Honors Thesis Draft Prospectus

Griffin

November 23, 2023

Abstract

In 1996, Taipei joined other major East Asian cities with the development of its own Mass Rapid Transit (MRT) system. The development of this system has many implications concerning the expenditure and income of the average Taiwanese citizen. Adding transportation data that includes the number of stations in a district in Taipei at a given time, to the Family Income Survey from the years 1991 - 2006, I find statistically significant increases in measures of rent expenditure in districts that are connected to more MRT stations. The model for this question includes time and location fixed effects, controlling for age, income, education, and other characteristics of the household head. More exploration will be done using a regression discontinuity design to estimate the impact of elderly assistance programs on expenditure, which may also tease out a measure of elasticity for transportation in Taiwan during this time period.

1 Introduction

Taiwan underwent massive structural changes to its economy and political system in the second half of the 20th century. On the back of these major reformations, Taiwan experienced a period of rapid growth known as the Taiwan Miracle. This resulted in some growing pains that public transportation infrastructure had to catch up to. Prior to 1996, there the main public transportation method was by bus; because of perilous traffic issues and congestion, the Taiwanese government began planning for the development of its own Mass Rapid Transit System in Taipei. Since its opening in 1996, it has received regular and substantial expansions, up to the present day.

Today, the MRT is an integral part of Taipei's public transportation infrastructure, with connections to neighboring New Taipei City, and to the Taoyuan International Airport. While this represents an immense transformation, it also implies changes in the expenditure of Taiwanese citizens, and potentially the cost of living.

2 Literature Review

There is a substantial body of literature regarding the impacts of public transportation infrastructure on pricing and inequality.

3 Data and limitations

In this exploration, I'm using the Family Income Survey. Conducted annually, it collects information at the household level and individuals in said households. It includes information on income, education, marital status, employment status, and expenditures on different items. The limitation of this data set is its geographic granularity. Using the survey from 1991 to 2006 is motivated by two main factors. The first being that the MRT station opened in 1996, starting in 1991 provides 5 years of pre-period data. The reason for stopping at 2006 is due to a change in the survey at this time, namely, the abstraction of geographic data. Prior to 2007, the survey collected information on the specific district a household resides in, while in 2007 and onward, data is only collected at the county or city level.
In terms of the transportation data, which is not included in the original data set, I collected information on the number of stations that exist in each district in each year and added this to the survey data. The number of stations in a district serves as a continuous treatment variable, and the main explanatory variable I'm interested in observing.
The main limitation of this data set is the lack of geographic granularity, as I'm observing 12 districts across a period of 15 years; this leads to large standard errors, many of the relationships demonstrated in section 4 straddle 0 in the 95 percent confidence interval.

4 Methodology and Findings

The main methodology and model I've utilized thus fair is an ordinary least squares regression with time and location fixed effects. I control for various household head characteristics such as income, education, marital status, occupation, and employment status. I'm treating the number of stations that exist within the same district at a given time as a household as the main explanatory variable. My current specification is as follows:

$$Outcome_{i,t} = \alpha + \lambda Stations_{i,t} + \beta_1 Area_i + \beta_2 Time_t + \beta_3 X + \epsilon$$

Where *Stations* is the number of stations in the district the household resides in at a current time, $Area_i$ is the district, $Time_t$ is the year, and X is a matrix of control variables for the household head characteristics mentioned above.
I found statistically significant decreases in expenditures on motor vehicle pre-

miums, increases in rent and water, increases in imputed rent or self owned, and decreases in gas charge.

5 Further Exploration, Regression Discontinuity

Taiwan, like the United States, has elderly assistance programs. These programs include assistance on healthcare spending, daycare, a national pension system, and a discount program of 50 percent off on transportation fares. All of these are of interest, but the last one is the most relevant to my question. The next way I'm going to explore this data set is through a regression discontinuity design, comparing those just above and below the age cutoff to receive these benefits.
This exploitation would allow me to see a few things that I believe are interesting and relevant. Firstly, it would allow me to see how many seniors in Taiwan take advantage of these programs. Secondly, it may allow me to estimate a measure of elasticity of this public good in Taipei. The mathematical representation of idea for this is as follows:

Average expenditure at 64 years old: $S = E(P * N)$, where S is spending, P is the price of a bus or MRT ride, and N is the amount of times a person utilizes public transportation at this price.

Turning 65, their spending equation changes: $S^* = E(P^* * N)$, where $P^* = \frac{P}{2}$ and represents the price that seniors face. This is the average expenditure right after crossing the threshold. Interpretation of for this can be shown through this example: what does it mean when these two spending equations equal each other?

$S = S^*$

The only way that this is possible is if the number of rides increases two-fold to offset the discount in the price by half. Similarly, if S^* is exactly half of S, this implies that the individual has no change to their utilization of public transportation when their price is decreases by half. This may allow me to estimate the elasticity of demand of public transportation In Taipei during this time period. Where:

$\epsilon_i = \frac{\Delta\%Q_i}{\Delta\%P_i}$

As mentioned above, the cutoff for the regression design is when the individual turns 65 years old. So, the metric model for this approach is going to be as follows:
$Y = \alpha + \beta\tau + \delta X + \epsilon$

Where τ is is the treatment effect of going across the threshold. It will be important to keep in mind external validity concerns, as this is only calculating the Local Average Treatment Effect.