Phillip Hicks  
April 9, 2020

My third paper will look into the relationship between transportation (Uber and taxi rides) and the real estate value of New York City neighborhoods. The main thesis is that people are quicker to move than prices, so increases in activity will preempt increase in valuation. Property values will be sourced from New York City tax valuations and will not include type-1 properties as they cannot increase more than 6% in value in one year or 20% over five. Uber and taxi data will be taken from the NYC OpenData website and will be used to find both the starting point and destination. Macroeconomic variables based on the past paper will also be included. These will account for overall market changes and general growth. Lastly, population variables for the pickup location (wealth and age) may also be considered if cleaning the taxi data is completed fast enough. The theory here is that younger or more wealthy people might have a greater impact on determining the next “high growth” neighborhood, so rides originating from specific demographic areas may be more informative.

The biggest challenge of this project is data preparation. Work was already done aggregating property valuations into means based on neighborhoods, however the taxi data is far bigger and more difficult to work with than the property data. Six month periods can have nearly 100 million rows with 19 columns before being modified, so even downloading and importing csv files takes time. However, once the unless columns are removed and pick/drop-off coordinates are changed to neighborhoods, the process of regressing values onto the covariates should be fairly easy.

While *Forecasting Real Housing Price Growth in the Eighth District States* (2007) by David E. Rapach and Jack K. Strauss was used to select macroeconomic variables in the last paper, no literature could be found relating transportation to shifts in neighborhood values. In fact, little work could be found addressing changes in neighborhood valuations in general. Instead, most work either focuses on cities as a whole or specific properties.

The main outcome of this project will be variable significance and overall prediction accuracy from the initial regression. Depending on time remaining, further work could be done using more advanced models for increase prediction accuracy.