New York State Broadband Access Model Writeup

October, 2024

Introduction

This is a small project studying broadband availability in New York State. To complete this task, I utilize Jupyter notebook, alongside 2019 data from the Federal Communications Commission (FCC) on broadband deployment across the US, and 2015 US Census Bureau Cartographic Boundary files found here. [1] I first set up a Jupyter notebook and utilizing an API download both the New York State data and Cartographic Boundary data in a Json format. I then filter the data for completeness, and convert the results into a data frame to be used in the next stage of the analyses.

Question 1

Table 1 below, presents the mean and median number of broadband providers per Census block at the 15 digit FIPS block level. To get these results, I first drop duplicate provider ids for each block, and then utilize the mean, median and count functions in python to find the requisite data points. For each FIPs block, the unique number of providers (identified by 'provider-id') is calculated. From the resulting FIP-count pairs, the mean number of providers is 5.85 and the median number of providers is 6.0. Given that the mean number of providers is not an integer, there are 0 census blocks that have exactly the mean number of providers within them. There are 143,225 blocks with the median number of providers.

Table 1: Descriptive statistics for N.Y. broadband providers per U.S. Census block

At the 15-digit FIPs block level the:	Total
Mean number of providers:	5.85
Median number of providers:	6
Number of blocks with mean number of providers:	0
Number of blocks with median number of providers:	143,225

Note: Each observation is at the 15-digit FIPS level. As a result of the mean not being an integer, there is 0 number of blocks with exactly the mean number of providers.

[1] https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.2015.htmllist-tab-1556094155

Question 2

At an aggregate level, I find mixed evidence that the number of broadband providers is positively related to how fast the available internet is in New York State. However, this result is endogenous, with there being substantial variation in internet speed across different geographic locations in New York State (as seen in the choropleth maps in Figure 1). To attain such results, I ran two separate OLS regressions, using two different measures of internet speed, maximum advertised upload and maximum advertised download respectively. For clarity, note that download speed encompasses downloading material (i.e. a video), and upload speed encompasses uploading material (i.e. a file to GitHub).

Table 2 and 3 each show the results of the two OLS regressions with different dependent variables. Table 2 has the dependent variable variable as the average maximum advertised upload speed (avg maxadup), and includes two covariates, being, the number of unique internet service providers in each geographic area and the availability of fiber optic connections in the area. The average maximum advertised upload speed is 14.48 MBPS, and an additional provider is associated with -0.38 MBPS decrease in the variable of interest, controlling for the proportion of providers using optic fiber. If the proportion of optic fiber provider increase from 70% to 80%, the associated increase in advertised speed is 45.4 MBPS. This model had an R^2 of 0.491 meaning that just under half of the percentage of variance in the outcome is explained by the model. I see a more positive trend in Table 3's results, which replaces the dependent variable with the average maximum advertised download speed (avg maxaddown). The average maximum advertised download speed is 40.88 MBPS, and an additional provider increases the variable by 18.43 MBPS controlling for the proportion of providers using optic fiber. Suppose the proportion of providers using optic fiber increase from 70% to 80%, the average maximum advertised download speed is associated with an increase of 23.11 MBPS. This model had an R^2 of 0.255 meaning that just under one quarter of the percentage of variance in the outcome is explained by the model. Altogether, there are mixed evidence on relationship between number of providers and internet speed. While the results of Table 2 and 3, especially the positive relationship seen in Table 3, suggest that number of providers is a positive factor in internet speed, the technology adopted by the providers and population density are likely better determinants.

Table 2: OLS Regression results for maximum upload speed across New York State

Dependent Variable	avg_maxadup
Intercept	14.4804***
	(0.346)
Number of providers	-0.3835***
	(0.063)
Fiber optic supply	454.4220***
	(0.898)
Oberservations	350,169
R^2	0.491

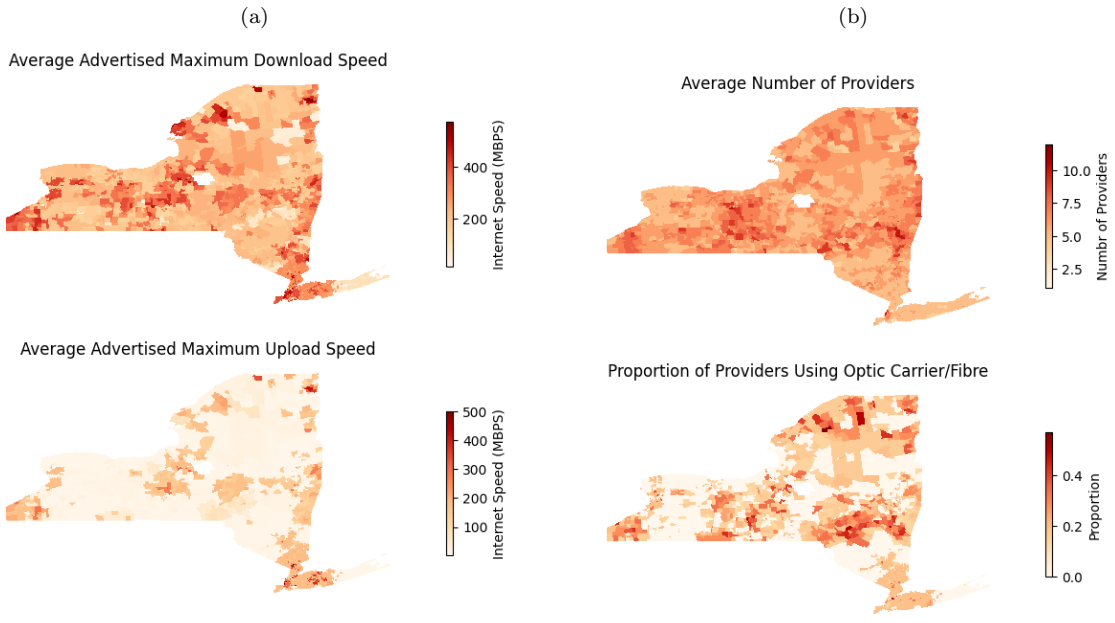
Note: OLS Regression results. Each observation is at the 15-digit FIPS level. Each observation is a region-industry-year. All variables are calculated using FCC data. Standard errors assume that the covariance matrix of the errors is correctly specified. **Statistical significance:** *** 0.01, ** 0.5, * 0.1.

Table 3: OLS Regression results for maximum download speed across New York State

Dependent Variable	avg_maxaddown
Intercept	40.8834***
	(0.552)
Number of providers	18.4345***
	(0.557)
Fiber optic supply	231.1196***
	(6.659)
Oberservations	350,169
R^2	0.255

Note: OLS Regression results. Each observation is at the 15-digit FIPS level. Each observation is a region-industry-year. All variables are calculated using FCC data. Standard errors assume that the covariance matrix of the errors is correctly specified. **Statistical significance:** *** 0.01, ** 0.5, * 0.1.

Figure 1: Choropleth maps of broadband services and coverage in New York State



Question 3

Examining variation in available internet speed by geography, I find that there is substantial variation in upload and download speeds across different locations in New York State, as indicated by the choropleth maps in Figure 1.[2] Specifically, I find that specific geographic areas, that have high numbers of internet providers, also have high advertised upload and download speeds. Specifically, the areas that have higher population densities, namely metropolitan areas, such as New York City, Albany, and Rochester, have higher numbers of providers servicing those regions, and higher upload and download speeds. More remote regions, especially those focused primarily on agriculture, such as in the finger lakes region, have a substantially lower number of providers servicing their areas and lower internet speeds overall. Similarly, as my previous regressions confirmed, the areas that have faster upload and download speeds are also where internet providers have rolled out fiber optic cable. Such a result is expected, as fiber optic cable is the considered the fastest commercially available means of wired connection, and thus, fiber optic coverage and broadband speed are expected to be positively associated with one another, thereby serving as good confirmation of our larger results. More broadly, this result may be indicative of a positive effect of competition in regions where there is a larger consumer market for broadband connection. Specifically, high density population locations where the internet market is bigger, such as in New York City, Albany, and Rochester, inevitably place more demand on broadband providers to provide better speeds, infrastructure, and services, thereby promoting competition amongst providers to offer better products to the population.

[2]Note: The choropleth map is at census block group level.