Broadband Data Validation: Comparing U.S. Broadband Coverage

DSPG Team: Kat Savchyn (Georgetown University), Sarah McDonald (Marymount University), Raghav Sawhney (Virginia Tech)

SDAD Team: Teja Pristavec, Josh Goldstein, Stephanie Shipp

SPONSOR: John Pender, Peter Hahn, Zachary Keeler, USDA Economic Research Service

Introduction

The broad purpose of this project explores whether broadband development affects rural prosperity, as measured by changes in property values and quality of life indicators in rural communities. The first component of the project conducted a comprehensive literature review of value modeling consisting of factors other than broadband that shape property values. We found relevant environmental, neighborhood, structural, and locational attributes that influence property prices. The second component of this project explores methods of validating Federal Communications Commission (FCC) broadband data. Following the Consolidated Appropriations Act, the United States Department of Agriculture (USDA) created the ReConnect Program, which provided $600 million in loans and grants for broadband improvement in rural areas. The FCC provides data on broadband, FCC Form 477, which informs broadband funding policies, but has known limitations.

Data

FCC FORM 477

- Receives self-reported data from broadband providers, potentially introducing error if advertised coverage estimates are reported rather than actual estimates.
- If a census tract had just one subscriber, the entire census tract is assumed to have broadband.
- Claims that broadband is not available to 25 million Americans.

American Community Survey

- The survey provides more detailed socioeconomic information than the decennial Census form.
- Table B28002: Presence and Types of Internet Subscriptions in Household for the 2013-2017 Five-Year Summary provided data specific to broadband.
- Receives self-reported data from households but does not include group quarters.

BroadbandNow

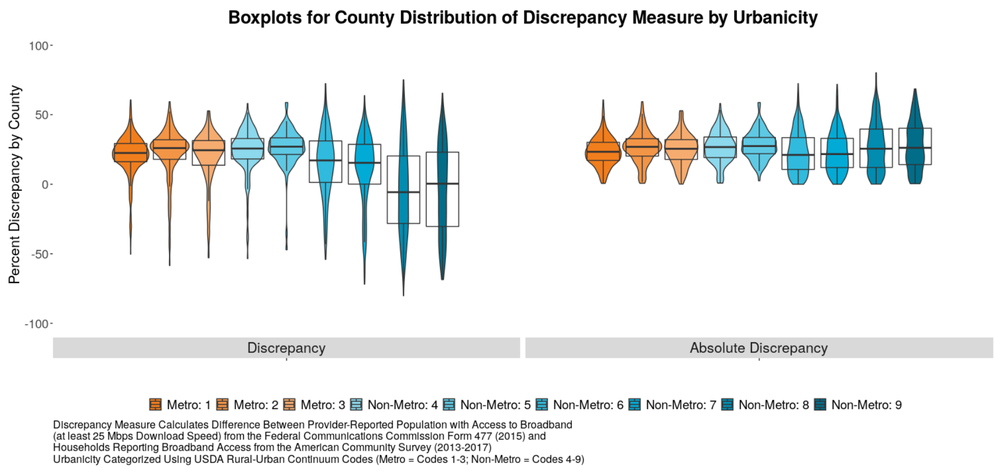
- Online search engine which lists information for broadband availability on the zip code and city level.
- Provides information regarding the number of providers, maximum and average upload and download speeds as well as the change in broadband speeds over time.

Microsoft Airband

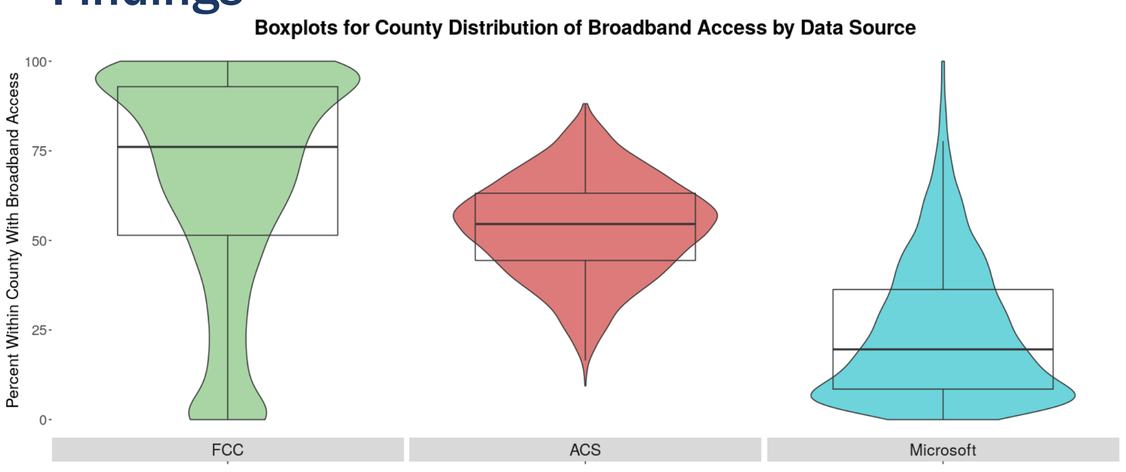
- Finds that around 163 million Americans do not use the internet at 25 Mbps, or broadband speed.
- Microsoft calculated this by analyzing its servers' logs when electronic devices downloaded Microsoft Windows and/or Office updates, as well as when devices used the Microsoft's Bing search engine and Xbox gaming consoles.

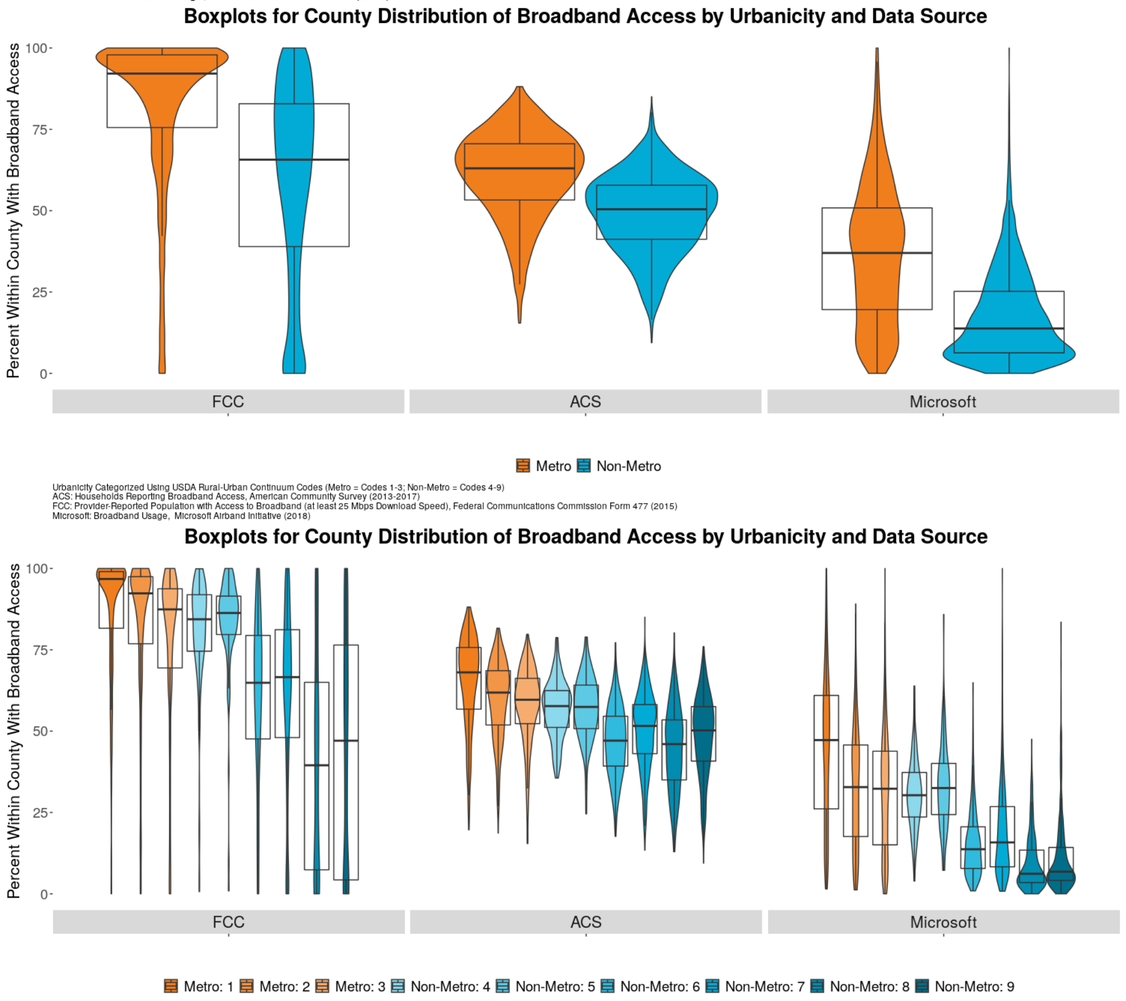
Validation Methodology

- Data was scraped and aggregated at the County, City, Census Tract, and Block Group levels in order to observe differences in coverage.
- FCC coverage metrics were developed utilizing a threshold of 25 Mbps download speed and supplemented with Decennial Population data to determine % of population with broadband access. The absolute difference between this metric and other data sources identifies level of discrepancy:
    1) ACS 007 – Access to Broadband (Excluding Cellular and Satellite)
    2) FCC Internet Service Connections of at least 10 Mbps (Bins) [Block Group, Census Tract]
    3) Microsoft Usage Percentage [County]
    4) BroadbandNow Coverage [City level]
- Data examined by USDA-specified Rural-Urban Continuum Codes (RUCC) – RUCC Codes distinguish metropolitan counties by the population size of their metro area, and nonmetropolitan counties by degree of urbanization and adjacency to a metro area.



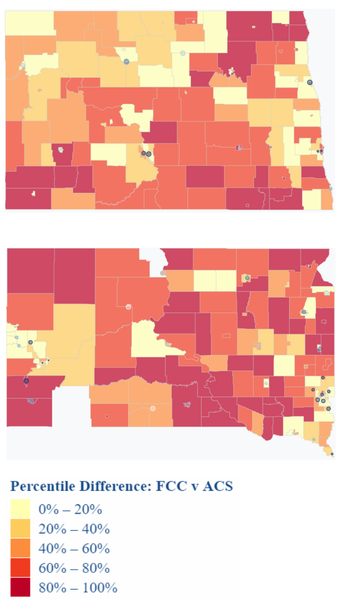
Findings





Case Study:

Broadband coverage discrepancy map for North Dakota (ND) and South Dakota (SD) at block group level



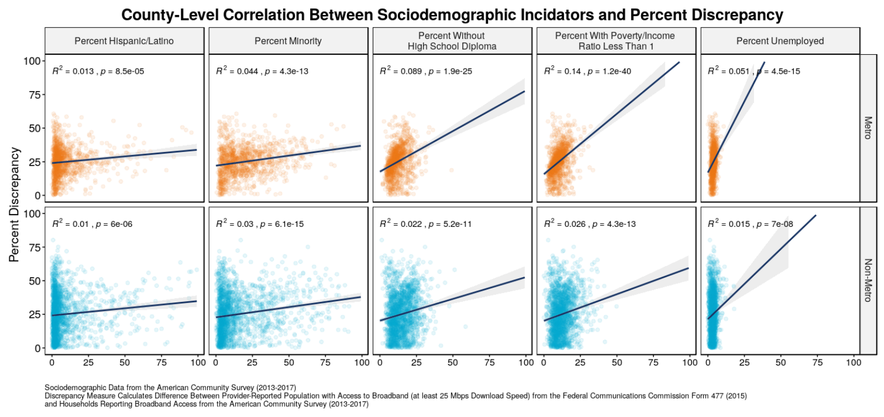
- Explored the broadband coverage discrepancy between the FCC and ACS data in North and South Dakota to understand the effect of state boundaries and policies on broadband availability.
- We notice higher coverage discrepancy in the more rural parts of ND and SD.
- The southwest section of ND is majorly urban compared to the rest of the state and we can clearly see a difference in broadband coverage across the northeast portion of SD which is purely rural.
- North Dakota is a largely rural state that has experienced a surge in broadband availability mainly due to the Broadband Technology Opportunities Program (BTOP) as well as a large-scale introduction of rural cooperatives (co-ops) that are made up of local service providers in the area[1].

Discussion & Further Research

We find that the FCC, ACS, and Microsoft Airband data give very inconsistent estimates of broadband accessibility. However, they all estimate lower coverage in rural areas compared to urban areas.

Next steps:

- Explore the relationship between percent discrepancy and sociodemographic factors for counties by urbanicity. We provide our preliminary exploration of discrepancy and sociodemographic factors correlations below. Sociodemographic factors are more likely to influence discrepancies in metro than non-metro areas.



- Use FCC, ACS, and Microsoft data to obtain model-based broadband access predictions based on sociodemographic information (e.g. urbanicity, unemployment).

UNIVERSITY of VIRGINIA
BIOCOMPLEXITY INSTITUTE
Data Science for the Public Good

http://bband.policy-analytics.net/