Harding, Hutchinson, Jackson, Jerauld, Jones, Lyman, McPherson, Mellette, Miner, Moody, Perkins, Potter, Roberts, Shannon, Stanley, Sully, Tripp, Turner, Union, Yankton	60	68,516	797,628	159,684
University of Arkansas System	Arkansas, Ashley, Baxter, Benton, Boone, Bradley, Calhoun, Carroll, Chicot, Clark, Clay, Cleburne,	75	52,035	2,946,242	400,405

Grant (State)	Counties	Number of Counties	Land Area (mi^2)	Population (June 30, 2011)	Population without Availability (June 30, 2011)
(Arkansas)	Cleveland, Columbia, Conway, Craighead, Crawford, Crittenden, Cross, Dallas, Desha, Drew, Faulkner, Franklin, Fulton, Garland, Grant, Greene, Hempstead, Hot Spring, Howard, Independence, Izard, Jackson, Jefferson, Johnson, Lafayette, Lawrence, Lee, Lincoln, Little River, Logan, Lonoke, Madison, Marion, Miller, Mississippi, Monroe, Montgomery, Nevada, Newton, Ouachita, Perry, Phillips, Pike, Poinsett, Polk, Pope, Prairie, Pulaski, Randolph, Saline, Scott, Searcy, Sebastian, Sevier, Sharp, St. Francis, Stone, Union, Van Buren, Washington, White, Woodruff, Yell				
Zayo Bandwidth (Indiana)	Allen, Bartholomew, Dearborn, Delaware, Elkhart, Fayette, Grant, Howard, Jefferson, Kosciusko, Lake, LaPorte, Madison, Monroe, Porter, Sullivan, Vanderburgh, White *Added due to connected CAI*: Gibson, Johnson, Marion, Tippecanoe, Wabash	23	9,703	3,653,508	89,868
Total		**408**	**260,861**	**31,961,430**	**2,708,290**

12.2 Final Report Appendix D.2 Selecting Control Counties

12.2.1 Final Report Table 47. Potential Control Counties in the United States

Final Report Table 47 summarizes the process used by ASR to identify potential control counties in the United States. The "Table 47" R script reproduces this table, which is saved in the "Table 47" Excel file. Table 47 below presents the results of the R script. *Final Report Table 47* and the "Table 47" R script are identical to *Final Report Table 11* and the "Table 11" R script, respectively. See Section 9.2.1 for a full description of *Final Report Table 11* and the "Table 11" R script.

Table 47. Potential Control Counties in the United States

All counties and equivalents in the United States	**3,234**	
Less counties in Alaska, Hawaii, and outlying areas	3,109	
Less counties not in proposed BTOP service area	2,640	
Less counties in awarded BTOP or BIP grant service area[196]	**884**	**Potential Controls**

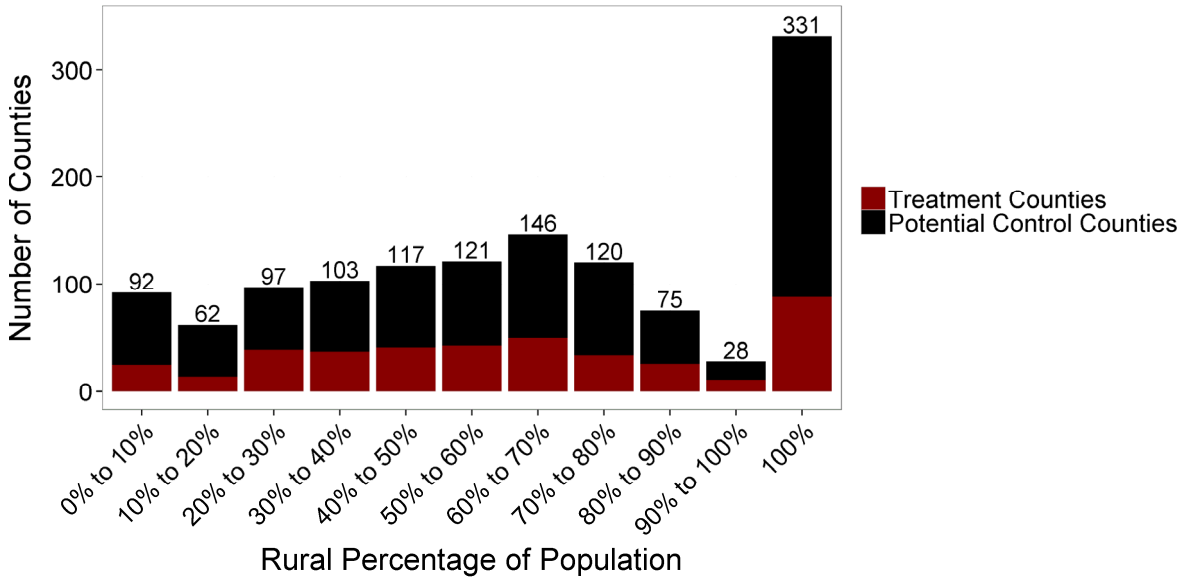
12.2.2 Final Report Figure 11. Distribution of Rural Percentage of Population in Treatment and Potential Control Counties

Final Report Figure 11 summarizes the rurality of all treatment and potential control counties in the United States that are considered in the "Matching" statistical analysis. Rurality is the percentage of the total population within a county that resides in a rural area. The "Figure 11" R script reproduces this figure using the following prepared data:

- "Treatment and Potential Control Counties" identifies every county included in NBM data according to the steps used to identify potential control counties. These data are derived from data published by the USDA, ProPublica, and NTIA, as well as data collected by ASR from individual grantees and other public sources.[197] Section 4.28 provides a full description of "Treatment and Potential Control Counties."
- "NBM Population Statistics" is a county-level table containing population, demographic, and broadband availability statistics derived from NBM data provided to ASR by NTIA.[198] Section 4.12 provides a full description of "NBM Population Statistics." The following county-level attributes from the June 30, 2011 NBM release are used:
  - rural percentage of population

The script loads the prepared data described above. Treatment and potential control counties are identified in the NBM data; all counties without one of these designations are removed. The continuous values of rural percentage of population are converted into a categorical variable representing each 10-percentage-point interval between 0 percent and 100 percent. An eleventh categorical level is added for counties with a rural percentage of population of 100 percent. A stacked bar plot filled by treatment or potential control is then created and saved to the "Figure 11" PNG file. Figure 11 below presents the results of the R script.

Figure 11. Distribution of Rural Percentage of Population in Treatment and Potential Control Counties



12.3 Final Report Appendix D.3 Developing Sensitivity Analysis Control Groups

12.3.1 Final Report Table 48. Census Blocks and Counties with Reported Decreases in the Broadband Availability Rate over the Study Period

Final Report Table 48 summarizes the census blocks, populated census blocks, and counties in the contiguous United States in which the broadband availability rate decreased from the June 30, 2011 release to the June 30, 2013 release. The "Table 48" R script reproduces these tables using the following prepared data:

- "NBM Population Statistics" is a county-level table containing population, demographic, and broadband availability statistics derived from NBM data provided to ASR by NTIA.[199] Section 4.12 provides a full description of "NBM Population Statistics." The table used the following county-level variables from the June 30, 2011 and June 30, 2013 NBM releases:
  - availability, forward-looking availability, and backward-looking availability rates

The script uses the following input files:

- The "NBM Census Blocks with Availability Rate Decreases by County" Excel file contains county-level counts of populated and unpopulated census blocks. Counts include blocks with decreases in availability from the June 30, 2011 to the June 30, 2013 NBM releases. Data are derived from the June 30, 2011 and June 30, 2013 NBM releases. Section 3.26 provides a full description of "NBM Census Blocks with Availability Rate Decreases by County." The script uses the following variables:
  - count of census blocks

- counts of census blocks with any decrease in the availability rate, decrease greater than or equal to 1 percentage point, and decrease greater than or equal to 5 percentage points for the NOFA and NTIA definitions of broadband

The script loads the prepared data and reads the input file described above. The block data are aggregated into two tables: one for all census blocks, and another for populated census blocks only. In the county data, changes between releases are calculated and Boolean indicators denoting decreases, decreases of more than a percentage point, and decreases of more than 5 percentage points are created. County-level Boolean indicators and the count of counties are then aggregated.

Aggregated results for census blocks, populated census blocks, and counties are combined and percentage values are calculated. A summary table is then created and saved to the "Table 48" Excel file. Table 48 below presents the results of the R script.

Table 48. Census Blocks and Counties with Reported Decreases in the Broadband Availability Rate over the Study Period

Description		Census Blocks		Populated Blocks		Counties	
		N	%	N	%	N	%
Total		**11,007,989**	**100.00%**	**6,166,982**	**100.00%**	**3,109**	**100.00%**
NOFA	Decrease in availability rate	166,657	1.51%	166,657	2.70%	1,695	54.52%
NOFA	Decrease in availability rate ≥ 1%	166,211	1.51%	166,211	2.70%	1,157	37.21%
NOFA	Decrease in availability rate ≥ 5%	164,358	1.49%	164,358	2.67%	518	16.66%
NTIA	Decrease in availability rate	235,309	2.14%	235,309	3.82%	1,385	44.55%
NTIA	Decrease in availability rate ≥ 1%	234,972	2.14%	234,972	3.81%	1,091	35.09%
NTIA	Decrease in availability rate ≥ 5%	233,333	2.12%	233,333	3.78%	638	20.52%

12.3.2 Final Report Table 49. Counties with Reported Decreases in the Broadband Availability Rate over the Study Period

Final Report Table 49 summarizes the counties in the contiguous United States in which the broadband availability rate decreased from the June 30, 2011 release to the June 30, 2013 release using three different availability rates (availability, forward looking, and backward looking). The "Table 49" R script reproduces these tables using the following prepared data:

- "NBM Population Statistics" is a county-level table containing population, demographic, and broadband availability statistics derived from NBM data provided to ASR by NTIA.[200] Section 4.12 provides a full description of "NBM Population Statistics." The table used the following county-level variables from the June 30, 2011 and June 30, 2013 NBM releases:
  - availability, forward-looking availability, and backward-looking availability rates

The R script loads the prepared data. Changes between releases are calculated for each of the availability rates and Boolean indicators denoting decreases, decreases of more than a percentage point, and decreases of more than 5 percentage points are created. A summary table is then created and saved to the "Table 49" Excel file. Table 49 below presents the results of the R script.

Table 49. Counties with Reported Decreases in the Broadband Availability Rate over the Study Period

Description		Availability		Forward Looking		Backward Looking	
		N	%	N	%	N	%
NOFA	Decrease in availability rate	1,695	54.52%	116	3.73%	94	3.02%
	Decrease in availability rate ≥ 1%	1,157	37.21%	7	0.23%	9	0.29%
	Decrease in availability rate ≥ 5%	518	16.66%	0	0.00%	0	0.00%
NTIA	Decrease in availability rate	1,385	44.55%	117	3.76%	91	2.93%
	Decrease in availability rate ≥ 1%	1,091	35.09%	10	0.32%	11	0.35%
	Decrease in availability rate ≥ 5%	638	20.52%	0	0.00%	0	0.00%

12.4 Final Report Appendix D.4 Matching Results

12.4.1 Final Report Table 50. Descriptive Statistics of Matching Variables

Final Report Table 50 provides descriptive statistics of the variables used in different matching specifications. The "Table 50" R script reproduces this table using the following statistical analysis data:

- "Matching" contains the results of twelve specifications of treatment-control matching using nearest neighbor matching with restrictions. The results are stored as tables with one column identifying the treatment county FIPS code and a second with the matched control county FIPS code. "Matching" also contains the matching data set, a county-level table of NBM population statistics, ACS (2006-2010) statistics, and county matching group classification. Section 5.1 provides a full description of "Matching." The following results and data are used:
  - NTIA Primary, NTIA Sensitivity, NTIA Forward-Looking Primary, NTIA Forward-Looking Sensitivity, NTIA Backward-Looking Primary, NTIA Backward-Looking Sensitivity, NOFA Primary, NOFA Sensitivity, NOFA Forward-Looking Primary, NOFA Forward-Looking Sensitivity, NOFA Backward-Looking Primary, and NOFA Backward-Looking Sensitivity matching results
  - NBM data from the June 30, 2011 release:
    - availability, forward-looking availability, and backward-looking availability
    - total, rural, minority, over-60, and in-poverty populations
  - ACS (2006-2010) data on non-English-speaking population

The R script loads the statistical analysis results described above. The relevant broadband, population, and vulnerable population data are extracted for each set of matching results. The data are then combined, reshaped, summarized (mean and median), and formatted. Several columns in the output table represent the same data: Base and Backward Looking matching results are identical by design; and data for Treatment counties are identical for Primary and Sensitivity matching results. Descriptive statistics for these duplicative columns were combined. A summary table of results is then created and written to the "Table 50" Excel file. Table 50 below presents the results of the R script.

Table 50. Descriptive Statistics of Matching Variables

Characteristic		Statistic	Base and Backward Looking			Forward Looking		
			Treatment	Primary Control	Sensitivity Control	Treatment	Primary Control	Sensitivity Control
NOFA	Availability	Mean	90.22%	90.26%	90.23%	85.67%	85.79%	85.77%
		Median	94.56%	94.71%	94.33%	90.03%	90.29%	90.06%
	Rurality	Mean	60.25%	59.73%	60.47%	60.25%	60.25%	60.84%
		Median	61.10%	60.96%	62.39%	61.10%	61.69%	63.36%
	Minority	Mean	16.36%		15.58%	16.36%		15.22%
		Median	9.62%		8.84%	9.62%		8.97%
	Over 60 Years Old	Mean	25.92%		25.39%	25.92%		25.46%
		Median	25.50%		24.81%	25.50%		24.63%
	Poverty	Mean	17.79%		17.07%	17.79%		17.11%
		Median	17.28%		16.31%	17.28%		16.25%
	Non-English	Mean	1.95%		2.03%	1.95%		2.12%
		Median	1.22%		1.43%	1.22%		1.46%
NTIA	Availability	Mean	73.17%	73.10%	73.10%	68.49%	68.53%	68.52%
		Median	82.85%	82.79%	82.88%	77.19%	77.72%	77.16%
	Rurality	Mean	60.25%	61.46%	62.16%	60.25%	60.88%	62.80%
		Median	61.10%	65.27%	66.08%	61.10%	61.01%	64.16%
	Minority	Mean	16.36%		15.59%	16.36%		16.06%
		Median	9.62%		10.90%	9.62%		10.12%
	Over 60 Years Old	Mean	25.92%		25.26%	25.92%		25.19%
		Median	25.50%		24.44%	25.50%		24.37%
	Poverty	Mean	17.79%		16.87%	17.79%		17.23%
		Median	17.28%		16.18%	17.28%		16.41%
	Non-English	Mean	1.95%		2.09%	1.95%		2.05%
		Median	1.22%		1.46%	1.22%		1.33%

12.4.2 Final Report Figure 12. Number of Different Control Counties Selected for Treatment Counties

Final Report Figure 12 summarizes the variation in matching results based on different matching specifications. The "Figure 12" R script reproduces this figure using the following statistical analysis:

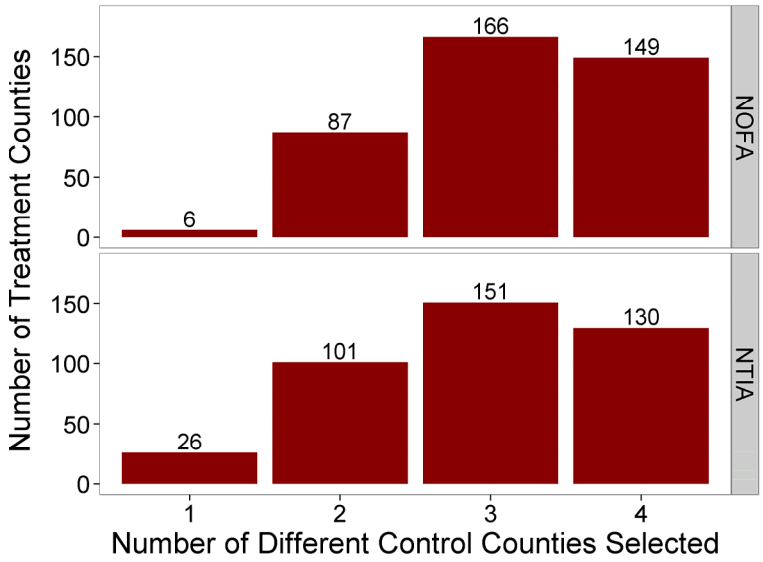
- "Matching" contains the results of twelve specifications of treatment-control matching using nearest neighbor matching with restrictions. The results are stored as tables with one column identifying the treatment county FIPS code and a second with the matched control county FIPS code. "Matching" also contains the matching data set, a county-level table of NBM population

statistics, ACS (2006-2010) statistics, and county matching group classification. Section 5.1 provides a full description of "Matching." The following results and data are used:

- NTIA Primary, NTIA Sensitivity, NTIA Forward-Looking Primary, NTIA Forward-Looking Sensitivity, NTIA Backward-Looking Primary, NTIA Backward-Looking Sensitivity, NOFA Primary, NOFA Sensitivity, NOFA Forward-Looking Primary, NOFA Forward-Looking Sensitivity, NOFA Backward-Looking Primary, and NOFA Backward-Looking Sensitivity matching results
- NBM data from the June 30, 2011 release:
  - availability, forward-looking availability, and backward-looking availability
  - total, rural, minority, over-60, and in-poverty populations
- ACS (2006-2010) data on non-English-speaking population

The R script loads the statistical analysis results described above. All twelve matching results are combined into a single table, and the number of different selected control counties for each treatment county is tabulated. A bar chart is created using this tabulation and written to the "Figure 12" PNG file. Figure 12 below presents the results of the R script.

Figure 12. Number of Different Control Counties Selected for Treatment Counties



12.5 Final Report Appendix D.5 Difference-in-Differences Estimates

12.5.1 Final Report Table 51. Availability Rates, Differences, and Difference-in-Differences for Matched Pair Groups

Final Report Table 51 summarizes the June 30, 2011 and June 30, 2011 availability rates, differenced availability rates, and difference-in-differences for the twelve matching results. The "Table 51" R script reproduces this table using the following statistical analysis:

- "Effect on Availability" contains a program-level table containing June 30, 2011 and June 30, 2011 availability rates, differenced availability rates, and difference-in-differences for the twelve matching results. Each row in the table represents the treatment and control availability rates, differences, and difference-in-differences for a unique combination of the definition of

broadband, any adjustments made to the data, the matching group, and the release date (differences and difference-in-differences are identical between releases). Section 5.2 provides a full description of the analysis leading to these results. The following results are used in this analysis:

- NTIA Primary, NTIA Sensitivity, NTIA Forward-Looking Primary, NTIA Forward-Looking Sensitivity, NTIA Backward-Looking Primary, NTIA Backward-Looking Sensitivity, NOFA Primary, NOFA Sensitivity, NOFA Forward-Looking Primary, NOFA Forward-Looking Sensitivity, NOFA Backward-Looking Primary, and NOFA Backward-Looking Sensitivity availability rates, differenced availability rates, and difference-in-differences estimates of the effect of BTOP

The R script loads the statistical analysis results discussed above. The table is reshaped and formatted. After these manipulation steps, a table of results is written to the "Table 51" Excel file. Table 51 below presents the results of the R script.

Table 51. Availability Rates, Differences, and Difference-in-Differences for Matched Pair Groups

Definition	Match	Group	Release	Availability	Forward Looking	Backward Looking
NOFA	Primary	Treatment	2011	96.71%	95.25%	96.71%
			2013	96.53%	96.53%	97.98%
			Difference	**-0.18%**	**1.28%**	**1.26%**
		Control	2011	96.79%	95.51%	96.79%
			2013	96.06%	96.55%	97.87%
			Difference	**-0.73%**	**1.04%**	**1.08%**
		Difference-in-Differences		**0.55%**	**0.24%**	**0.18%**
	Sensitivity	Treatment	2011	96.71%	95.25%	96.71%
			2013	96.53%	96.53%	97.98%
			Difference	**-0.18%**	**1.28%**	**1.26%**
		Control	2011	96.50%	95.14%	96.50%
			2013	95.92%	96.29%	97.68%
			Difference	**-0.57%**	**1.15%**	**1.18%**
		Difference-in-Differences		**0.39%**	**0.13%**	**0.08%**
NTIA	Primary	Treatment	2011	91.53%	89.72%	91.53%
			2013	94.40%	94.40%	96.19%
			Difference	2.87%	4.68%	4.66%
		Control	2011	92.28%	90.57%	92.28%
			2013	93.16%	93.41%	95.62%
			Difference	**0.88%**	**2.84%**	**3.34%**
		Difference-in-Differences		**2.00%**	**1.84%**	**1.32%**
	Sensitivity	Treatment	2011	91.53%	89.72%	91.53%
			2013	94.40%	94.40%	96.19%
			Difference	**2.87%**	**4.68%**	**4.66%**
		Control	2011	91.77%	89.53%	91.77%
			2013	92.75%	92.48%	95.39%
			Difference	**0.98%**	**2.95%**	**3.62%**
		Difference-in-Differences		**1.89%**	**1.73%**	**1.04%**

12.5.2 Final Report Table 52. Estimated Total Population with Broadband Availability Due to BTOP

Final Report Table 52 summarizes the effects of BTOP on broadband availability in terms of population for the twelve matching results. That is, the estimated difference-in-differences effects of

BTOP times the population in the BTOP service area. The "Table 52" R script reproduces this table using the following statistical analysis:

- "Effect on Availability" contains a program-level table containing affected population estimates for the twelve different matching scenarios. Each row in the table represents the affected populations in the evaluation study sample, the rest of BTOP, and all of BTOP for a unique combination of the definition of broadband, any adjustments made to the data, and the matching group. Section 5.2 provides a full description of the analysis leading to these results. The following results are used in this analysis:
  - NTIA Primary, NTIA Sensitivity, NTIA Forward-Looking Primary, NTIA Forward-Looking Sensitivity, NTIA Backward-Looking Primary, NTIA Backward-Looking Sensitivity, NOFA Primary, NOFA Sensitivity, NOFA Forward-Looking Primary, NOFA Forward-Looking Sensitivity, NOFA Backward-Looking Primary, and NOFA Backward-Looking Sensitivity affected population estimates

The R script loads the statistical analysis results discussed above. The table is reshaped and formatted. After these manipulation steps, a table of results is written to the "Table 52" Excel file. Table 52 below presents the results of the R script.

Table 52. Estimated Total Population with Broadband Availability Due to BTOP

Definition	Match Type	Grants	Availability	Forward Looking	Backward Looking
NOFA	Primary	Evaluation Study Sample	176,402	78,944	58,873
		Rest of BTOP	1,009,800	451,906	337,014
		All BTOP	**1,186,202**	**530,850**	**395,887**
	Sensitivity	Evaluation Study Sample	126,713	42,423	27,306
		Rest of BTOP	725,358	242,845	156,311
		All BTOP	**852,071**	**285,268**	**183,617**
NTIA	Primary	Evaluation Study Sample	645,510	594,469	428,691
		Rest of BTOP	3,695,167	3,402,993	2,454,009
		All BTOP	**4,340,677**	**3,997,462**	**2,882,700**
	Sensitivity	Evaluation Study Sample	612,576	559,618	336,942
		Rest of BTOP	3,506,641	3,203,486	1,928,799
		All BTOP	**4,119,217**	**3,763,104**	**2,265,742**

12.6 Final Report Appendix D.6 Developing Confidence Intervals

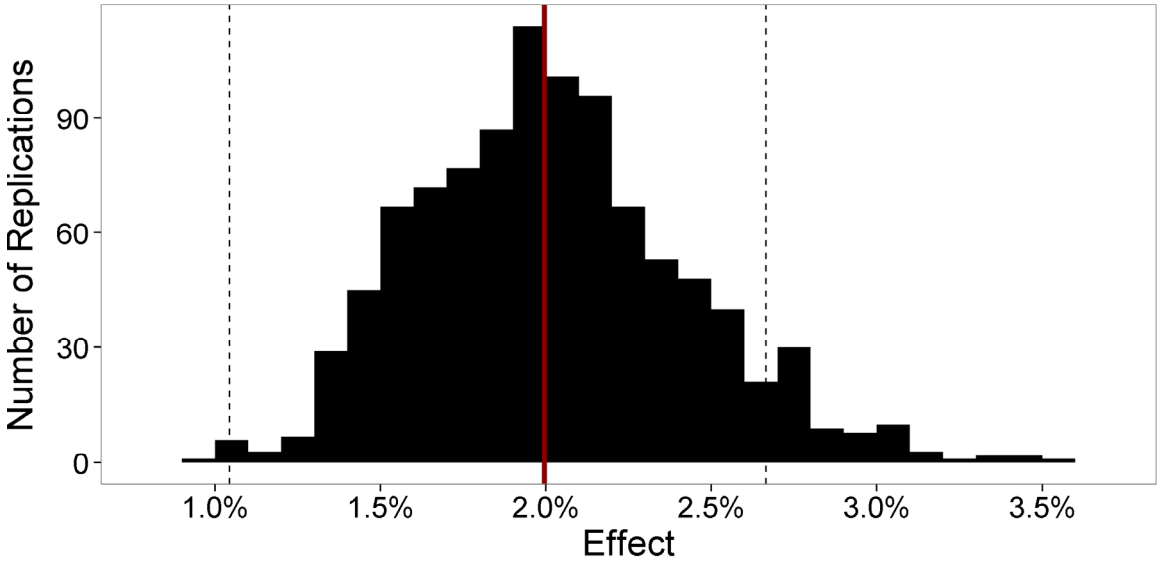
12.6.1 Final Report Figure 13. Resampled Effect and Confidence Bands for NTIA Broadband Availability

Final Report Figure 13 summarizes results of the resampling method ASR uses to examine the robustness of difference-in-differences estimates to individual treatment-to-control matches for the NTIA definition of broadband based on Primary matching results using unadjusted data (the base case). The "Figure 13" R script reproduces this figure using the following statistical analysis:

- "Effect on Availability" contains a table with difference-in-differences estimates of the base case effect of BTOP for 1,000 resampled replications. The table also contains the estimated effect using the full evaluation study sample and the lower and upper 95 percent confidence limits around this effect based on the resampling results. Section 5.2 provides a full description of the analysis leading to these results.

The R script loads the statistical analysis results described above. A histogram is created using the resampling results and written to the "Figure 13" PNG file. Figure 13 below presents the results of the R script.

Figure 13. Resampled Effect and Confidence Bands for NTIA Broadband Availability



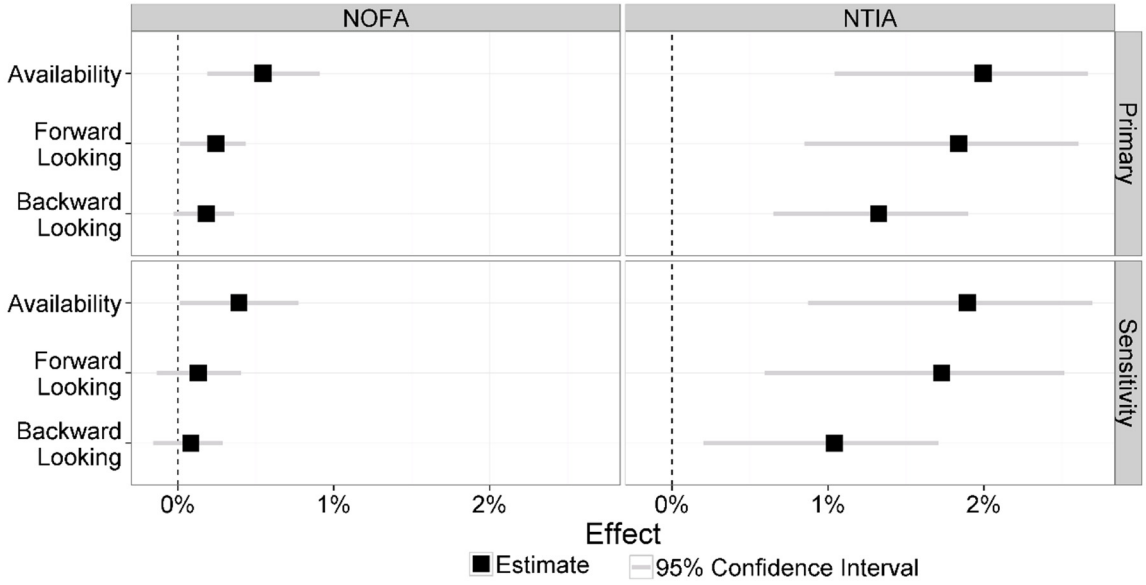
12.6.2 Final Report Figure 14. Comparison of Estimated Effects with Confidence Intervals

Final Report Figure 14 summarizes the robustness of difference-in-differences estimates to individual treatment-to-control matches according to the resampling method used by ASR for the twelve matching results. The "Figure 14" R script reproduces this figure using the following statistical analysis:

- "Effect on Availability" contains a table with difference-in-differences estimates of the effect of BTOP and the lower and upper 95 percent confidence limits of this effect for the twelve different matching results. Section 5.2 provides a full description of the analysis leading to these results.

The R script loads the statistical analysis results described above. The data are reshaped, and a forest plot based on the reshaped data is created and written to the "Figure 14" PNG file. Figure 14 below presents the results of the R script. ASR used an image editor to change the orientation of the legend from vertical to horizontal.

Figure 14. Comparison of Estimated Effects with Confidence Intervals



12.6.3 Final Report Table 53. Difference-in-Differences Estimates and Lower and Upper Confidence Values for Matched Pair Groups

Final Report Table 53 summarizes the robustness of difference-in-differences estimates to individual treatment-to-control matches according to the resampling method used by ASR for the twelve matching results. The "Table 53" R script reproduces this table using the following statistical analysis:

- "Effect on Availability" contains a table with difference-in-differences estimates of the effect of BTOP and the lower and upper 95 percent confidence limits of this effect for the twelve different matching results. Section 5.2 provides a full description of the analysis leading to these results.

The R script loads the statistical analysis results described above. The data are reshaped and written to the "Table 53" Excel file. Table 53 below presents the results of the R script.

Table 53. Difference-in-Differences Estimates and Lower and Upper Confidence Values for Matched Pair Groups

Match Type	Adjustment	NOFA			NTIA		
		L95	Est.	U95	L95	Est.	U95
Primary	Availability	0.19%	0.55%	0.91%	1.04%	2.00%	2.66%
	Forward Looking	0.01%	0.24%	0.44%	0.85%	1.84%	2.61%
	Backward Looking	-0.03%	0.18%	0.36%	0.65%	1.32%	1.90%
Sensitivity	Availability	0.01%	0.39%	0.77%	0.87%	1.89%	2.70%
	Forward Looking	-0.14%	0.13%	0.41%	0.59%	1.73%	2.52%
	Backward Looking	-0.16%	0.08%	0.29%	0.20%	1.04%	1.71%

12.7 Final Report Appendix D.8 Long-Term Impact of BTOP CCI Infrastructure on GDP

12.7.1 Final Report Table 55. Extrapolated Total Benefit from Increased Output Due to BTOP (Annual, Million USD)

Czernich et al. (2011) estimated the introduction of broadband availability in twenty Organisation for Economic Co-operation and Development (OECD) countries increased GDP by 2.7 to 3.9 percent.[201] Using this estimate, census blocks with no broadband availability that later received availability would experience an increase in GDP. If 1 percent of the population in a county resided in such census blocks, county GDP would be expected to increase by 0.027 percent.[202] LECG Ltd. (2009) estimated that a 1 percentage point increase in broadband availability raised productivity by 0.1 percent in countries with medium to high levels of information and communications technology.[203]

Final Report Table 55 presents the results of these extrapolations. ASR used the estimated effects of BTOP on broadband availability in conjunction with publicly available data to extrapolate long-term benefits due to BTOP. The "D.8 Long-term Impact of BTOP CCI Infrastructure on GDP" R script loads statistical analysis results and prepared data, extrapolates benefits, and creates a summary table for the *Final Report*.

ASR used the following statistical analysis to extrapolate the benefits of increased GDP due to BTOP:

- "Extrapolation Table" contains a table representing every estimated effect of BTOP for every county in the BTOP service area. That is, the table contains one record for every combination of county FIPS code, definition of broadband, adjustments made to the data, and match type, with values for the estimated effect, the rate of adoption by households with availability, and whether or not the county is in an evaluation study sample service area. Section 5.4 provides a full description of the analysis leading to these results.

The script uses the following prepared data to extrapolate the benefits of increased GDP due to BTOP:

- "NBM Population Statistics" is a county-level table containing population, demographic, and broadband availability statistics derived from NBM data provided to ASR by NTIA.[204] Section 4.12 provides a full description of "NBM Population Statistics." The following county-level attribute from the June 30, 2011 NBM release was used:
  - population
- "Gross County Product" contains a county-level table of estimates of county-level GDP based on decompositions of state-level GDP. Section 4.9 provides a full description of "Gross County Product." The following county-level attribute was used:
  - county GDP

The R script loads the statistical analysis and prepared data discussed above. ASR followed the same process to extrapolate the benefits of increased GDP due to BTOP based on the two different estimates produced by Czernich et al. (2011) and LECG Ltd. (2009). That is, the process below was carried out twice, once for each study coefficient:

- The extrapolation table was copied to a new table
- County GDP was joined to the table by county FIPS code
- County population was joined to the table by county FIPS code

- County GDP per capita was calculated
- The median of non-missing values of county GDP per capita was calculated
- Missing values of county GDP were replaced with county population times the median of non-missing values of county GDP per capita
- Benefit values (in millions) were extrapolated by multiplying county GDP by the effect of BTOP times the study coefficient, then dividing by 1 million

The extrapolation results for the two literature sources were then combined, and the benefits due to BTOP were aggregated to the evaluation study sample level (in sample, out of sample, and total). The combined, aggregated results were written to the "D.8 Long-term Impact of BTOP CCI Infrastructure on GDP - Table 55" Excel file. Table 55 below presents the results of the R script.

Table 55. Extrapolated Total Benefit from Increased Output Due to BTOP (Annual, Million USD)

Study	Grants	Match Type	Availability		Forward Looking		Backward Looking	
			NOFA	NTIA	NOFA	NTIA	NOFA	NTIA
Czernich et al. (2011)	Evaluation Study Sample	Primary	206	755	92	695	69	501
		Sensitivity	148	717	50	655	32	394
	Rest of BTOP	Primary	1,340	4,903	600	4,516	447	3,256
		Sensitivity	963	4,653	322	4,251	207	2,559
	All BTOP	Primary	1,546	5,659	692	5,211	516	3,758
		Sensitivity	1,111	5,370	372	4,906	239	2,954
LECG Ltd. (2009)	Evaluation Study Sample	Primary	764	2,797	342	2,576	255	1,857
		Sensitivity	549	2,654	184	2,425	118	1,460
	Rest of BTOP	Primary	4,963	18,161	2,221	16,725	1,656	12,061
		Sensitivity	3,565	17,234	1,194	15,744	768	9,479
	All BTOP	Primary	5,727	20,957	2,563	19,300	1,911	13,918
		Sensitivity	4,114	19,888	1,377	18,169	887	10,939

12.8 Final Report Appendix D.9 Long-term Impact of BTOP CCI Infrastructure on Employment

12.8.1 Final Report Table 56. Extrapolated Total Increase in Employment Due to BTOP

Kolko (2010) estimated that one standard deviation increase in broadband availability in ZIP Code Tabulation Areas (ZCTA) across the United States increased employment by 0.085 standard deviations.[205]

Gillett et al. (2006) estimated the introduction of broadband availability in ZIP Codes across the United States increased the employment growth rate by 1.44 percent over a four-year period.[206] This is equivalent to a 0.359 percent annual employment growth rate. Simplifying this, a 1

percentage point increase in broadband availability in a county increases the employment growth rate by 0.00359 percent.

Final Report Tables 56-59 present the results of these extrapolations. ASR used the estimated effects of BTOP on broadband availability in conjunction with publicly available data to extrapolate long-term benefits due to BTOP. The "D.9 Long-term Impact of BTOP CCI Infrastructure on Employment" R script loads statistical analysis results and prepared data, extrapolates benefits, and creates summary tables for the *Final Report*. The script uses the following statistical analysis to extrapolate the benefits of increased employment due to BTOP:

- "Extrapolation Table" contains a table representing every estimated effect of BTOP for every county in the BTOP service area. That is, the table contains with one record for every combination of county FIPS code, definition of broadband, adjustments made to the data, and match type, with values for the estimated effect, rate of adoption by households with availability, and whether or not the county is in an evaluation study sample service area. Section 5.4 provides a full description of the analysis leading to these results.

The script uses the following prepared data to extrapolate the benefits of increased GDP due to BTOP:

- "Labor Force Statistics" contains a county-level table with attributes derived from data published by BLS.[207] Section 4.11 provides a full description of "Labor Force Statistics." The following 2011 county-level attributes were used:
  - total employment
  - average annual wage
- "Standard Deviation of Availability Growth" is a national level table with attributes derived from NBM data provided to ASR by NTIA.[208] Section 4.26 provides a full description of "Standard Deviation of Availability Growth." The following national attributes were used:
  - standard deviation of county-level change in broadband from June 30, 2011 to June 30, 2013 for two definitions of broadband and three availability rates (availability, forward looking, and backward looking)
- "Standard Deviation of Employment Growth" is a national level table with attributes derived from data published by BLS.[209] Section 4.27 provides a full description of "Standard Deviation of Employment Growth." The following national attribute was used:
  - standard deviation of county-level change in total employment from 2011 to 2012

The R script loads the statistical analysis and prepared data discussed above. The literature estimates of Kolko (2010) and Gillett et al. (2006), discussed above, measured the effect of broadband on employment differently. ASR followed the following process to extrapolate benefits according to the results of Kolko (2010):

- The extrapolation table was copied to a new table representing estimates based on Kolko (2010)
- Labor force statistics were joined to the table by county FIPS code
- Standard deviations of availability growth were joined to the table by broadband definition and availability rate
- The effect of broadband was divided by standard deviation of availability growth to determine the effect in standard deviations
- The increase in employment growth due to BTOP was extrapolated by multiplying the effect in standard deviations by 0.085 times the standard deviation of employment growth
- The increase in total employment due to BTOP was extrapolated by total employment by the extrapolated increase in employment growth

- Benefit values (in millions) were extrapolated by multiplying the increase in total employment due to BTOP by average annual wage, then dividing by 1 million

ASR followed the following process to extrapolate benefits according to the results of Gillett et al. (2006):

- The extrapolation table was copied to a new table representing estimates based on Gillett et al. (2006)
- Labor force statistics were joined to the table by county FIPS code
- The annual increase in total employment due to BTOP was extrapolated by multiplying the effect of BTOP by 0.00359 times total employment
- Benefit values (in millions) were extrapolated by multiplying the increase in total employment due to BTOP by average annual wage, then dividing by 1 million

Because the two literature sources estimated the effect of broadband on employment differently, ASR reports extrapolations based on the sources separately. However, the script combines the results in order to ensure consistency between the two. The extrapolated increases in employment based on the results of Kolko (2010) were aggregated and written to the "D.9 Long-term Impact of BTOP CCI Infrastructure on Employment - Table 56" Excel file. Table 56 below presents the results of the R script.

Table 56. Extrapolated Total Increase in Employment Due to BTOP

Study	Grants	Match Type	Availability		Forward Looking		Backward Looking	
			NOFA	NTIA	NOFA	NTIA	NOFA	NTIA
Kolko (2010)	Evaluation Study Sample	Primary	1,804	3,386	815	3,686	759	2,754
		Sensitivity	1,296	3,213	438	3,470	352	2,164
	Rest of BTOP	Primary	10,423	19,563	4,707	21,298	4,385	15,911
		Sensitivity	7,487	18,565	2,529	20,050	2,034	12,506
	All BTOP	Primary	12,227	22,949	5,521	24,984	5,143	18,664
		Sensitivity	8,783	21,778	2,967	23,519	2,386	14,670

12.8.2 Final Report Table 57. Extrapolated One-Year Increase in Employment Due to BTOP

The extrapolated annual increases in employment based on the results of Gillett et al. (2006) were aggregated and written to the "D.9 Long-term Impact of BTOP CCI Infrastructure on Employment - Table 57" Excel file. Table 57 below presents the results of the R script.

Table 57. Extrapolated One-Year Increase in Employment Due to BTOP

Study	Grants	Match Type	Availability		Forward Looking		Backward Looking	
			NOFA	NTIA	NOFA	NTIA	NOFA	NTIA
Gillett et al. (2006)	Evaluation Study Sample	Primary	280	1,024	125	943	93	680
		Sensitivity	201	972	67	888	43	535
	Rest of BTOP	Primary	1,617	5,917	724	5,449	540	3,929
		Sensitivity	1,161	5,615	389	5,130	250	3,088
	All BTOP	Primary	1,897	6,941	849	6,392	633	4,610
		Sensitivity	1,362	6,587	456	6,017	294	3,623

12.8.3 Final Report Table 58. Extrapolated Total Increase in Income Due to Total Increase in Employment (Annual, Million USD)

The extrapolated economic benefits of increased employment due to BTOP based on the results of Kolko (2010) were aggregated and written to the "D.9 Long-term Impact of BTOP CCI Infrastructure on Employment - Table 58" Excel file. Table 58 below presents the results of the R script.

Table 58. Extrapolated Total Increase in Income Due to Total Increase in Employment (Annual, Million USD)

Study	Grants	Match Type	Availability		Forward Looking		Backward Looking	
			NOFA	NTIA	NOFA	NTIA	NOFA	NTIA
Kolko (2010)	Evaluation Study Sample	Primary	75	141	34	154	32	115
		Sensitivity	54	134	18	145	15	90
	Rest of BTOP	Primary	502	942	227	1,026	211	766
		Sensitivity	361	894	122	966	98	602
	All BTOP	Primary	577	1,084	261	1,180	243	881
		Sensitivity	415	1,028	140	1,111	113	693

12.8.4 Final Report Table 59. Extrapolated Total Benefit Due to One-Year Increase in Employment (Million USD)

The extrapolated annual economic benefits of increased employment due to BTOP based on the results of Gillett et al. (2006) were aggregated and written to the "D.9 Long-term Impact of BTOP CCI Infrastructure on Employment - Table 59" Excel file. Table 59 below presents the results of the R script.

Table 59. Extrapolated Total Benefit Due to One-Year Increase in Employment (Million USD)

Study	Grants	Match Type	Availability		Forward Looking		Backward Looking	
			NOFA	NTIA	NOFA	NTIA	NOFA	NTIA
Gillett et al. (2006)	Evaluation Study Sample	Primary	12	43	5	39	4	28
		Sensitivity	8	41	3	37	2	22
	Rest of BTOP	Primary	78	285	35	263	26	189
		Sensitivity	56	271	19	247	12	149
	All BTOP	Primary	90	328	40	302	30	218
		Sensitivity	64	311	22	284	14	171

12.9 Final Report Appendix D.10 Value to New Subscribers

12.9.1 Final Report Table 60. Extrapolated Total Increased Value to Consumers Due to BTOP (Annual, Million USD)

The Allen Consulting Group (2010) finds the value of broadband Internet access to the average American household is about 3.4 percent of average household income.[210] *Final Report Table 60* presents the results of these extrapolations. ASR used the estimated effects of BTOP on broadband availability in conjunction with publicly available data to extrapolate long-term benefits due to BTOP. The "D.10 Value to New Subscribers" R script loads statistical analysis results and prepared data, extrapolates benefits, and creates a summary table for the *Final Report*.

ASR used the following statistical analysis to extrapolate the benefits of increased employment due to BTOP:

- "Extrapolation Table" contains a table representing every estimated effect of BTOP for every county in the BTOP service area. That is, the table contains one record for every combination of county FIPS code, broadband definition, adjustments made to the data, and match type, with values for the estimated effect, rate of adoption by households with availability, and whether or not the county is in an evaluation study sample service area. Section 5.4 provides a full description of the analysis leading to these results.

ASR uses the following prepared data to extrapolate the benefits of increased GDP due to BTOP:

- "Household Statistics" contains a county-level table with attributes derived from NBM data provided to ASR by NTIA and data published by the Census Bureau.[211] Section 4.10 provides a full description of "Labor Force Statistics." The following county-level attributes were used:
  - NBM June 30, 2011 data: total households
  - ACS (2006-2010) data: average household income

The R script loads the statistical analysis and prepared data discussed above. ASR followed the following process to extrapolate benefits according to the results of Allen Consulting Group (2010):

- The extrapolation table was copied to a new table representing estimates based on Allen Consulting Group (2010)
- Household statistics were joined to the table by county FIPS code

- The increase in adopting households due to BTOP was extrapolated as the effect on the availability rate times the rate of adoption by households with availability times the total number of households
- Benefit values (in millions) were extrapolated by multiplying the increase in adopting households due to BTOP by average household income times 0.034, then dividing by 1 million

The extrapolated increases in employment based on the results of Allen Consulting Group (2010) were aggregated and written to the "D.10 Value to New Subscribers - Table 60" Excel file. Table 60 below presents the results of the R script.

Table 60. Extrapolated Total Increased Value to Consumers Due to BTOP (Annual, Million USD)

Grants	Match Type	Availability		Forward Looking		Backward Looking	
		NOFA	NTIA	NOFA	NTIA	NOFA	NTIA
Evaluation Study Sample	Primary	95	347	48	363	32	230
	Sensitivity	68	329	26	341	15	181
Rest of BTOP	Primary	628	2,298	315	2,373	210	1,526
	Sensitivity	451	2,181	169	2,234	97	1,200
All BTOP	Primary	723	2,645	363	2,735	241	1,757
	Sensitivity	519	2,510	195	2,575	112	1,381

Notes

[1] National Telecommunications and Information Administration, "BTOP Evaluation Study," *Program Reports*, 2013, http://www2.ntia.doc.gov/BTOP-Reports#evaluation.

[2] National Telecommunications and Information Administration, "Statement of Work for Broadband Technology Opportunities Program (BTOP) Evaluation Study," July 26, 2010, http://www2.ntia.doc.gov/files/btop_sow.pdf.

[3] ASR Analytics, *BTOP Evaluation Study Design* (Potomac, MD, January 30, 2012), http://www.ntia.doc.gov/files/ntia/publications/study_design_order_number_d10pd18645_-_revised_2012-01-30.pdf.

[4] R Core Team, "R: A Language and Environment for Statistical Computing" (Vienna, Austria, 2013), http://www.r-project.org/.

[5] Matt Dowle et al., "Data.table: Extension of Data.frame," *R Package Version 1.9.2*, 2014, http://cran.r-project.org/package=data.table.

[6] Adrian A. Dragulescu, "Xlsx: Read, Write, Format Excel 2007 and Excel 97/2000/XP/2003 Files," *R Package Version 0.5.5*, 2013, http://cran.r-project.org/package=xlsx.

[7] Adrian A. Dragulescu, "Xlsxjars: Package Required POI Jars for the Xlsx Package," *R Package Version 0.6.0*, 2014, http://cran.r-project.org/package=xlsxjars.

Simon Urbanek, "rJava: Low-Level R to Java Interface," *R Package Version 0.9-6*, 2013, http://cran.r-project.org/package=rJava.

[8] Hadley Wickham, "Reshaping Data with the Reshape Package," *Journal of Statistical Software* 21, no. 12 (2007): 1–20, http://www.jstatsoft.org/v21/i12/.

[9] Hadley Wickham, "The Split-Apply-Combine Strategy for Data Analysis," *Journal of Statistical Software* 40, no. 1 (2011): 1–29, http://www.jstatsoft.org/v40/i01/.

[10] Hadley Wickham, *ggplot2: Elegant Graphics for Data Analysis* (New York: Springer, 2009), http://had.co.nz/ggplot2/book.

[11] Hadley Wickham, "Scales: Scale Functions for Graphics," *R Package Version 0.2.3*, 2012, http://cran.r-project.org/package=scales.

[12] Duncan Murdoch, "Tables: Formula-Driven Table Generation," *R Package Version 0.7.64*, 2013, http://cran.r-project.org/package=tables.

[13] Frank E. Harrell Jr. and Charles Dupont, "Hmisc: Harrell Miscellaneous," *R Package Version 3.14-4*, 2014, http://cran.r-project.org/package=Hmisc.

[14] Deepayan Sarkar, *Lattice: Multivariate Data Visualization with R* (New York: Springer, 2008), http://lmdvr.r-forge.r-project.org.

Terry M. Therneau, "A Package for Survival Analysis in S," *R Package Version 2.37-7*, 2014, http://cran.r-project.org/package=survival.

Terry M. Therneau and Patricia M. Grambsch, *Modeling Survival Data: Extending the Cox Model* (New York: Springer, 2000), 0-387-98784-3.

Achim Zeileis and Yves Croissant, "Extended Model Formulas in R: Multiple Parts and Multiple Responses," *Journal of Statistical Software* 34, no. 1 (2010): 1–13, http://www.jstatsoft.org/v34/i01/.

[15] Garrett Grolemund and Hadley Wickham, "Dates and Times Made Easy with Lubridate," *Journal of Statistical Software* 40, no. 3 (2011): 1–25, http://www.jstatsoft.org/v40/i03/.

[16] Jasjeet S. Sekhon, "Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R," *Journal of Statistical Software* 42, no. 7 (June 2011).

[17] W.N. Venables and Brian D. Ripley, *Modern Applied Statistics with S*, 4th ed. (New York: Springer, 2002), http://www.stats.ox.ac.uk/pub/MASS4.

[18] Angelo Canty and Brian D. Ripley, "Boot: Bootstrap R (S-Plus) Functions," *R Package Version 1.3-9*, 2013, http://cran.r-project.org/web/packages/boot/index.html.

A.C. Davison and D. V. Hinkley, *Bootstrap Methods and Their Applications* (Cambridge, UK: Cambridge University Press, 1997).

[19] IMPLAN Group LLC, "Economic Impact Analysis," July 10, 2013, http://implan.com/V4/Index.php.

[20] U.S. Department of Energy, *Economic Impact of Recovery Act Investments in the Smart Grid*, 2013, http://www.smartgrid.gov/sites/default/files/doc/files/Smart Grid Economic Impact Report.pdf.

U.S. Department of Agriculture, *Estimating The Impact on Employment of USDA's Programs in the American Recovery and Reinvestment Act (ARRA)*, 2013.

Terry L. Clower and Bernard L. Weinstein, *Economic and Fiscal Impacts of the Corps of Engineers' Trinity River Vision Project in Tarrant County Texas*, 2005, http://www.trinityrivervision.org/docs/PDFDocuments/pr_05impacts.pdf.

[21] United States Department of Agriculture, *Advancing Broadband*, *Broadband Initiatives Program Awards Report*, January 2011, http://www.rurdev.usda.gov/supportdocuments/RBBreport_V5ForWeb.pdf.

[22] United States Department of Agriculture, *Advancing Broadband*.

[23] ProPublica, "Recovery Tracker Eye on the Stimulus," 2012, http://projects.propublica.org/recovery/gov_entities/12e2/list/4.

[24] ProPublica, "Recovery Tracker Eye on the Stimulus."

[25] For more information, visit http://www2.ntia.doc.gov.

[26] National Telecommunications and Information Administration, "NTIA's BTOP Map," *Connecting America's Communities* (Washington, DC, August 15, 2013), http://www2.ntia.doc.gov/BTOPmap/.

[27] National Telecommunications and Information Administration, "NTIA's BTOP Map."

National Telecommunications and Information Administration, "Post-Award Monitoring Database," April 11, 2014.

[28] United States Census Bureau, "2010 TIGER/Line® Shapefiles" (Washington, DC, June 19, 2012), http://www.census.gov/geo/www/cob/cbf_counties.html.

[29] United States Bureau of Economic Analysis, "Local Area Personal Income," *Regional Economic Accounts* (Washington, DC, April 01, 2014), http://www.bea.gov/regional/index.htm.

[30] United States Bureau of Economic Analysis, "Gross Domestic Product by State," *Regional Economic Accounts* (Washington, DC, April 01, 2014), http://www.bea.gov/regional/index.htm.

[31] United States Bureau of Labor Statistics, "Quarterly Census of Employment and Wages" (Washington, DC, April 01, 2014), http://www.bls.gov/cew/.

[32] United States Bureau of Labor Statistics, "Local Area Unemployment Statistics" (Washington, DC, April 01, 2014), http://www.bls.gov/lau/.

[33] United States Department of Labor and United States Bureau of Labor Statistics, *Geographic Profile of Employment and Unemployment, 2011*, *Bulletin 2774* (Washington, DC, September 2012), http://www.bls.gov/opub/gp/pdf/gp11full.pdf.

[34] National Telecommunications and Information Administration, "NTIA's BTOP Map."

[35] National Telecommunications and Information Administration, "BTOP Map Fact Sheet," *Connecting America's Communities* (Washington, DC, March 2012), http://www2.ntia.doc.gov/BTOPmap/data/btop-map-fact-sheet.pdf.

[36] National Telecommunications and Information Administration, "NTIA's BTOP Map."

[37] National Telecommunications and Information Administration, "BTOP Map Fact Sheet."

[38] United States Census Bureau, "2010 TIGER/Line® Shapefiles."

[39] National Telecommunications and Information Administration, "NTIA's BTOP Map."

[40] National Telecommunications and Information Administration, "BTOP Map Fact Sheet."

[41] National Telecommunications and Information Administration, "NTIA's BTOP Map."

[42] United States Census Bureau, "2010 TIGER/Line® Shapefiles."

[43] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[44] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[45] ASR Analytics, *Clearwave Communications Comprehensive Community Infrastructure*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, December 03, 2013).

ASR Analytics, *Executive Office of the State of West Virginia Comprehensive Community Infrastructure*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, March 25, 2014).

ASR Analytics, *Lane Council of Governments Comprehensive Community Infrastructure*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, April 04, 2014).

ASR Analytics, *Massachusetts Technology Park Comprehensive Community Infrastructure*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, March 19, 2014).

ASR Analytics, *MCNC Comprehensive Community Infrastructure*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, March 04, 2014).

ASR Analytics, *Merit Network, Inc. Comprehensive Community Infrastructure*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, March 04, 2014).

ASR Analytics, *Mid-Atlantic Broadband Cooperative Comprehensive Community Infrastructure*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, February 06, 2014).

ASR Analytics, *OneCommunity Comprehensive Community Infrastructure*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, April 21, 2014).

ASR Analytics, *OSHEAN, Inc. Comprehensive Community Infrastructure*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, March 10, 2014).

ASR Analytics, *South Dakota Network Comprehensive Community Infrastructure*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, February 05, 2014).

ASR Analytics, *University of Arkansas System Comprehensive Community Infrastructure*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, March 11, 2014).

ASR Analytics, *Zayo Group Comprehensive Community Infrastructure*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, February 06, 2014).

[46] For more information, visit http://www2.ntia.doc.gov.

National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[47] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[48] In some cases, geocoding institution names and addresses or institution names and location names (e.g., cities) was required to determine geographic coordinates.

United States Census Bureau, "2010 TIGER/Line® Shapefiles."

[49] ASR Analytics, *Clearwave Communications Comprehensive Community Infrastructure*.

ASR Analytics, *Executive Office of the State of West Virginia Comprehensive Community Infrastructure*.

ASR Analytics, *Lane Council of Governments Comprehensive Community Infrastructure*.

ASR Analytics, *Massachusetts Technology Park Comprehensive Community Infrastructure*.

ASR Analytics, *MCNC Comprehensive Community Infrastructure*.

ASR Analytics, *Merit Network, Inc. Comprehensive Community Infrastructure*.

ASR Analytics, *Mid-Atlantic Broadband Cooperative Comprehensive Community Infrastructure*.

ASR Analytics, *OneCommunity Comprehensive Community Infrastructure*.

ASR Analytics, *OSHEAN, Inc. Comprehensive Community Infrastructure*.

ASR Analytics, *South Dakota Network Comprehensive Community Infrastructure*.

ASR Analytics, *University of Arkansas System Comprehensive Community Infrastructure*.

ASR Analytics, *Zayo Group Comprehensive Community Infrastructure*.

[50] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[51] Not all of these variables exist as independent fields in the Summary File. For example, populations of particular age ranges reported in the Summary File are added together to produce the age ranges listed here.

United States Census Bureau, "2005-2009 ACS 5-Year Summary File," *American Community Survey* (Washington, DC, December 14, 2010), http://www.census.gov/acs/www/data_documentation/2009_release/.

[52] United States Census Bureau, "2006-2010 ACS 5-Year Summary File," *American Community Survey* (Washington, DC, December 08, 2011), http://www.census.gov/acs/www/data_documentation/2010_release/.

[53] United States Census Bureau, "County Business Patterns" (Washington, DC, April 01, 2014), http://www.census.gov/econ/cbp/.

[54] United States Census Bureau, "2010 TIGER/Line® Shapefiles."

[55] United States Census Bureau, "County Business Patterns."

[56] United States Department of Labor, "Changes in Basic Minimum Wages in Non-Farm Employment Under State Law: Selected Years 1968 to 2013," *Wage and Hour Division (WHD)* (Washington, DC, December 2013), http://www.dol.gov/whd/state/stateMinWageHis.htm.

[57] Federal Communications Commission, "Census Tract Information Mapped for Internet Access Services Faster than 200 Kbps in at Least One Direction," *Local Telephone Competition and Broadband Deployment* (Washington, DC, March 04, 2014), http://transition.fcc.gov/wcb/iatd/comp.html.

[58] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[59] Defunded grant status was taken from the 19th BTOP Quarterly Program Status Report.

National Telecommunications and Information Administration, *Broadband Technology Opportunities Program (BTOP) Quarterly Program Status Report*, February 12, 2014, http://www.ntia.doc.gov/files/ntia/publications/ntia_btop_19th_quarterly_report.pdf.

[60] National Telecommunications and Information Administration and Federal Communications Commission, "State Broadband Initiative," May 12, 2014.

[61] For more information, visit http://www2.ntia.doc.gov/sbdd.

[62] GeoLytics, "Estimates Professional," July 10, 2014, http://geolytics.com/USCensus,Estimates-Professional,Products.asp.

[63] National Telecommunications and Information Administration and Federal Communications Commission, "State Broadband Initiative."

[64] For more information, visit http://www2.ntia.doc.gov/sbdd.

[65] GeoLytics, "Estimates Professional."

[66] Description is available for SBA grants only.

National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[67] For PCC grants, this was based on the program name only. For SBA grants, this was based on program name and description.

[68] Description is available for SBA grants only.

National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[69] For PCC grants, this was based on the program name only. For SBA grants, this was based on program name and description.

[70] ASR Analytics, "C.K. Blandin Foundation: Minnesota Intelligent Rural Communities Sustainable Broadband Adoption (SBA)," *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, June 18, 2012).

ASR Analytics, “California Emerging Technology Fund: Broadband Awareness and Adoption Sustainable Broadband Adoption (SBA),” *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, June 18, 2012).

ASR Analytics, “Cambridge Housing Authority Community Computer Centers Public Computer Centers (PCC),” *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, June 18, 2012).

ASR Analytics, “City of Chicago: Smart Chicago Sustainable Broadband Adoption (SBA),” *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, June 18, 2012).

ASR Analytics, “Connect Arkansas Expanding Broadband Use in Arkansas through Education Sustainable Broadband Adoption (SBA),” *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, June 18, 2012).

ASR Analytics, “Delaware Division of Libraries: Job/Learning Labs Public Computer Center (PCC),” *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, June 18, 2012).

ASR Analytics, “Florida A&M University: The FAMU Center for Public Computing and Workforce Development Public Computer Center (PCC),” *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, June 18, 2012).

ASR Analytics, “Foundation for California Community Colleges: California Connects Sustainable Broadband Adoption (SBA),” *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, June 18, 2012).

ASR Analytics, “Future Generations Graduate School: Equipping West Virginia’s Fire and Rescue Squads with Technology and Training to Serve Communities Sustainable Broadband Adoption (SBA),” *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, June 18, 2012).

ASR Analytics, “Las Vegas Clark County Urban League: Access to Computer Technology and Instruction in Online Networking (ACTION) Public Computer Center (PCC),” *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, June 18, 2012).

ASR Analytics, “Michigan State University: Evidence Based Computer Centers II Public Computer Centers (PCC),” *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, June 18, 2012).

ASR Analytics, “SC State Board for Technical and Comprehensive Education: Reach for Success Public Computer Centers (PCC),” *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, June 18, 2012).

ASR Analytics, “Technology For All, Inc.: Texas Connects Coalition Public Computer Center (PCC),” *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, June 18, 2012).

ASR Analytics, “Urban Affairs Coalition: Freedom Rings Sustainable Broadband Adoption (SBA),” *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, June 18, 2012).

ASR Analytics, “WorkForce West Virginia: One-Stop Public Computer Center Modernization Public Computer Center (PCC),” *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, June 18, 2012).

[71] ASR Analytics, *Progress towards BTOP Goals: Interim Report on PCC and SBA Case Studies*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*

(Potomac, MD, October 15, 2012), http://www.ntia.doc.gov/report/2012/progress-towards-btop-goals-interim-report-pcc-and-sba-case-studies.

ASR Analytics, *Short-Term Economic Impacts Report*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, September 2013).

[72] National Telecommunications and Information Administration, "NTIA's BTOP Map."

[73] For more information, visit http://www2.ntia.doc.gov.

[74] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[75] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[76] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[77] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[78] National Telecommunications and Information Administration, "Application Search," *Broadband USA*, 2014, http://www.ntia.doc.gov/legacy/broadbandgrants/applications/results.htm.

[79] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[80] ASR Analytics, *Progress towards BTOP Goals: Interim Report on PCC and SBA Case Studies*.

ASR Analytics, *Progress towards BTOP Goals: Second Interim Report on CCI, PCC, and SBA Case Studies*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, 2014).

ASR Analytics, *Short-Term Economic Impacts Report*.

[81] Additional data were required for the SC Reach for Success PCC grant. Section 3.40 provides additional detail.

[82] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[83] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[84] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[85] ASR Analytics, *Progress towards BTOP Goals: Interim Report on PCC and SBA Case Studies*.

[86] United States Census Bureau, "2006-2010 ACS 5-Year Summary File."

[87] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[88] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

ASR Analytics, *Clearwave Communications Comprehensive Community Infrastructure*.

ASR Analytics, *Executive Office of the State of West Virginia Comprehensive Community Infrastructure*.

ASR Analytics, *Lane Council of Governments Comprehensive Community Infrastructure*.

ASR Analytics, *Massachusetts Technology Park Comprehensive Community Infrastructure*.

ASR Analytics, *MCNC Comprehensive Community Infrastructure*.

ASR Analytics, *Merit Network, Inc. Comprehensive Community Infrastructure*.

ASR Analytics, *Mid-Atlantic Broadband Cooperative Comprehensive Community Infrastructure*.

ASR Analytics, *OneCommunity Comprehensive Community Infrastructure*.

ASR Analytics, *OSHEAN, Inc. Comprehensive Community Infrastructure*.

ASR Analytics, *South Dakota Network Comprehensive Community Infrastructure*.

ASR Analytics, *University of Arkansas System Comprehensive Community Infrastructure*.

ASR Analytics, *Zayo Group Comprehensive Community Infrastructure*.

[89] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[90] United States Bureau of Economic Analysis, "Local Area Personal Income."

United States Bureau of Economic Analysis, "Gross Domestic Product by State."

United States Census Bureau, "County Business Patterns."

[91] United States Bureau of Economic Analysis, "Are Gross Domestic Product (GDP) Estimates Available at the County Level?," *Frequently Asked Questions* (Washington, DC, January 09, 2008), http://www.bea.gov/faq/index.cfm?faq_id=458.

United States Bureau of Economic Analysis, "Gross Domestic Product by State."

[92] Paul W. Bauer and Yoonsoo Lee, *Estimating GSP and Labor Productivity By State*, Policy Discussion Paper (Cleveland, OH, March 2006), http://www.clevelandfed.org/research/policydis/pdpno16.pdf.

John D. Barreca, J. Matthew Fannin, and Joshua D. Detre, *Estimating GDP at the Parish (County) Level: An Evaluation of Alternative Approaches*, Research Bulletin, January 24, 2012, https://www.lsuagcenter.com/en/communications/publications/Publications+Catalog/research/Economics/Estimating-GDP-at-the-Parish-County-Level-An-Evaluation-of-Alternative-Approaches.htm.

[93] United States Census Bureau, "2006-2010 ACS 5-Year Summary File."

National Telecommunications and Information Administration and Federal Communications Commission, "State Broadband Initiative."

[94] United States Bureau of Labor Statistics, "Quarterly Census of Employment and Wages."

United States Bureau of Labor Statistics, "Local Area Unemployment Statistics."

[95] National Telecommunications and Information Administration and Federal Communications Commission, "State Broadband Initiative."

[96] United States Census Bureau, "2005-2009 ACS 5-Year Summary File."

[97] All grants that include the Virgin Islands in their service areas serve the Virgin Islands only

[98] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[99] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[100] United States Bureau of Labor Statistics, "Quarterly Census of Employment and Wages."

United States Bureau of Labor Statistics, "Local Area Unemployment Statistics."

United States Department of Labor and United States Bureau of Labor Statistics, *Geographic Profile of Employment and Unemployment, 2009*, *Bulletin 2748* (Washington, DC, July 2010), http://www.bls.gov/opub/gp/pdf/gp09full.pdf.

United States Census Bureau, "2005-2009 ACS 5-Year Summary File."

[101] Implicit is the assumption that state-level averages uniformly describe averages for lower geographic levels within the state.

[102] Implicit is the assumption that state-level averages uniformly describe averages for lower geographic levels within the state.

[103] All grants that include the Virgin Islands in their service areas serve the Virgin Islands only.

[104] ASR Analytics, "C.K. Blandin Foundation: Minnesota Intelligent Rural Communities Sustainable Broadband Adoption (SBA)."

ASR Analytics, "California Emerging Technology Fund: Broadband Awareness and Adoption Sustainable Broadband Adoption (SBA)."

ASR Analytics, "Cambridge Housing Authority Community Computer Centers Public Computer Centers (PCC)."

ASR Analytics, "City of Chicago: Smart Chicago Sustainable Broadband Adoption (SBA)."

ASR Analytics, "Connect Arkansas Expanding Broadband Use in Arkansas through Education Sustainable Broadband Adoption (SBA)."

ASR Analytics, "Delaware Division of Libraries: Job/Learning Labs Public Computer Center (PCC)."

ASR Analytics, "Florida A&M University: The FAMU Center for Public Computing and Workforce Development Public Computer Center (PCC)."

ASR Analytics, "Foundation for California Community Colleges: California Connects Sustainable Broadband Adoption (SBA)."

ASR Analytics, "Future Generations Graduate School: Equipping West Virginia's Fire and Rescue Squads with Technology and Training to Serve Communities Sustainable Broadband Adoption (SBA)."

ASR Analytics, "Las Vegas Clark County Urban League: Access to Computer Technology and Instruction in Online Networking (ACTION) Public Computer Center (PCC)."

ASR Analytics, "Michigan State University: Evidence Based Computer Centers II Public Computer Centers (PCC)."

ASR Analytics, "SC State Board for Technical and Comprehensive Education: Reach for Success Public Computer Centers (PCC)."

ASR Analytics, "Technology For All, Inc.: Texas Connects Coalition Public Computer Center (PCC)."

ASR Analytics, "Urban Affairs Coalition: Freedom Rings Sustainable Broadband Adoption (SBA)."

ASR Analytics, "WorkForce West Virginia: One-Stop Public Computer Center Modernization Public Computer Center (PCC)."

National Telecommunications and Information Administration, "NTIA's BTOP Map."

[105] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[106] ASR Analytics, *Progress towards BTOP Goals: Interim Report on PCC and SBA Case Studies*.

[107] ASR Analytics, *Progress towards BTOP Goals: Interim Report on PCC and SBA Case Studies*.

ASR Analytics, *Progress towards BTOP Goals: Second Interim Report on CCI, PCC, and SBA Case Studies*.

ASR Analytics, *C.K. Blandin Foundation Sustainable Broadband Adoption Round 2*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, August 15, 2013).

ASR Analytics, *California Emerging Technology Fund Sustainable Broadband Adoption Round 2*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, September 09, 2013).

ASR Analytics, *Cambridge Housing Authority Public Computer Center Round 2*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, July 10, 2013).

ASR Analytics, *Connect Arkansas Sustainable Broadband Adoption Round 2*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, September 10, 2013).

ASR Analytics, *Delaware Department of State (Delaware Division of Libraries) Public Computer Center Round 2*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, December 19, 2013).

ASR Analytics, *Florida Agricultural and Mechanical University Public Computer Center Round 2*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, September 13, 2013).

ASR Analytics, *Foundation for California Community Colleges Sustainable Broadband Adoption Round 2*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, September 12, 2013).

ASR Analytics, *Future Generations Sustainable Broadband Adoption Round 2*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, August 08, 2013).

ASR Analytics, *Las Vegas-Clark County Urban League Public Computer Center Round 2*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, August 01, 2013).

ASR Analytics, *Michigan State University Public Computer Center Round 2*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, August 29, 2013).

ASR Analytics, *Smart Chicago Sustainable Broadband Adoption Round 2*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, July 10, 2013).

ASR Analytics, *South Carolina State Board for Technical and Comprehensive Education Public Computer Center Round 2*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, September 11, 2013).

ASR Analytics, *Technology For All, Inc. Public Computer Center Round 2*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, September 16, 2013).

ASR Analytics, *Urban Affairs Coalition Sustainable Broadband Adoption Round 2*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, September 10, 2013).

ASR Analytics, *WorkForce West Virginia Public Computer Center Round 2*, *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)* (Potomac, MD, September 23, 2013).

[108] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[109] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[110] National Telecommunications and Information Administration, "NTIA's BTOP Map."

[111] National Telecommunications and Information Administration, "NTIA's BTOP Map."

National Telecommunications and Information Administration, "Post-Award Monitoring Database."

Samantha Becker et al., *Opportunity for All: How the American Public Benefits from Internet Access at U.S. Libraries* (Washington, DC: Institute of Museum and Library Services, March 2010), http://www.gatesfoundation.org/learning/Pages/us-libraries-report-opportunity-for-all.aspx.

[112] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[113] The following types of locations are listed in CAC, but are not included: Unreported, Community College, Institution of Higher Education, For-Profit Organization, Government Facility, Medical or Healthcare Provider, Public Safety, and School (K-12).

National Telecommunications and Information Administration, "NTIA's BTOP Map."

[114] Becker et al., *Opportunity for All: How the American Public Benefits from Internet Access at U.S. Libraries*.

[115] Becker et al., *Opportunity for All: How the American Public Benefits from Internet Access at U.S. Libraries*, 37.

[116] Federal Communications Commission, "Census Tract Information Mapped for Internet Access Services Faster than 200 Kbps in at Least One Direction."

National Telecommunications and Information Administration and Federal Communications Commission, "State Broadband Initiative."

[117] National Telecommunications and Information Administration, "NTIA's BTOP Map."

[118] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[119] There was no online PPR for Connected Nation, Inc., Public Adoption through Libraries (OPAL II): 2013-Q4. The entry most consistent with previous values was used.

[120] National Telecommunications and Information Administration and Federal Communications Commission, "State Broadband Initiative."

[121] United States Bureau of Labor Statistics, "Local Area Unemployment Statistics."

[122] ASR Analytics, *BTOP Evaluation Study Design*.

[123] United States Census Bureau, "2006-2010 ACS 5-Year Summary File."

[124] National Telecommunications and Information Administration and Federal Communications Commission, "State Broadband Initiative."

[125] "Treatment and Potential Control Counties" contains additional tables, but they are not used in the "Matching" analysis and not referenced in this section

[126] United States Department of Agriculture, *Advancing Broadband*.

ProPublica, “Recovery Tracker Eye on the Stimulus.”

National Telecommunications and Information Administration, “NTIA’s BTOP Map.”

National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

United States Census Bureau, “2010 TIGER/Line® Shapefiles.”

National Telecommunications and Information Administration, “Application Search.”

National Telecommunications and Information Administration and Federal Communications Commission, “State Broadband Initiative.”

[127] Sekhon, “Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R.”

[128] Section 5.2 provides additional discussion on difference-in-differences.

[129] National Telecommunications and Information Administration, “NTIA’s BTOP Map.”

National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

United States Census Bureau, “2010 TIGER/Line® Shapefiles.”

National Telecommunications and Information Administration, “Application Search.”

National Telecommunications and Information Administration and Federal Communications Commission, “State Broadband Initiative.”

[130] In this document, a “key” is a variable used to identify observations in a “data.table” R table. ASR used a key, or a combination of keys, to identify unique observations in these tables. Tables are sorted in ascending order by their keys.

[131] Canty and Ripley, “Boot: Bootstrap R (S-Plus) Functions.”

Davison and Hinkley, *Bootstrap Methods and Their Applications*.

[132] National Telecommunications and Information Administration and Federal Communications Commission, “State Broadband Initiative.”

[133] Canty and Ripley, “Boot: Bootstrap R (S-Plus) Functions.”

Davison and Hinkley, *Bootstrap Methods and Their Applications*.

[134] Federal Communications Commission, “Census Tract Information Mapped for Internet Access Services Faster than 200 Kbps in at Least One Direction.”

National Telecommunications and Information Administration and Federal Communications Commission, “State Broadband Initiative.”

[135] National Telecommunications and Information Administration, “NTIA’s BTOP Map.”

National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

United States Census Bureau, “2010 TIGER/Line® Shapefiles.”

National Telecommunications and Information Administration, “Application Search.”

National Telecommunications and Information Administration and Federal Communications Commission, “State Broadband Initiative.”

[136] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[137] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[138] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[139] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[140] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[141] CCI grants do not report training activities.

[142] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[143] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[144] National Telecommunications and Information Administration, “NTIA’s BTOP Map.”

[145] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[146] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

United States Census Bureau, “2005-2009 ACS 5-Year Summary File.”

[147] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[148] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[149] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[150] For more information, visit http://www2.ntia.doc.gov

[151] The PPRs do not differentiate in-kind and cash match contributions. ASR treats all contributions as cash.

[152] Although IMPLAN uses the term “IMPLAN industry sector,” this report uses the term “industry sector” for consistency.

[153] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[154] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[155] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[156] United States Department of Agriculture, *Advancing Broadband*.

ProPublica, “Recovery Tracker Eye on the Stimulus.”

National Telecommunications and Information Administration, “NTIA’s BTOP Map.”

National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

United States Census Bureau, “2010 TIGER/Line® Shapefiles.”

National Telecommunications and Information Administration, “Application Search.”

National Telecommunications and Information Administration and Federal Communications Commission, “State Broadband Initiative.”

[157] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[158] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[159] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[160] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[161] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[162] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[163] National Telecommunications and Information Administration, “NTIA’s BTOP Map.”

[164] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[165] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[166] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[167] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

United States Census Bureau, “2005-2009 ACS 5-Year Summary File.”

[168] The coefficient is statistically significant at the 5 percent level.

Peter Kuhn and Hani Mansour, “Is Internet Job Search Still Ineffective?,” *Forthcoming in The Economic Journal* (December 2013): 29, doi:10.1111/ecoj.12119.

[169] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[170] United States Bureau of Labor Statistics, “Local Area Unemployment Statistics.”

United States Bureau of Labor Statistics, “Quarterly Census of Employment and Wages.”

United States Census Bureau, “2005-2009 ACS 5-Year Summary File.”

United States Department of Labor and United States Bureau of Labor Statistics, *Geographic Profile of Employment and Unemployment, 2009*.

[171] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

National Telecommunications and Information Administration, “NTIA’s BTOP Map.”

Becker et al., *Opportunity for All: How the American Public Benefits from Internet Access at U.S. Libraries*.

[172] Becker et al., *Opportunity for All: How the American Public Benefits from Internet Access at U.S. Libraries*.

[173] Becker et al., *Opportunity for All: How the American Public Benefits from Internet Access at U.S. Libraries*, 71.

[174] Becker et al., *Opportunity for All: How the American Public Benefits from Internet Access at U.S. Libraries*, A3.38–9.

[175] Becker et al., *Opportunity for All: How the American Public Benefits from Internet Access at U.S. Libraries*, A3.40.

[176] ASR Analytics, *Michigan State University Public Computer Center Round 2*, 12.

ASR Analytics, *Florida Agricultural and Mechanical University Public Computer Center Round 2*, 8.

[177] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[178] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

National Telecommunications and Information Administration, "NTIA's BTOP Map."

Becker et al., *Opportunity for All: How the American Public Benefits from Internet Access at U.S. Libraries*.

[179] Becker et al., *Opportunity for All: How the American Public Benefits from Internet Access at U.S. Libraries*, 71.

[180] The coefficient is statistically significant at the 5 percent level.

Kuhn and Mansour, "Is Internet Job Search Still Ineffective?," 29.

[181] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

[182] United States Bureau of Labor Statistics, "Local Area Unemployment Statistics."

United States Bureau of Labor Statistics, "Quarterly Census of Employment and Wages."

United States Census Bureau, "2005-2009 ACS 5-Year Summary File."

United States Department of Labor and United States Bureau of Labor Statistics, *Geographic Profile of Employment and Unemployment, 2009*.

[183] National Telecommunications and Information Administration, "Post-Award Monitoring Database."

United States Census Bureau, "2005-2009 ACS 5-Year Summary File."

[184] Becker et al., *Opportunity for All: How the American Public Benefits from Internet Access at U.S. Libraries*.

[185] Becker et al., *Opportunity for All: How the American Public Benefits from Internet Access at U.S. Libraries*, 71.

[186] Becker et al., *Opportunity for All: How the American Public Benefits from Internet Access at U.S. Libraries*, A3.38–9.

[187] Becker et al., *Opportunity for All: How the American Public Benefits from Internet Access at U.S. Libraries*, A3.40.

[188] ASR Analytics, *C.K. Blandin Foundation Sustainable Broadband Adoption Round 2*, 8.

ASR Analytics, *Foundation for California Community Colleges Sustainable Broadband Adoption Round 2*, 13.

ASR Analytics, *Connect Arkansas Sustainable Broadband Adoption Round 2*, 8.

[189] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

[190] Grimes, Ren, and Stevens (2009) estimated twelve specifications of the effect of connectivity on firm productivity. Estimated productivity increases due to broadband adoption ranged from 7.6 percent to 13 percent. All twelve estimates of the effect of broadband adoption were statistically significant at the 5 percent level.

Arthur Grimes, Cleo Ren, and Phillip Stevens, “The Need for Speed: Impacts of Internet Connectivity on Firm Productivity” (2009): 33, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1604247.

[191] National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

United States Census Bureau, “2005-2009 ACS 5-Year Summary File.”

[192] Quantria Strategies LLC, *Effective Federal Income Tax Rates Faced By Small Businesses in the United States* (Cheverly, MD: United States Small Business Administration Office of Advocacy, April 2009), http://www.sba.gov/advocacy/effective-federal-income-tax-rates-faced-small-businesses-united-states.

[193] National Telecommunications and Information Administration and Federal Communications Commission, “State Broadband Initiative.”

[194] Suffolk County, Massachusetts was removed from the service area because it contained no connected CAIs and is different from the listed counties in size and demographic makeup.

[195] All service area counties were determined by connected CAIs and BTOP-funded fiber routes.

[196] 1,269 counties fell within the service area of an awarded BTOP grant only, 183 fell within a BIP service area only, and 292 fell within both a BTOP and BIP service area, for a total of 1,744 unique counties.

[197] United States Department of Agriculture, *Advancing Broadband*.

ProPublica, “Recovery Tracker Eye on the Stimulus.”

National Telecommunications and Information Administration, “NTIA’s BTOP Map.”

National Telecommunications and Information Administration, “Post-Award Monitoring Database.”

United States Census Bureau, “2010 TIGER/Line® Shapefiles.”

National Telecommunications and Information Administration, “Application Search.”

National Telecommunications and Information Administration and Federal Communications Commission, “State Broadband Initiative.”

[198] National Telecommunications and Information Administration and Federal Communications Commission, “State Broadband Initiative.”

[199] National Telecommunications and Information Administration and Federal Communications Commission, “State Broadband Initiative.”

[200] National Telecommunications and Information Administration and Federal Communications Commission, “State Broadband Initiative.”

[201] Nina Czernich et al., “Broadband Infrastructure and Economic Growth,” *The Economic Journal* 121, no. 552 (May 12, 2011): 20, doi:10.1111/j.1468-0297.2011.02420.x.

[202] Czernich et al. (2011) estimated three specifications of the effect of broadband introduction on GDP. The estimated effects ranged from 2.7 percent to 3.9 percent. All three effects were statistically significant at the 1 percent level.

Czernich et al., “Broadband Infrastructure and Economic Growth,” 521.

[203] LECG Ltd., *Economic Impact of Broadband: An Empirical Study* (London, UK, February 22, 2009), http://www.connectivityscorecard.org/images/uploads/media/Report_BroadbandStudy_LECG_March6.pdf.

[204] National Telecommunications and Information Administration and Federal Communications Commission, “State Broadband Initiative.”

[205] This estimate is statistically significant at the 1 percent level.

Jed Kolko, *Broadband and Local Growth*, August 21, 2010, 38, http://ssrn.com/abstract=1680597.

[206] The coefficient is statistically significant at the 10 percent level, with a *p*-value of 5.2 percent.

Sharon E. Gillett et al., *Measuring Broadband’s Economic Impact*, *National Technical Assistance, Training, Research, and Evaluation Project #99-07-13829* (Washington, DC: United States Department of Commerce, February 28, 2006), 25, http://cfp.mit.edu/publications/CFP_Papers/Measuring_bb_econ_impact-final.pdf.

[207] United States Bureau of Labor Statistics, “Local Area Unemployment Statistics.”

United States Bureau of Labor Statistics, “Quarterly Census of Employment and Wages.”

[208] National Telecommunications and Information Administration and Federal Communications Commission, “State Broadband Initiative.”

[209] United States Bureau of Labor Statistics, “Local Area Unemployment Statistics.”

[210] The Allen Consulting Group (2010) estimated the average value of a broadband Internet connection to an Australian household to be A$148 per week. The Allen Consulting Group (2010) also states that United States Internet users spend about half the time online that Australian Internet users spend online. Therefore, ASR used half of the estimated benefit to Australians. According to the Australian Bureau of Statistics (2011), average weekly gross household income from wages and salaries in Australia was A$2,173 in 2009 to 2010. The percent of income value was computed as (148 × 0.5) ÷ 2173.

The Allen Consulting Group, *Quantifying the Possible Economic Gains of Getting More Australian Households Online*, November 2010, 25, http://www.acilallen.com.au/cms_files/acgquantifyingonline2010.pdf.

Australian Bureau of Statistics, *Household Income and Income Distribution, Australia, 2009-10* (Canberra, ACT, August 30, 2011), 26, http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/6523.0Main+Features12009-10.

[211] National Telecommunications and Information Administration and Federal Communications Commission, “State Broadband Initiative.”

United States Census Bureau, “2006-2010 ACS 5-Year Summary File.”

Glossary

Acronym	Definition
ACS	American Community Survey
APR	Annual Performance Progress Report
ASR	ASR Analytics, LLC
BEA	United States Bureau of Economic Analysis
BIP	Broadband Initiatives Program
BLS	United States Bureau of Labor Statistics
BTOP	Broadband Technology Opportunities Program
CAC	Connecting America's Communities
CAI	Community Anchor Institution
CBP	County Business Patterns
CCI	Comprehensive Community Infrastructure
COTR	Contracting Officer's Technical Representative
CSV	Comma Separated Value
DOL	United States Department of Labor
FCC	Federal Communications Commission
FIPS	Federal Information Processing Standards
FTE	Full-Time Equivalent
FTTH	Fiber to the Home
FTTP	Fiber to the Premises
GCP	Gross County Product
GDP	Gross Domestic Product
GSP	Gross State Product
IMPLAN	Impact Analysis for Planning
JRE	Java Runtime Environment
LAPI	Local Area Personal Income
LAUS	Local Area Unemployment Statistics
LSAD	Legal/Statistical Area Description
MAF	Master Address File
NAICS	North American Industry Classification System
NBM	National Broadband Map
NOFA	Notice of Funds Availability

Acronym	Definition
NTIA	National Telecommunications and Information Administration
OECD	Organisation for Economic Co-operation and Development
PCC	Public Computer Centers
PNG	Portable Network Graphics
POP	Point of Presence
PPR	Quarterly Performance Progress Report
QCEW	Quarterly Census of Employment and Wages
RDA	R Data File
SBA	Sustainable Broadband Adoption
SBP	State Business Patterns
SCTCS	South Carolina Technical College System
SOW	Statement of Work
TIGER	Topologically Integrated Geographic Encoding and Referencing
USD	United States Dollar
USDA	United States Department of Agriculture
ZCTA	ZIP Code Tabulation Area

Bibliography

Allen Consulting Group, The. *Quantifying the Possible Economic Gains of Getting More Australian Households Online*, November 2010. http://www.acilallen.com.au/cms_files/acgquantifyingonline2010.pdf.

ASR Analytics. *BTOP Evaluation Study Design*. Potomac, MD, January 30, 2012. http://www.ntia.doc.gov/files/ntia/publications/study_design_order_number_d10pd18645_-_revised_2012-01-30.pdf.

———. *C.K. Blandin Foundation Sustainable Broadband Adoption Round 2. Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, August 15, 2013.

———. "C.K. Blandin Foundation: Minnesota Intelligent Rural Communities Sustainable Broadband Adoption (SBA)." *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, June 18, 2012.

———. *California Emerging Technology Fund Sustainable Broadband Adoption Round 2. Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, September 09, 2013.

———. "California Emerging Technology Fund: Broadband Awareness and Adoption Sustainable Broadband Adoption (SBA)." *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, June 18, 2012.

———. "Cambridge Housing Authority Community Computer Centers Public Computer Centers (PCC)." *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, June 18, 2012.

———. *Cambridge Housing Authority Public Computer Center Round 2. Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, July 10, 2013.

———. "City of Chicago: Smart Chicago Sustainable Broadband Adoption (SBA)." *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, June 18, 2012.

———. *Clearwave Communications Comprehensive Community Infrastructure. Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, December 03, 2013.

———. "Connect Arkansas Expanding Broadband Use in Arkansas through Education Sustainable Broadband Adoption (SBA)." *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, June 18, 2012.

———. *Connect Arkansas Sustainable Broadband Adoption Round 2. Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, September 10, 2013.

———. *Delaware Department of State (Delaware Division of Libraries) Public Computer Center Round 2. Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, December 19, 2013.

———. "Delaware Division of Libraries: Job/Learning Labs Public Computer Center (PCC)." *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, June 18, 2012.

———. *Executive Office of the State of West Virginia Comprehensive Community Infrastructure. Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, March 25, 2014.

———. "Florida A&M University: The FAMU Center for Public Computing and Workforce Development Public Computer Center (PCC)." *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, June 18, 2012.

———. *Florida Agricultural and Mechanical University Public Computer Center Round 2. Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, September 13, 2013.

———. *Foundation for California Community Colleges Sustainable Broadband Adoption Round 2. Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, September 12, 2013.

———. "Foundation for California Community Colleges: California Connects Sustainable Broadband Adoption (SBA)." *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, June 18, 2012.

———. "Future Generations Graduate School: Equipping West Virginia's Fire and Rescue Squads with Technology and Training to Serve Communities Sustainable Broadband Adoption (SBA)." *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, June 18, 2012.

———. *Future Generations Sustainable Broadband Adoption Round 2. Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, August 08, 2013.

———. *Lane Council of Governments Comprehensive Community Infrastructure. Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, April 04, 2014.

———. "Las Vegas Clark County Urban League: Access to Computer Technology and Instruction in Online Networking (ACTION) Public Computer Center (PCC)." *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, June 18, 2012.

———. *Las Vegas-Clark County Urban League Public Computer Center Round 2. Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, August 01, 2013.

———. *Massachusetts Technology Park Comprehensive Community Infrastructure. Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, March 19, 2014.

———. *MCNC Comprehensive Community Infrastructure*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, March 04, 2014.

———. *Merit Network, Inc. Comprehensive Community Infrastructure*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, March 04, 2014.

———. *Michigan State University Public Computer Center Round 2*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, August 29, 2013.

———. "Michigan State University: Evidence Based Computer Centers II Public Computer Centers (PCC)." *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, June 18, 2012.

———. *Mid-Atlantic Broadband Cooperative Comprehensive Community Infrastructure*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, February 06, 2014.

———. *OneCommunity Comprehensive Community Infrastructure*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, April 21, 2014.

———. *OSHEAN, Inc. Comprehensive Community Infrastructure*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, March 10, 2014.

———. *Progress towards BTOP Goals: Interim Report on PCC and SBA Case Studies*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, October 15, 2012. http://www.ntia.doc.gov/report/2012/progress-towards-btop-goals-interim-report-pcc-and-sba-case-studies.

———. *Progress towards BTOP Goals: Second Interim Report on CCI, PCC, and SBA Case Studies*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, 2014.

———. "SC State Board for Technical and Comprehensive Education: Reach for Success Public Computer Centers (PCC)." *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, June 18, 2012.

———. *Short-Term Economic Impacts Report*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, September 2013.

———. *Smart Chicago Sustainable Broadband Adoption Round 2*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, July 10, 2013.

———. *South Carolina State Board for Technical and Comprehensive Education Public Computer Center Round 2*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, September 11, 2013.

———. *South Dakota Network Comprehensive Community Infrastructure*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, February 05, 2014.

———. *Technology For All, Inc. Public Computer Center Round 2*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, September 16, 2013.

———. "Technology For All, Inc.: Texas Connects Coalition Public Computer Center (PCC)." *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, June 18, 2012.

———. *University of Arkansas System Comprehensive Community Infrastructure*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, March 11, 2014.

———. *Urban Affairs Coalition Sustainable Broadband Adoption Round 2*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, September 10, 2013.

———. "Urban Affairs Coalition: Freedom Rings Sustainable Broadband Adoption (SBA)." *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, June 18, 2012.

———. *WorkForce West Virginia Public Computer Center Round 2*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, September 23, 2013.

———. "WorkForce West Virginia: One-Stop Public Computer Center Modernization Public Computer Center (PCC)." *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, June 18, 2012.

———. *Zayo Group Comprehensive Community Infrastructure*. *Broadband Technology Opportunities Program Evaluation Study (Order Number D10PD18645)*. Potomac, MD, February 06, 2014.

Australian Bureau of Statistics. *Household Income and Income Distribution, Australia, 2009-10*. Canberra, ACT, August 30, 2011. http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/6523.0Main+Features12009-10.

Barreca, John D., J. Matthew Fannin, and Joshua D. Detre. *Estimating GDP at the Parish (County) Level: An Evaluation of Alternative Approaches*. Research Bulletin, January 24, 2012. https://www.lsuagcenter.com/en/communications/publications/Publications+Catalog/research/Economics/Estimating-GDP-at-the-Parish-County-Level-An-Evaluation-of-Alternative-Approaches.htm.

Bauer, Paul W., and Yoonsoo Lee. *Estimating GSP and Labor Productivity By State*. Policy Discussion Paper. Cleveland, OH, March 2006. http://www.clevelandfed.org/research/policydis/pdpno16.pdf.

Becker, Samantha, Michael D. Crandall, Karen E. Fisher, Bo Kinney, Carol Landry, and Anita Rocha. *Opportunity for All: How the American Public Benefits from Internet Access at U.S.*

Libraries. Washington, DC: Institute of Museum and Library Services, March 2010. http://www.gatesfoundation.org/learning/Pages/us-libraries-report-opportunity-for-all.aspx.

Canty, Angelo, and Brian D. Ripley. “Boot: Bootstrap R (S-Plus) Functions.” *R Package Version 1.3-9*, 2013. http://cran.r-project.org/web/packages/boot/index.html.

Clower, Terry L., and Bernard L. Weinstein. *Economic and Fiscal Impacts of the Corps of Engineers’ Trinity River Vision Project in Tarrant County Texas*, 2005. http://www.trinityrivervision.org/docs/PDFDocuments/pr_05impacts.pdf.

Czernich, Nina, Oliver Falck, Tobias Kretschmer, and Ludger Woessmann. “Broadband Infrastructure and Economic Growth.” *The Economic Journal* 121, no. 552 (May 12, 2011): 505–532. doi:10.1111/j.1468-0297.2011.02420.x.

Davison, A.C., and D. V. Hinkley. *Bootstrap Methods and Their Applications*. Cambridge, UK: Cambridge University Press, 1997.

Dowle, Matt, T. Short, S. Lianoglou, R. Saporta, A. Srinivasan, and E. Antonyan. “Data.table: Extension of Data.frame.” *R Package Version 1.9.2*, 2014. http://cran.r-project.org/package=data.table.

Dragulescu, Adrian A. “Xlsx: Read, Write, Format Excel 2007 and Excel 97/2000/XP/2003 Files.” *R Package Version 0.5.5*, 2013. http://cran.r-project.org/package=xlsx.

———. “Xlsxjars: Package Required POI Jars for the Xlsx Package.” *R Package Version 0.6.0*, 2014. http://cran.r-project.org/package=xlsxjars.

Federal Communications Commission. “Census Tract Information Mapped for Internet Access Services Faster than 200 Kbps in at Least One Direction.” *Local Telephone Competition and Broadband Deployment*. Washington, DC, March 04, 2014. http://transition.fcc.gov/wcb/iatd/comp.html.

GeoLytics. “Estimates Professional,” July 10, 2014. http://geolytics.com/USCensus,Estimates-Professional,Products.asp.

Gillett, Sharon E., William H. Lehr, Carlos A. Osorio, and Marvin A. Sirbu. *Measuring Broadband’s Economic Impact*. *National Technical Assistance, Training, Research, and Evaluation Project #99-07-13829*. Washington, DC: United States Department of Commerce, February 28, 2006. http://cfp.mit.edu/publications/CFP_Papers/Measuring_bb_econ_impact-final.pdf.

Grimes, Arthur, Cleo Ren, and Phillip Stevens. “The Need for Speed: Impacts of Internet Connectivity on Firm Productivity” (2009). http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1604247.

Grolemund, Garrett, and Hadley Wickham. “Dates and Times Made Easy with Lubridate.” *Journal of Statistical Software* 40, no. 3 (2011): 1–25. http://www.jstatsoft.org/v40/i03/.

Harrell Jr., Frank E., and Charles Dupont. “Hmisc: Harrell Miscellaneous.” *R Package Version 3.14-4*, 2014. http://cran.r-project.org/package=Hmisc.

IMPLAN Group LLC. “Economic Impact Analysis,” July 10, 2013. http://implan.com/V4/Index.php.

Kolko, Jed. *Broadband and Local Growth*, August 21, 2010. http://ssrn.com/abstract=1680597.

Kuhn, Peter, and Hani Mansour. “Is Internet Job Search Still Ineffective?” *Forthcoming in The Economic Journal* (December 2013). doi:10.1111/ecoj.12119.

LECG Ltd. *Economic Impact of Broadband: An Empirical Study*. London, UK, February 22, 2009. http://www.connectivityscorecard.org/images/uploads/media/Report_BroadbandStudy_LECG_March6.pdf.

Murdoch, Duncan. “Tables: Formula-Driven Table Generation.” *R Package Version 0.7.64*, 2013. http://cran.r-project.org/package=tables.

National Telecommunications and Information Administration. “Application Search.” *Broadband USA*, 2014. http://www.ntia.doc.gov/legacy/broadbandgrants/applications/results.htm.

———. *Broadband Technology Opportunities Program (BTOP) Quarterly Program Status Report*, February 12, 2014. http://www.ntia.doc.gov/files/ntia/publications/ntia_btop_19th_quarterly_report.pdf.

———. “BTOP Evaluation Study.” *Program Reports*, 2013. http://www2.ntia.doc.gov/BTOP-Reports#evaluation.

———. “BTOP Map Fact Sheet.” *Connecting America’s Communities*. Washington, DC, March 2012. http://www2.ntia.doc.gov/BTOPmap/data/btop-map-fact-sheet.pdf.

———. “NTIA’s BTOP Map.” *Connecting America’s Communities*. Washington, DC, August 15, 2013. http://www2.ntia.doc.gov/BTOPmap/.

———. “Post-Award Monitoring Database,” April 11, 2014.

———. “Statement of Work for Broadband Technology Opportunities Program (BTOP) Evaluation Study,” July 26, 2010. http://www2.ntia.doc.gov/files/btop_sow.pdf.

National Telecommunications and Information Administration, and Federal Communications Commission. “State Broadband Initiative,” May 12, 2014.

ProPublica. “Recovery Tracker Eye on the Stimulus,” 2012. http://projects.propublica.org/recovery/gov_entities/12e2/list/4.

Quantria Strategies LLC. *Effective Federal Income Tax Rates Faced By Small Businesses in the United States*. Cheverly, MD: United States Small Business Administration Office of Advocacy, April 2009. http://www.sba.gov/advocacy/effective-federal-income-tax-rates-faced-small-businesses-united-states.

R Core Team. “R: A Language and Environment for Statistical Computing.” Vienna, Austria, 2013. http://www.r-project.org/.

Sarkar, Deepayan. *Lattice: Multivariate Data Visualization with R*. New York: Springer, 2008. http://lmdvr.r-forge.r-project.org.

Sekhon, Jasjeet S. “Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R.” *Journal of Statistical Software* 42, no. 7 (June 2011).

Therneau, Terry M. “A Package for Survival Analysis in S.” *R Package Version 2.37-7*, 2014. http://cran.r-project.org/package=survival.

Therneau, Terry M., and Patricia M. Grambsch. *Modeling Survival Data: Extending the Cox Model*. New York: Springer, 2000. 0-387-98784-3.

U.S. Department of Agriculture. *Estimating The Impact on Employment of USDA’s Programs in the American Recovery and Reinvestment Act (ARRA)*, 2013.

U.S. Department of Energy. *Economic Impact of Recovery Act Investments in the Smart Grid*, 2013. http://www.smartgrid.gov/sites/default/files/doc/files/Smart Grid Economic Impact Report.pdf.

United States Bureau of Economic Analysis. “Are Gross Domestic Product (GDP) Estimates Available at the County Level?” *Frequently Asked Questions*. Washington, DC, January 09, 2008. http://www.bea.gov/faq/index.cfm?faq_id=458.

———. “Gross Domestic Product by State.” *Regional Economic Accounts*. Washington, DC, April 01, 2014. http://www.bea.gov/regional/index.htm.

———. “Local Area Personal Income.” *Regional Economic Accounts*. Washington, DC, April 01, 2014. http://www.bea.gov/regional/index.htm.

United States Bureau of Labor Statistics. “Local Area Unemployment Statistics.” Washington, DC, April 01, 2014. http://www.bls.gov/lau/.

———. “Quarterly Census of Employment and Wages.” Washington, DC, April 01, 2014. http://www.bls.gov/cew/.

United States Census Bureau. “2005-2009 ACS 5-Year Summary File.” *American Community Survey*. Washington, DC, December 14, 2010. http://www.census.gov/acs/www/data_documentation/2009_release/.

———. “2006-2010 ACS 5-Year Summary File.” *American Community Survey*. Washington, DC, December 08, 2011. http://www.census.gov/acs/www/data_documentation/2010_release/.

———. “2010 TIGER/Line® Shapefiles.” Washington, DC, June 19, 2012. http://www.census.gov/geo/www/cob/cbf_counties.html.

———. “County Business Patterns.” Washington, DC, April 01, 2014. http://www.census.gov/econ/cbp/.

United States Department of Agriculture. *Advancing Broadband*. *Broadband Initiatives Program Awards Report*, January 2011. http://www.rurdev.usda.gov/supportdocuments/RBBreport_V5ForWeb.pdf.

United States Department of Labor. “Changes in Basic Minimum Wages in Non-Farm Employment Under State Law: Selected Years 1968 to 2013.” *Wage and Hour Division (WHD)*. Washington, DC, December 2013. http://www.dol.gov/whd/state/stateMinWageHis.htm.

United States Department of Labor, and United States Bureau of Labor Statistics. *Geographic Profile of Employment and Unemployment, 2009*. *Bulletin 2748*. Washington, DC, July 2010. http://www.bls.gov/opub/gp/pdf/gp09full.pdf.

———. *Geographic Profile of Employment and Unemployment, 2011*. *Bulletin 2774*. Washington, DC, September 2012. http://www.bls.gov/opub/gp/pdf/gp11full.pdf.

Urbanek, Simon. “rJava: Low-Level R to Java Interface.” *R Package Version 0.9-6*, 2013. http://cran.r-project.org/package=rJava.

Venables, W.N., and Brian D. Ripley. *Modern Applied Statistics with S*. 4th ed. New York: Springer, 2002. http://www.stats.ox.ac.uk/pub/MASS4.

Wickham, Hadley. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer, 2009. http://had.co.nz/ggplot2/book.

———. “Reshaping Data with the Reshape Package.” *Journal of Statistical Software* 21, no. 12 (2007): 1–20. http://www.jstatsoft.org/v21/i12/.

———. “Scales: Scale Functions for Graphics.” *R Package Version 0.2.3*, 2012. http://cran.r-project.org/package=scales.

———. “The Split-Apply-Combine Strategy for Data Analysis.” *Journal of Statistical Software* 40, no. 1 (2011): 1–29. http://www.jstatsoft.org/v40/i01/.

Zeileis, Achim, and Yves Croissant. “Extended Model Formulas in R: Multiple Parts and Multiple Responses.” *Journal of Statistical Software* 34, no. 1 (2010): 1–13. http://www.jstatsoft.org/v34/i01/.