The Effect of New Housing Projects on Expenditures in New York Municipalities

Andrew Bates

October 11, 2018

Executive Summary

In this paper we estimate future expenditures for two municipalites in New York based on various projected demographic and income-related factors. These estimates are obtained from a linear model chosen via a stepwise regression procedure with AIC as the model selection criteria. The variables in the model are wealth per person, population, percent intergovernmental funding, and growth rate. The estimated expenditures for Warwick are 1, 2 for the years 2005 and 2025. The estimated expenditures for Monroe are 1, 2 for the years 2005 and 2025.

1 Introduction

Two New York towns, Warwick and Monroe, would like to estimate future expenditures triggered by new housing construction proposals. They are primarily interested in determining whether they need to increase funds to compensate for increased expenditures related to the housing projects. To construct these estimates, Warwick and Monroe obtained data on expenditures along with various demographic and income-related variables from several New York municiplaties. The data used in this study consisted of 916 observations of seven variables. The response variable is expenditure per person revenue from state and federal grants, population density, mean income per person, and growth rate.

2 Analysis

2.1 Exploratory Data Analysis

The *insert reason here. e.g. skewness* of growth rate suggested a log transformation. However, growth rate contains some negative values as well as some zero values. For this reason, we used the following modified log transformation:

$$p\log(\text{growth rate}) = \begin{cases} \log(growthrate + 0.15) & \text{if growth rate} > 0 \\ -\log(-growthrate + 0.15) & \text{if growth rate} < 0. \end{cases}$$

Note that since this is a one-to-one transformation ...

In this dataset there were two measures of the size of a municipality: population and population density. Unsurprisingly, the correlation between these two measures was relatively high (0.67). In the interest of parsimony and to mitigate possible colinearity issues, we decided to consider only one of these variables to construct our model. Population density had a moderate correlation with mean income per person (0.49) whereas population had a relatively low correlation with income (0.29). To hedge against problems with colinearity, we decided to consider population for the model building process.

2.2 Modeling and Diagnostics

3 Conclusion