Selection Rule for Higher Cap Level

Vinod Bakthavachalam

November 26, 2014

The 295 city libraries identified as outliers based on size and user levels (values in the top 5% of the IMLS sample) represent over 30% of the total reported users of public wifi. That is 295 out of the 9024 libraries in the sample with valid data values (nonnegative reported user and size levels). To state this another way 3% of libraries represent almost a third of total users. Any economically efficient selection rule will send money to where it is most needed, so the optimal rule should send money to those libraries that serve a majority of the user base. Therefore, when designing a cap selection rule to identify which libraries should be eligible for the higher cap, it is important to ensure that these libraries are selected.

Natural expectations are that user levels should correlate highly with population levels. The inclusion of the outlier libraries under any population selection rule can be assessed using the LOCALE categories within the IMLS data, which assigns a value based on the group the library falls into: rural, suburban, town, or city. These groups then breakdown into subcategories. The relevant values for city libraries are as follows:

11 - inside a principal city with population of 250,000 or more.

12 - inside a principal city with population less than 250,000 and greater than or equal to 100,000.

13 - inside a principal city with population less than 100,000.

For a selection rule based on population levels above 250,000 and IMLS city classifications (LOCALE codes 11, 12, and 13) only 71 of the 295 outliers or 24% of the outliers would be included. The number of outliers that would be included in this selection rule can be examined by looking at the breakdown of LOCALE codes for the 295 city libraries:

Table 1: Breakdown of Outlier City Libraries

LOCALE Type:	11	12	13
Library Freq:	71	107	117

Relaxation of the population cutoff to lower levels also does not solve this issue. Using a cutoff of 100,000, so that libraries in cities with a population over 100,000 will be eligible for the increased cap, leads to inclusion of 178 of the 295 outliers or 60%. A large amount of libraries would be missed under this rule, but using a selection rule of over 100,000 does capture 27% out of the 30% of all users served by these outlier city libraries. The majority of the reported users are in cities with reported user levels above 100,000.

Turning to libraries not classified as within principal cities as defined by the IMLS data (suburb, town, and rural libraries), there are an additional 341 outliers with high user levels per square foot. If a selection rule of libraries includes areas outside of IMLS city classifications with a population over 250,000 (particularly in IMLS classification 21 - Suburb, outside a principal city and inside an urbanized area with a population of 250,000 or more), then the rule will yield eligibility for an additional 122 outlier libraries serving an additional 10% of the reported wifi user base. In addition, this formula used in conjunction with cities over a population of 250,000 can capture 300 of the total 636 outliers across all cities, suburbs, towns, and rural libraries and represents a large portion of the user base.

A much more natural selection rule is based on user levels, which was used in the identification of the outliers in the first place. Using user levels instead of population levels aligns the selection rule to the outcome of interest, which is precisely to provide increased funding for public wifi service to those libraries that need it, which should be those libraries who serve the most users. Including all libraries with a reported user level in the top 5% of the data will ensure all outlier libraries are included. This ensures that additional libraries with heavy usage of wifi which are not necessarily in large cities can also qualify for needed funding, ensuring that the top 7% of libraries which serve over 60% of all public wifi users receive the requisite funding. Based on the data this would lead to including libraries with user levels above 126,234.8. This selection rule is the most equitable one as it restricts the higher cap to those who likely need it the most and does not bias towards a

certain type of library or location, including libraries that reside in town and rural locations.

An argument could be made that a selection rule based on user levels does not lead to the efficient outcome because user levels do not necessarily correlate with the cost of providing the service. For example, a rural location with less users than a city library could incur much higher costs of providing service solely due to library location. Another criteria for efficiency could therefore be sending money to libraries where costs of providing the service are higher. This rule has a couple shortcomings. First off, there is no good measure of cost of providing wifi in the data. There is a variable that includes the cost of providing service, but it also includes other operating expenses that could skew the inclusion group. Libraries could have high expenses even with low cost of providing wifi due solely to other operating costs being higher. Second, a library could incur high costs of providing wifi due to inefficiency and not because it serves a large number of people. It does not make sense to give extra funding to libraries that do not have many users. The funding should target areas where user levels are high to help the most people possible.

Overall, the best way to identify libraries eligible for the increased cap is through reported user levels. This is the most direct way. Selection rules based on other criteria like population level or cost bias towards one type of location. A population selection rule with a cutoff across the large city and suburb IMLS categories can capture a large percentage of wifi users, but it does not reach all libraries who need the funding and may miss libraries who are not outliers and otherwise eligible for additional funding. User levels is a more direct measure of wifi demand and so allocating funds to libraries with high demand will maximize the impact of this policy.