Headway's Effect on Rail Transit Ridership in the U.S. and Its Policy Implications

A Thesis
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Master of Public Policy
in Public Policy

By

Alva Dennis Lytton, B.A.

Washington, DC
April 10, 2015
# Table of Contents

Introduction........................................................................................................................................ 1

Background......................................................................................................................................... 2

Literature Review............................................................................................................................... 3

American Public Transportation Association (APTA) Operational Data................................. 5

Descriptive Statistics and Graphs From the APTA Dataset .......................................................... 9

Multivariate OLS Regression of Ridership on Headway .............................................................. 17

Natural Experiment Via Headway Change at One Transit Agency ............................................. 22

Suggestions for Further Research Along These Lines................................................................. 28

Policy Implications for Existing Systems — How Often Should We Run?......................... 29

Policy Implications for New Systems — What Should Be Built?............................................. 32

Conclusions...................................................................................................................................... 36

Bibliography ...................................................................................................................................... 39
LIST OF FIGURES

Figure A. July 2014 Ridership of 31 Leading U.S. Rail Transit Systems ............... 11

Figure B. July 2014 Ridership of 30 Leading U.S. Rail Transit Systems without New York City.......................................................... 11

Figure C. July 2014 Ridership of 31 Leading U.S. Rail Transit Systems Divided by Track Mile.................................................................................................. 12

Figure D. Ridership Frequency Distribution (without New York)...................... 13

Figure E. Ridership Per Track Mile Frequency Distribution ............................ 13

Figure F. Minimum Headway Frequency Distribution ..................................... 13

Figure G. Scatterplot of Minimum Headway and Total Ridership for 2013 of 74 Rail Transit Systems .............................................................................. 14

Figure H. Scatterplot of the Natural Log of Total Passengers and Minimum Headway for 2013 of 74 Rail Transit Systems ................................. 15

Figure I. Scatterplot of the Natural Log of Total Passengers and Minimum Headway for 2013 of 74 Rail Transit Systems ................................. 16

Figure J. Hypothesized Relationship Between Headway and Ridership ............ 17

Figure K. Map of the Los Angeles Metro Rail system ...................................... 23

Figure L. Los Angeles Metro Rail advertising for headway reductions ............. 32
LIST OF TABLES

Table 1. Ridership and Other Select Operational Metrics for 31 U.S. Rail Transit Systems for June 2014 ......................................................... 9

Table 2. Basic Descriptive Statistics for the 31 Rail Transit Systems Discussed in Table A ................................................................. 10

Table 3. OLS Regression Table ................................................................................. 19

Table 4. Headway Changes Made to Los Angeles Metro Rail from 2011 to 2013 .... 24

Table 5. Results of Difference-in-Difference Analysis of Los Angeles Metro Rail Headway Changes ........................................................................... 25

Table 6. Results of DiD Analysis With a Comparison to Predicted Values From the OLS Regression Models ............................................................. 26

Table 7. Summary of Revenues and Expenses as a Result of Increasing Headway and author's own work ................................................................. 30

Table 8. Cost Drivers of Building Rail Transit Modes, adapted with author's observations and (Vuchic, 2005) ...................................................................... 33

Table 8. (con't) Notes From Previous Table ............................................................. 34

Table 9. Deficiencies and Possible Remedies in Empirical Approaches ............... 37
INTRODUCTION

A very basic experience of taking a train, any kind of train anywhere in the world, is waiting for it to pick you up. You might look at the timetable and do most of the waiting at home and not at the station\(^1\). You might come to the station just as a train is showing up and do little waiting. Or you might curse as your train is leaving. That the train is often not leaving exactly when you safely make it to the platform is a fundamental experiential difference between the train and the private automobile. That the train usually travels slower than a private automobile\(^2\), sometimes at the same speed, and sometimes faster is another important experiential difference. That individuals try to maximize their preferences for transportation as for other goods is a fundamental way to start an analysis that even the most anti-economics minded observers will agree to. So how much of a statistically significant positive effect, if any, does increasing train frequency (decreasing headway) have on ridership? Is it cost effective to increase train frequency on new or existing train systems?

Americans use trains and other public transportation at a lower rate than many in the developed world (Greendex 2014: Consumer Choice and the Environment – A Worldwide Tracking Survey, 2014, p. 114). However, there are still many different rail transit systems in this country with widely varying ridership, operational characteristics, and demographic criteria. Fortunately, data exists on key factors regarding the operational and demographic characteristics that make a sound econometric and policy analysis possible.

---

\(^1\) With the rise of smart phone based next bus or train arrival technology, not having to wait at the platform is often possible now.

\(^2\) Accounting for waiting for the train and its stops, as I'll discussing later in regards to *average speed.*
BACKGROUND

The train, a series of cars on a fixed guide-way, had existed for centuries since Roman times — mostly to haul materials out of mines. But a revolution occurred in Britain in the first years of the 19th century when the steam locomotive was added to propel a train. Also, advances in metallurgy allowed steel wheels to contact steel rails by about 1820. The steam locomotive — expanding water vapor heated by biomass like wood or later coal — pulling passenger and freight cars was a vast improvement over most inland canals that flourished through Europe and the U.S. in the generations before the railroad. Steam propulsion, however, could not provide the fast, frequent, and clear (coal soot in tunnels) power needed by growing cities such as New York or London. The advent of the electric traction motor in 1880 (Thomas Edison's first installation in the U.S.), the multiple-unit control system in 1898 (allowing a "train" of self-propelled cars to be controlled by one motorman), and the diesel engine by the 1920s vastly improved a trains ability to move people more quickly.

After the end of the Second World War, however, passenger rail suffered a decline across the U.S. which only bottomed out in the early 1970s. This occurred in Western Europe and Japan as well. However, the fall of passenger train travel in the U.S. was much more severe owing to much less financial public support for trains in the United States as well public support for the Interstate Highway System over a continental sized country (Vuchic, 2005).

By the 1970s public funding for rail transit began at both the federal and state levels. The Urban Mass Transit association was created in 1964 (became the Federal Transit Administration in 1991). At the state and local level bankrupt private firms that ran buses and rail transit were taken over or recreated by public transit agencies.
Of particular note is the rise of light rail. Light rail is a form of rail transit that straddles the space between full Metro that is separated from traffic and streetcar or trolley that is in "mixed-flow" with other cars. Although it had antecedents earlier in the interurbans in the U.S. or the German Stadbahn, University of Pennsylvania transportation researcher Vukan Vuchic finds that light rail in its various implementations can be thought of as a continuum from fully grade-separated fast Metro systems to slower streetcar/tramway systems (2007). Light rail thus provides a cheaper and more versatile technique for achieving rail systems in the U.S. and would lead to their expansion from the last few decades of the 20th century through to the present.

Later in the paper I will return to the continuum of rail transit modes. The costs and benefits of train "intensity" I will illuminate with my empirical work.

**Literature Review**

In University of Pennsylvania researcher Vukon Vuchic's popular transit economics textbook he discusses the basic approach to the use of multivariate linear regression in transit demand modeling (2005). The various scholarly articles I have surveyed thus far approach the problem typically using OLS (ordinary least squares) regression at the station level (Chan, 2013) & (Mangan, 2013), from one station to another (Choi, 2012) & (Zhao, 2014), and at the line or route level (Delbosc, 2013). Timms, interestingly, takes a step back from modeling and suggests alternative philosophical approaches to the problem (2008). Also Chan's work contains a table summarizing the various approaches to modeling transit ridership and what variables and econometric techniques that each approach uses, building on Vuchic's basic description in his textbook.
Also of interest, though not an approach to model transit ridership, is Belmonte's approach (2014). He used an OLS regression linear probability model (LPM) to find the probability that any individual will use transit with a focus on income level as an explanatory variable. His heavily data driven work drew from the U.S. Census Bureau's American Community Survey (ACS), American Public Transportation Association (APTA) operational data, and the Texas A&M Transportation Institute’s 2012 Urban Mobility Report.

**My Approach's Contribution**

Most recent papers have approached the modeling problem at the station level or analysis of a single line or transit system. A few have compared routes across many systems, encompassing rail and bus modes. My approach will look at multiple rail transit systems in the United States with reference to operational and basic demographic metrics. I will drive that data into an OLS multivariate regression approach with my key variable as headway. In its most conceptual econometric form my approach is:

\[
\frac{\text{Ridership}}{\text{unit of length}} = \beta_0 + \beta_1 \text{Headway} + \beta_2 \text{Other Operational Controls} + \beta_3 \text{Demographic Controls} + \mu
\]

Expressing ridership per unit of length of the rail lines for the dependent variable allows us to compare transit systems with widely varying total ridership with a common metric. Others have used this approach (see especially (Delbosc, 2013)). Headway is my primary explanatory variable of interest. Control variables can be broadly classed as other operational variables (discussed below) and demographic characteristics.

In a parallel approach I will use panel data related to the headway change of one line of one rail transit system and draw conclusions from that. My policy analysis following the econometrics will be informed by my experience and knowledge of the industry.
AMERICAN PUBLIC TRANSPORTATION ASSOCIATION (APTA) OPERATIONAL DATA

APTA Dataset Described

APTA is a nonprofit organization which serves as an advocate for the advancement of public transportation programs and initiatives in the United States. They are a leader in coordinating management and technical cooperation and knowledge sharing between transit providers.

APTA collects extensive data on many operational, economic, and demographic aspects of public transportation. They briefly describe it as:

APTA produces a number of specific products that serve as the resource for industry statistics, including the annual Fact Book. The Fact Book provides a wide range of data for all public transportation systems, including those that do not report to the National Transit Database. As such, this particular resource is considered the most comprehensive in the transit industry. Other statistical products generated by APTA, including the vehicle database, fare report, and infrastructure report, are based in large part on data provided to APTA by its members.

In addition to APTA products, a number of outside data resources provide useful agency specific information. Most notably, the National Transit Database serves as a common resource for public transportation systems. (American Public Transportation Association, 2014)

The dataset has a panel or longitudinal characteristic to it. That, is for example, it collects monthly ridership data for its member systems. However, for my OLS regression work I will not be using this characteristic of the data. Rather, I will be looking at a "snap-shop" in time of ridership data to drive my regression results.
Key Variables from the APTA Dataset

My main explanatory variable from the APTA data is:

- **Headway** — The time between trains. Such as a 10 minute headway (or six trains an hour) or a five minute headway (or 12 trains an hour) (Vuchic, 2005).
  
  - Note the often confusing distinction between headway (time between trains) and frequency (trains-per-hour). They are actually inverses of one other (that is, flipping over the numerator and denominator). Frequency is the more convenient for to use for mathematically.
  
  - Note that APTA does not collect headways for individual lines but rather the minimum headway for a transit system's lines is ever achieved in regular service.

Several other variables that I have an intuition are good controls variables for an OLS regression include:

- **Size of System** — The size of the systems such as in miles or kilometers. Special care should be taken as to whether the statistic is referencing miles of track or route miles or some other similar formulation. Confusion in this regard can lead to vastly over or undercounting the size of the system.

- **Number of Stations** — Combined with the size of the system this can produce a ratio metric that expresses the average distance between stations. For many rail transit systems station spacing can be very tight such as having two or more stations in a mile, even more for streetcars. For longer distance systems like American commuter rail or German S-Bahn the station spacing can be much longer such as several miles between stations.
- **Ridership** — Can be expressed a number of ways, such as total system, a particular line, boardings at a certain station, or passengers per kilometer or mile. Can also be expressed as rush hour, monthly, or yearly aggregates.

- **Riding Habit** — Defined as the ratio of annual transit rides to the population of the served area, indicating how much of the population utilizes the transit system (Vuchic, p. 43).

- **Line** — A particular service on a track. While it generally follows a path something like a twisted line it may also be in the form of a "circle line" and may include express or skip-stop service.
  
  - Note that APTA does not collect line-by-line data. However, transit agencies typically publish line-by-line data on their websites or otherwise make them publically available.

- **Speed** — The metric for speed can have varying definitions. The maximum speed of the train vehicle in service versus the average speed over the length of the line, for instance (Vuchic, p. 629). The latter is always lower than the former and often amazingly so. The case of San Francisco's MUNI system a severe example, with an average 8.1 miles-per-hour (City and County of San Francisco: Office of the Controller - City Services Auditor, 2014).
  
  - In my study of the APTA data I discovered that both of these metrics have serious problems when applied to econometric study as I will discuss below.

- **Number of Tracks** — Most systems have two tracks, like a two lane road. However, a few have three or more tracks allowing them to run express service. (The New York Subway is the prime example of the expressed tracked system). Many have sections
"single-track" bottlenecks that prevent headways from ever dropping below approximately 20 minutes.

- **Connecting Modes** — APTA’s statistics track the number of other services, buses, intercity trains, other train lines, that a system connects to. Presumably connecting services encourage ridership.

- **Disabled Access** — While all new systems in the U.S. are 100% compliant with the Americans With Disabilities Act (ADA) many older ones are not. APTA tracts the proportion of each system’s stations that are ADA compliant.

- **Fare Structure** — Rail transit fares may be distance based, flat (no regard for how far you travel), and have various discount categories such as seniors and the disabled.

- **Number of restrooms and childcare facilities at stations.**

- **Financial operations metrics, especially:**
  - **Farebox Recovery Ratio (FRR)** — This is the proportion of the system's operating costs that are covered by fares paid by passengers. Nearly no passenger rail system in the world exceed a one-to-one ratio. Many are significantly closer to zero than they are to one. The New York Subway's 0.7 ratio is considered good by American standards.

- The APTA statistics also cross-reference some demographic information such as the population of the system's service area and standard geographic urbanized coding, among others.

---

3 Only about 100 of the New York Subway's more than 420 stations are ADA compliant (MTA New York City Transit, 2015).

4 There is some consensus around the assertion that high-speed rail such as the French TGV, German ICE, or Japanese Shinkansen can achieve a ratio of greater than one. Also the Honk Kong MTR rail systems is sometimes quoted as have a ratio greater than one. However, that they can achieve this without some built in public subsidy is debated.
**DESCRIPTIVE STATISTICS AND GRAPHS FROM THE APTA DATASET**

**Initial Impressions of the Data**

To get an initial impression of what the operational data looks like I took 31 U.S. rail transit systems and laid out some of the key operational variables. I then ran some basic descriptive statistics on some of those variables, frequency distributions, and bar graphs to get a sense of the data as well.

**Table 1 — Ridership and Other Select Operational Metrics for 31 U.S. Rail Transit Systems for June 2014**

<table>
<thead>
<tr>
<th>Name &amp; City/Region</th>
<th>Ridership (June 2014)</th>
<th>Average Daily Ridership</th>
<th>Size (in miles of tracks)</th>
<th>Ridership Per Track Mile</th>
<th>Maximum Speed</th>
<th>Average Speed</th>
<th>Minimum Headway (in minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>San Diego Sprinter DMU</td>
<td>212,900</td>
<td>7,097</td>
<td>44.0</td>
<td>4,838.64</td>
<td>.</td>
<td>.</td>
<td>80</td>
</tr>
<tr>
<td>Cleveland RTA Rapid Transit, LRT and HRT</td>
<td>655,779</td>
<td>21,893</td>
<td>68.46</td>
<td>9,579.01</td>
<td>.</td>
<td>.</td>
<td>7</td>
</tr>
<tr>
<td>San Jose VTA LRT</td>
<td>928,773</td>
<td>30,792</td>
<td>79.6</td>
<td>11,605.19</td>
<td>55</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>Pittsburgh &quot;The T&quot;</td>
<td>703,085</td>
<td>23,436</td>
<td>51.2</td>
<td>13,752.13</td>
<td>50</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Dallas Area Rapid Transit</td>
<td>2,388,882</td>
<td>79,402</td>
<td>152.0</td>
<td>15,621.64</td>
<td>80</td>
<td>25</td>
<td>8.73</td>
</tr>
<tr>
<td>St. Louis Metrolink</td>
<td>1,491,724</td>
<td>49,724</td>
<td>90.5</td>
<td>16,485.14</td>
<td>59</td>
<td>29</td>
<td>6</td>
</tr>
<tr>
<td>Salt Lake City TRAX and Streetcar</td>
<td>1,526,349</td>
<td>50,945</td>
<td>71.2</td>
<td>21,465.59</td>
<td>65</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Baltimore Light Rail and HRT</td>
<td>1,941,592</td>
<td>64,720</td>
<td>86.9</td>
<td>22,342.89</td>
<td>50</td>
<td>22</td>
<td>8.3</td>
</tr>
<tr>
<td>Buffalo New York Metro Rail</td>
<td>324,087</td>
<td>10,803</td>
<td>14.1</td>
<td>22,984.89</td>
<td>.</td>
<td>.</td>
<td>10</td>
</tr>
<tr>
<td>Charlotte LYNX LRT</td>
<td>432,536</td>
<td>14,418</td>
<td>18.6</td>
<td>23,254.62</td>
<td>.</td>
<td>.</td>
<td>10</td>
</tr>
<tr>
<td>New Orleans RTA Streetcar</td>
<td>663,007</td>
<td>22,100</td>
<td>25.28</td>
<td>26,226.54</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>Phoenix Metro Light Rail</td>
<td>1,042,363</td>
<td>34,745</td>
<td>39.18</td>
<td>26,604.47</td>
<td>55</td>
<td>.</td>
<td>12</td>
</tr>
<tr>
<td>San Diego Trolley LRT</td>
<td>3,258,041</td>
<td>108,601</td>
<td>102.6</td>
<td>31,754.79</td>
<td>50</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>Portland MAX LRT and Streetcar</td>
<td>3,425,200</td>
<td>114,173</td>
<td>104.4</td>
<td>32,808.43</td>
<td>55</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>New Jersey Light Rail</td>
<td>1,729,593</td>
<td>57,653</td>
<td>50.4</td>
<td>34,317.32</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>Miami Metrorail</td>
<td>1,686,597</td>
<td>56,220</td>
<td>47.8</td>
<td>35,284.46</td>
<td>58</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Seattle, Washington Link Light Rail</td>
<td>1,160,600</td>
<td>38,687</td>
<td>31.2</td>
<td>37,198.72</td>
<td>55</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>San Juan, Puerto Rico Tren Urbano</td>
<td>770,349</td>
<td>25,678</td>
<td>20.62</td>
<td>37,358.31</td>
<td>62</td>
<td>.</td>
<td>8</td>
</tr>
<tr>
<td>Philadelphia (except regional rail)</td>
<td>7,310,000</td>
<td>250,333</td>
<td>194.33</td>
<td>38,645.60</td>
<td>.</td>
<td>.</td>
<td>3</td>
</tr>
<tr>
<td>Denver Regional Transportation District</td>
<td>2,123,379</td>
<td>70,779</td>
<td>47.1</td>
<td>45,082.36</td>
<td>55</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>San Francisco BART</td>
<td>11,145,026</td>
<td>371,501</td>
<td>209.04</td>
<td>53,315.28</td>
<td>70</td>
<td>.</td>
<td>15</td>
</tr>
<tr>
<td>San Francisco MUNI LRT, Cablecar, and Streetcar</td>
<td>4,102,450</td>
<td>136,748</td>
<td>74.8</td>
<td>54,845.59</td>
<td>52</td>
<td>8.1</td>
<td></td>
</tr>
<tr>
<td>Atlanta MARTA</td>
<td>5,706,723</td>
<td>190,224</td>
<td>103.7</td>
<td>55,031.08</td>
<td>.</td>
<td>.</td>
<td>10</td>
</tr>
<tr>
<td>Los Angeles Metro Rail LRT and HRT</td>
<td>8,345,713</td>
<td>278,324</td>
<td>150.8</td>
<td>55,369.45</td>
<td>70</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Minneapolis METRO</td>
<td>1,381,469</td>
<td>48,049</td>
<td>24.71</td>
<td>55,839.49</td>
<td>.</td>
<td>.</td>
<td>10</td>
</tr>
<tr>
<td>Houston Metrorail</td>
<td>1,020,711</td>
<td>34,024</td>
<td>14.3</td>
<td>60,968.56</td>
<td>40</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Chicago &quot;L&quot; subway and elevated</td>
<td>20,191,271</td>
<td>675,042</td>
<td>221.8</td>
<td>91,033.68</td>
<td>55</td>
<td>18</td>
<td>1.5</td>
</tr>
<tr>
<td>Washington, DC Metrorail</td>
<td>23,356,655</td>
<td>797,899</td>
<td>212.6</td>
<td>112,590.10</td>
<td>59</td>
<td>33</td>
<td>3</td>
</tr>
<tr>
<td>Boston MBTA subway and Light Rail</td>
<td>20,371,952</td>
<td>679,065</td>
<td>129.3</td>
<td>157,555.70</td>
<td>50</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>New York - New Jersey PATH</td>
<td>6,537,001</td>
<td>217,900</td>
<td>28.6</td>
<td>228,566.47</td>
<td>55</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>New York City Subway</td>
<td>230,419,298</td>
<td>7,680,648</td>
<td>658.6</td>
<td>349,862.28</td>
<td>50</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

---

5 All data from the OLS regression are from APTA's June 2014 Ridership Report.
Some initial observations of the above:

- Maximum speed is probably not a good metric at all. It is intuitively obvious that the time one spends at a speed is far more important. That data point is likely put into APTA's report at the insistence of transit agencies to make their systems seem faster.

- Average speed is widely not reported to APTA by the transit agencies. Firstly the number is rather embarrassing, especially in the case of San Francisco MUNI where their average system speed is widely panned in the media as being simply a brisk walking speed. Moreover, scheduled average speed — obtained by comparing dividing the distract of a line and dividing it by the timetable running time — will surely be slower than the actual speed as trains and buses in mixed flow (streetcars) are often delayed. Until recently active GPS tracking of buses and trains was not available. Though it may be available now it is not surprising that it is under-reported to APTA.
Figure A — July 2014 Ridership of 31 Leading U.S. Rail Transit Systems

Figure B — July 2014 Ridership of 30 Leading U.S. Rail Transit Systems without New York City
Take note from figures A and B that is clear is that rail transit use in the United States is dominated by New York City. In fact, the above graph and tables confirms the widely reported statistic that most rail transit trips in America take place in New York City. That New York is a significant outlier here will have consequences for my models and whether to include it or not. In the descriptive statistics you will note that leaving New York out of the analysis greatly reduces the standard deviation and begins to close the gap between the mean and median values.

Figure C — July 2014 Ridership of 31 Leading U.S. Rail Transit Systems Divided by Track Mile

In Figure C I divided the ridership by the size of the system in miles of track, thus getting a metric that accounts for the fact that some systems are bigger than others. Simply put, how many passengers are there for each mile of track? You will note above that New York still was the highest by this metric but was not the same level of outlier as it was for total ridership.
The three histograms above show frequency distribution of: a) total ridership without New York City (Figure D), b) ridership per track mile (Figure E), and c) minimum headway (Figure F). You will note, similar to the bar graphs, that in the first two histograms that there is a noticeable skew to the right with only a few "heavy" systems on that right side of the axis. New York was left out of the Figure D ridership histogram to keep it legible. The basic skew pattern remains when ridership is divided by miles of track. Finally the headways histogram (Figure F)
Econometric Technique

Below I will look at some scatterplots with an eye to the intuition of my first empirical approach, OLS regression. Generally, my main variable of interest on the dependent side is ridership. On the right hand side my independent variable interest is train frequency. Take note that train frequency is commonly expressed in customer literature as headway (time between trains, usually in minutes). This is the intuitive experience of taking a train and waiting for it. However, forming the X-axis this way for econometrics is more confusing. Therefore, I express headway below in "trains-per-hour". As this makes the X axis "rise" to the right in an intuitive manner it is the common way that headway is expressed in operational planning literature (Vuchic, 2005).

Figure G — Scatterplot of Minimum Headway and Total Ridership for 2013of 74 Rail Transit Systems

---

6 From APTA's 2014 Fact Book, "Operating Data".
In the scatterplot in Figure G each point is a rail transit system graphed on a two-dimensional axis by total ridership (y-axis) and minimum headway (x-axis). Note that there is no pattern between headway and total system ridership can be discerned. Only the New York Subway, Washington Metrorail, and the Chicago rail transit system are discernible above the "mass" of U.S. rail transit systems. This was to be expected as New York City dominates the total ridership count. Furthermore, these data points vary greatly in size of the track network, some only several miles large while the largest are well over a thousand track miles.

Figure H — Scatterplot of the Natural Log of Total Passengers and Minimum Headway for 2013 of 74 Rail Transit Systems

---

7 The Y-axis is in a conventional arithmetic scale. Values are omitted from each y-axis "tick" for clarity.
The problem of skewed data can often be rectified by converting the data into a natural log form (Wooldridge, 2013). Furthermore, as I will outline further below, having a regression formula in the form of a natural logarithmic dependent variable with arithmetic or "level" independent variables has significant advantages for interpretation. In Figures H and I this is done. Figure H expresses the natural log of the total ridership (Pax) while Figure I shows it in terms of the natural log of passengers per track mile.

In both plots we now see a distinction pattern that follows the intuition. Namely that as headway increases ridership increases. Interestingly, the plot seems to be concave down. That is
ridership increases as headway increases at a decreasing rate. The intuition simply stated is that adding more trains, past a certain "saturation point" has rapidly diminishing marginal returns.

**MULTIVARIATE OLS REGRESSION OF RIDERSHIP ON HEADWAY**

**Intuition Restated**

Restated graphically my intuition as to the relationship between headway and ridership is as follows:

**Figure J — Hypothesized Relationship Between Headway and Ridership**

More trains → More Passengers

Simply put more trains (reduced headway, more frequency) are positively correlated with more passengers. Therefore the converse must also be true, fewer trains (greater headway, less frequency) are correlated with fewer passengers.

While I also suspect there is a diminishing return on this formula based on the scatterplot above, the technique I will outline below only produces coefficients capable of expressing a linear relationship. I think this will be a very good approximately of the truth in any case and it is on that basis upon which I proceed.

**Basic Setup**

Based on the intuition that was refined from the descriptive statistics above I setup the following regression formula:

\[
\ln \left( \frac{\text{pax}}{\text{track mile}} \right) = \beta_0 + \beta_1 \text{headway (in trains per hour)} + \beta_i \text{controls} + u
\]
Expressed in words — the natural log of passengers per track mile is equal to a constant plus a coefficient times the headway, plus a coefficient times the other control variables plus an uncertainty term.

The control variables that will be use in the models below are as follows:

- **Stations** — Number of Stations
- **StationsPerMile** — Stations per track mile
- **Ratio of size and stations**
- **ExprTrks** — Express Tracks?
  - Zero if one or two mainline tracks
  - One if three or more mainline tracks
- **FRR** — Farebox Recovery Ratio (FRR)
  Obtained by the ratio of the below.
  - Fares – The fare revenues collected during the most recent closed-out annual report year.
  - Operating Expenses – For the most recent closed-out annual report year.
- **System Type**
  - Dummies for:
    - CR - Commuter Rail
    - HR - Heavy Rail
    - LR - Light Rail
    - SR - Streetcar
    - MG - Monorail/Automated Guideway
    - YR - Hybrid Rail.
    - Reference group is LR — Light Rail.
- **popUZA** — Urbanized Area Population – The population of the urbanized area as reported by the U.S. Census. This is the population count for the year the census was made. No population updates are made until the next Census (2020).
- **areaUZA** — Urbanized Area Square Miles – The square miles of the urbanized area as reported by the U.S. census.
- **areaSA** — Property Service Area – The square miles of the area served by the property, as reported by the transit property. This may include some rural (non-urbanized) areas.
- **popSA** — Property Service Area Population – The population of the area served by the property, as reported by the transit property.
- **SystemSize** — Size of System in Track Miles (APTA metric)
- **Pax** — Total Ridership (close out 2013)
- **AvTripLength** — The ratio of Passenger Miles FY per Unlinked Passenger trips.

**Results**

The results of the OLS regression were processed using the APTA data and the statistical program STATA. They are expressed on the following page:
Table 3 — OLS Regression Table

Dependent Variable — Natural Log of Passengers Per Track Mile

<table>
<thead>
<tr>
<th></th>
<th>Model 1 Naive</th>
<th>Model 2 Naive without New York Subway¹</th>
<th>Model 3 Begin Full Model</th>
<th>Model 4 Without New York Subway¹</th>
<th>Model 5 Without top five ridership systems²</th>
<th>Model 6 Without bottom five ridership systems³</th>
<th>Model 7 Without top five in station spacing</th>
<th>Model 8 Without bottom five in station spacing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Headway (trains per hours)</td>
<td>.0647795 (.0151241)***</td>
<td>.0615691 (.0143655)***</td>
<td>.0159287 (.0057367)*</td>
<td>.0154143 (.0057975)**</td>
<td>.0127595 (.0063293)**</td>
<td>.0135931 (.0056913)**</td>
<td>.0166744 (.0053651)***</td>
<td>.0122942 (.0055156)**</td>
</tr>
<tr>
<td>Controls</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.2759</td>
<td>0.2581</td>
<td>0.9030</td>
<td>0.8977</td>
<td>0.8930</td>
<td>0.9341</td>
<td>0.8893</td>
<td>0.9265</td>
</tr>
<tr>
<td>N (Observations)</td>
<td>74</td>
<td>73</td>
<td>74</td>
<td>73</td>
<td>69</td>
<td>69</td>
<td>69</td>
<td>69</td>
</tr>
</tbody>
</table>

Coefficients with standard errors in parenthesis for headway for each model.

*p<0.1*  
*p<0.05**  
*p<0.01***

¹ The New York subway accounts for 55% of all rail transit boarding in the United States.
² The top five ridership systems account for 73.1% of all rail transit boardings. They are San Francisco BART, Boston MBTA subway, Chicago CTA subway, Washington, DC WMATA Metrorail, and MTA New York City Subway.
³ The bottom five systems account for only 4.63 x 10⁻⁵% of the total U.S. rail transit ridership.
Notes on Interpretation

- The "naive" models were simply a regression of the natural log of ridership per track mile on headway with no other controls or right-hand side variables. The following seven models built upon that with a variety of robustness checks such as leaving out high and low ridership systems, and systems with unusual station spacing characteristics.

- As the form of the regression is a natural log dependent variable and an arithmetic independent variable the coefficient on Headway is understood to be subject to a "semi-elastic" interpretation (Wooldridge, 2013).

- That is a one train per hour increase in headway is correlated with a $100\cdot \beta_1\%$ increase in ridership per track mile$^8$.

- For instance, adding one more train per hour under Model 3 is correlated with an $\approx 1.6\%$ in ridership.

Significant Results

- Headway for all models was positively correlated with ridership. This varied from a high about 6% for the naive models to about 1.5% for the "fully-controlled" models.

- All models had at least a p-value of 10% or lower for Headway with very high R-squared values throughout.

Problems with this OLS Regression Approach

The sample size for this regression is rather small, 74 observations ($N$). This is partly due to the nature of the question and the way this data is gathered by APTA at the level of transit system and not for each individual line of rail transit system as noted earlier. Therefore the number of rail transit systems report was never going to much higher than 74 even if their dataset did not

---

$^8$ Actually the precise interpretation is $\%\Delta y=100(e^{\beta}-1)$. However, $100\cdot \beta_1\%$ is approximately correct for betas between -0.1 and 0.1. See (Wooldridge, 2013)
have any missing or incorrect data. Some agencies did not report any data. In general, more observations leads to greater econometric validity.

Another factor is that headway is, along with the other data, self reported by the agency. A fine point on the definition of headway is whether lines that come together to share a common truck should have their headway counted for each individual branch (more time between trains) or for the trunk (less time between trains). Judging from my inspection of the data and comparing it to some agencies published timetables no consistent pattern for this rule was followed.

Moreover, the minimum headway achieved on any part of an individual transit system's network was only an approximation of the universe that I actually wanted. The universe that I believe will have more explanatory power would be the one where data for each line of a transit system is available as to its minimum headway. This would cause a significant multiplier effect in the sample size. For instance, the New York Subway has about 24 individual lines or services. Washington, DC Metrorail has five. Los Angeles Metro Rail has five as well. This data exists and is accessible but it is not collected in a central source\textsuperscript{9}.

Finally there may be some concern addressed regarding the actual direction of causality. That is — does ridership actually cause headway changes instead of the other way around? It can be argued that rail transit planners respond to increased ridership by scheduling more trains. Later in my development of the natural experiment difference-in-difference model I discuss this issue with regards to the true exogeneity of the headway independent variable.

In the next section I will go into detail on how I used a limited set of panel data regarding one transit system to capture the possible effect in the change in ridership over time via a natural experiment.

\textsuperscript{9} Therefore it was not practical data to obtain in the context of a master's thesis.
Background on Natural Experiments

In general in the social sciences econometric techniques are performed because experimentation akin to the physical or life sciences is not possible. That is, providing a treatment, controlling for confounding effects with a control group, and identifying differences between the control group and the treatment group that are due to the treatment.

A tempting technique would be to find a time a policy change occurred and measure the results from the policy change and interpret them. This intuitive technique to simply see what happened after a policy change occurred is often called a "pre-post" or "before-after" study design. However, because it lacks a control group with which to compare the treatment group it is generally methodologically flawed (Remler & Ryzin, 2011). Quite simply, without a properly designed control group you do not know whether the change was the result of the treatment of some other factor.

A possible fruitful approach them would be to combine the treatment, in our case the headway change, with a control group that is similar enough to the treatment group in all important respects so that we may know over time whether the treatment is really correlated with the change in the treatment group's state. Natural experimental design for changes in minimum wage policy have used another adjacent state's employment data where the change in policy did not occur as a control group for instance (Card & Krueger, 1994).

Therefore, it might be valid to proceed by finding a rail transit line where the headway changed and comparing it to another rail transit line in another city or state over the same time period. However, as I will outline below I have a more fruitful treatment and control group pair.
Background of Los Angeles Metro Rail system

The L.A. Metro Rail system is an urban rail system consisting of both light rail and heavy rail (subway) lines. It's four light rail and two heavy rail (fully grade separated) total currently almost 90 route miles and over 350,000 boardings on a typical weekday (Los Angeles County Metropolitan Transportation Authority, 2015).

Figure K — Map of the Los Angeles Metro Rail system
**Longitudinal Data From Los Angeles Metro Rail Description**

Metro Rail, like all transit agencies per U.S. federal guidelines, collects panel data on the ridership of each of its individual bus and rail transit lines. The data is in the form of a monthly ridership report expressed in average weekday, Sunday, and Saturday ridership for a given calendar month. This data is actually publically available on their website in the form of an interactive web based tool\(^{10}\). It is also available from their planning office in an easy to manipulate data based form\(^{11}\).

**Los Angeles Metro Rail Repeatedly Conducts a Natural Experiment in Headway Change**

Over approximately two years Los Angeles Metro Rail changed the frequency of one or more of their rail transit lines while leaving the others constant. The changes they made are as follows:

<table>
<thead>
<tr>
<th>Date</th>
<th>Line(s)</th>
<th>Headway Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>June 27, 2011</td>
<td>Gold</td>
<td>• Peak (rush-hour) headway was decreased from 10 minutes to six. Other lines keep the same headway.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Increase is from six trains per hour to 10. Net increase of four trains per hour.</td>
</tr>
<tr>
<td>Nov. 13, 2011</td>
<td>Blue, Purple, &amp; Red</td>
<td>• Nighttime (after 8pm) headway was decreased from 20 minutes to 10. Other lines keep the same headway.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Increase is from three trains per hour to six. Net increase of three trains per hour.</td>
</tr>
<tr>
<td>June 23, 2013</td>
<td>Gold &amp; Expo</td>
<td>• Nighttime (after 8pm) headway was decreased from 20 minutes to 10. Other lines keep the same headway.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Increase is from three trains per hour to six. Net increase of three trains per hour.</td>
</tr>
</tbody>
</table>

---

\(^{10}\) Go to [www.metro.net/news/ridership-statistics](http://www.metro.net/news/ridership-statistics). Then click on link for "Interactive Ridership Stats — Line Level Trends and historical info."  

\(^{11}\) Provided to the author in Microsoft Excel and Stata formats.
Results from the Difference-in-Difference Analysis

The basic formula to perform a difference-in-difference comparison is:

\[ Treatment\ Effect = (E_2 - E_1) - (C_2 - C_1) \]

Where \( E_2 \) = The post-treatment state of the treatment group.
\( E_1 \) = The pre-treatment state of the treatment group.
\( C_2 \) = The post-treatment state of the control group.
\( C_1 \) = The pre-treatment state of the control group. (Remler & Ryzin, 2011)

For each of the headway changes listed in Table D I performed the procedure as outlined above and expressed the change as a percentage from the pre-test time period to the post-test time period. My time period sub-1 under the above formula was the last full month before the treatment. My time period sub-2 under the above formula was one year after the time period in sub-1. This is a year-over-year or put another way a "one year later", analysis. The purpose of comparing year-over-year as opposed to looking at the first full month of the treatment is to control for seasonal travel patterns which can be quite pronounced in transit ridership. Factors such as the dampening of travel during the winter and when school is not in session are usually be significant (A Time Series Analysis of Transit Ridership, 2002).

Table 5 — Results of Difference-in-Difference Analysis of Los Angeles Metro Rail Headway Changes

<table>
<thead>
<tr>
<th>Date</th>
<th>Line(s)</th>
<th>Difference-in-Difference Year-Over-Year Analysis Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>June 27, 2011</td>
<td>Gold</td>
<td>7.73% greater increase in ridership than the control</td>
</tr>
<tr>
<td>Nov. 13, 2011</td>
<td>Blue, Purple, &amp; Red</td>
<td>9.57% greater increase in ridership than the control</td>
</tr>
<tr>
<td>June 23, 2013</td>
<td>Gold &amp; Expo</td>
<td>17.98% greater increase in ridership than the control</td>
</tr>
</tbody>
</table>
As expected, all rail lines that experienced a headway increase "treatment" showed a greater percentage increase in ridership than the control group rail lines.

**How This Compares to the Prediction from the OLS Cross Sectional Data**

It would be interesting to see how this approach compares with what the regression approach from earlier would predict for this "experiment". That is, what will we get if we "plug" the headway increase numbers from the natural experiments into the regression formulas developed earlier? As the log-level model is subject to a semi-elastic percentage interpretation is this fairly easy.

**Table 6 — Results of DiD Analysis With a Comparison to Predicted Values From the OLS Regression Models**

<table>
<thead>
<tr>
<th>Date</th>
<th>Line(s)</th>
<th>DiD Results</th>
<th>Increase Predicted My Naive Model #1</th>
<th>Increase Predicted My Full Model #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>June 27, 2011</td>
<td>Gold</td>
<td>7.73% increase</td>
<td><strong>25.91%</strong></td>
<td>6.67%</td>
</tr>
<tr>
<td>Nov. 13, 2011</td>
<td>Blue, Purple, &amp; Red</td>
<td>9.57% increase</td>
<td><strong>19.43%</strong></td>
<td>4.78%</td>
</tr>
<tr>
<td>June 23, 2013</td>
<td>Gold &amp; Expo</td>
<td>17.98% increase</td>
<td><strong>19.43%</strong></td>
<td>4.78%</td>
</tr>
</tbody>
</table>

**Problems with this Approach**

The difference-in-difference or natural experiment approach has some intuitive appeal to it. However, some concern may be raised about its econometric validity. I deal briefly with two of those issues below — internal validity and external validity.

**Internal Validity**

As the term implies, internal validity is the ability of the experiment to reveal genuine causal effects. For that to be to be true two conditions must be met. Firstly the independent variable must truly be exogenous. In the context of a natural experiment it means that the variation in the
independent variable cannot be driven by anything related to the outcome. Secondly, the treatment and control groups must be truly comparable or homogenous (Remler & Ryzin, 2011).

With regards to the first condition I believe a valid case can be made for its satisfaction. The independent variable, headway, for these natural experiments were driven by a political decision to increase headways and to promote that decision publically shortly before the headway increase was commenced (Hymon, 2011). There does not seem to be any path for the independent variable to be driven by any other variable related to the outcome. In fact, as 2011 was still during the Great Recession other variables were likely driving the ridership down and not up.

With regards to the second condition of the homogeneity of the control group to the treatment group. The map in Figure K shows that these rail lines are adjacent to each other in the same city. They share the characteristic of passing through a variety of neighborhoods and socio-economic conditions. They are also share a common fare structure, common marketing and branding, and a common macro-economic situation specific to the Los Angeles area. Other studies, such as the aforementioned one regarding the minimum wage and fast food workers, have used adjacent states as controls for natural experiments and have been praised for their validity (Card & Krueger, 1994).

External Validity

Whether the results of the Los Angeles "headway" experiment can be generalized to other rail transit systems in the U.S. (or even internationally) cannot be answered easily. Relevant factors in this regard might include:

- How much are Los Angeles train riders like other American train riders?

---

12 According to the press release: "It’s a demonstration project that aims to boost ridership and better serve entertainment, cultural and sporting venues in addition to restaurants, hotels and stores." The extra trains have not be rescinded to date.
• How similar are Los Angeles headways to headways of other systems?

• How similar is the percentage of transit ridership use versus private automobile use in Los Angeles to the rest of the U.S.?

• Are the length of people's commutes in other cities similar? The shorter the commute, the less you will be positively affected by trains running more often.

• American transit systems can be broadly categorized between "legacy" systems built about a century ago and newer ones built in the later third of the 20th century. Perhaps that is a useful breakpoint in how generalizable the behavior of potential rail transit passengers are as patterns of riding behavior may be different among older systems.

SUGGESTIONS FOR FURTHER RESEARCH ALONG THESE LINES

There are likely some ways my research question and methods could go deeper if expanded upon. I outline them briefly below.

OLS Regression of Headway

Ultimately a more fine grained version of my regression would be to regress the maximum headway for every individual line of a system and make the dependent variable the ridership for that line. This would greatly increase the number of observations for the regression without even going cross-country. (And of course going international would increase the N even more). Data for this approach exists but it is not collected into a central source. Rather, it is collected by each transit agency. This data in the U.S. is not considered protected government information and is openly available upon request. In fact it is often published online, though not always in a consistent data form. Therefore, it is labor intensive to reformat and "clean-up".
Ultimately, other researchers could likely use better control variables and more advanced econometric techniques than "simple" OLS regression to improve the econometric validity of this general approach.

Headway Natural Experiments
I thought to choose the headway natural experiment from Los Angeles Metro Rail in this case as I am a recent employee of the agency in rail operations management. Therefore I was familiar with this change as well as factors that I assert make the other lines good controls. There are other times this change occurred both within the same city as well as adjacent cities or regions. More work in this area pairing headway changes with different controls will likely increase confidence in whether the findings can be generalized, held to be externally valid, and a confidence interval around a coefficient can be established.

**POLICY IMPLICATIONS FOR EXISTING SYSTEMS — HOW OFTEN SHOULD WE RUN?**
Running trains more often is correlated with more ridership. This will intuitively cause both more revenue and more expenses. What those might be can be usefully summarized as follows:
### Table 7 — Summary of Revenues and Expenses as a Result of Increasing Headway (Vuchic, 2005) and author's own work

<table>
<thead>
<tr>
<th>Revenues</th>
<th>Expenses</th>
</tr>
</thead>
</table>
| - Income from fares  
  o Are a large portion of fares paid by unlimited monthly passes? This would mitigate (lower) potential fare gains. | - Operating Personnel Wages  
  o Most types of trains require a crew of one or two so this expense will be going up with more trains. |
| - Financial operating assistance (generally from the local, state, and federal governments):  
  o Is the assistance flat?  
  o It the assistance formulaic so that it goes up for more ridership?  
  o Is a special grant available to pay for this extra service initiative? | - Fuel and power expenses  
  o This must rise to run more trains. |
| - Advertising Revenue:  
  o Will this supplemental, usually minor, revenue stream go up or remain steady in response to the program to increase headway?  
  o Perhaps local business can pay for the extra service, especially the nighttime service by local "nightlife" venues. | - Maintenance and repairs.  
  o Rail vehicles have regular maintenance that must be performed after a fixed amount of miles are driven. Running more trains will drive up train maintenance costs. |
| - Filming Revenue:  
  o Los Angeles Metro Rail, as it is located in the center of the American film industry, generates revenue form "movie shoots". These are often done at night when trains run less or not at all. Taking away filming opportunities by running train more often could reduce this revenue source. | - Fare-collection expenses  
  o This probably will remain flat as especially modern electronic fare collection equipment will not incur extra costs due to a modest increase in "processing". |
|  | - Information, advertising, marketing  
  o Is the agency paying for advertising about more trains? Or is relying on free press releases, blogs, and other "free" media. |

### Attempting to Quantify the Costs and Benefits of More Trains on Existing Systems

Vuchic (2005) suggests the following broad formula for determining the efficiency of a rail transit system.

\[
\eta = \frac{\text{Output quantity produced}}{\text{Resource quantity expended}}
\]

Vuchic finds that the advantage of this "unit-less" ratio form to be that it "... permits comparison of systems of different sizes or magnitude, which would not be possible by..."
comparing absolute values of individual measures alone." (2005). The top half of the equation could simply be increased ridership or a metric such as ridership per mile, per seat-mile, vehicle miles traveled, or some kind of qualitative measure. The denominator will likely be expenses related to operations that can be given a monetary value in a more straightforward manner. Under this formula, the greater the value for \( \eta \) then the greater the efficiency in which output quantity produced in the numerator is being produced. A value of \( \eta \) close to zero conversely asserts that little of the output quantity produced is being made for a large expenditure of resources, relatively.

Building on Vuchic's work, I propose the following formula to conceptualize the efficiency of a headway increase on an existing rail transit system:

\[
B = \frac{\Delta pax \times NewRevPerPax}{ExpHdwyReduc} \times QualFactor
\]

Where:
- \( B > 1 \) means headway reduction is efficient
- \( B < 1 \) means headway reduction is inefficient
- \( \Delta pax \) = the change in passengers from the headway reduction
- NewRevPerPax = The new revenue generated by the headway reduction
- ExpHdwyReduc = How much will the headway reduction cost, such as increases in maintenance and labor.
- QualFactor = A "multiplier" to account for the qualitative value of the headway reduction

Note that without the qualitative factor so long the expense of the new service is greater than the numerator the headway reduction is inefficient. Therefore the qualitative factor "saves" a headway reduction that cannot be justified "on the numbers".

How Can These Qualitative Factors "Make It Worth It"

Waiting for the train longer dampens ridership. That seems a clear message of the empirical research of this paper. Rail transit development is often coupled with efforts to change land-use
around stations to make them more dense, to encourage economic development, and "get people out of their cars" (Transportation Research Board of the National Academies, 2004). Los Angeles Metro's motivation behind increasing service at night seems not to have been related to increasing revenue but rather to these factors.

Figure L — Los Angeles Metro Rail advertising for headway reductions

**POLICY IMPLICATIONS FOR NEW SYSTEMS — WHAT SHOULD BE BUILT?**

Most Americans use informal terms to discuss types of trains. Metros, subways, commuter trains, high-speed rail, trolley, tram, or simply "the train" are common ones. Those terms are rooted in what Vuchic calls the "Family of Rail Transit Modes" (Vuchic, 2007). These costs are driven by the "intensity" of several design and infrastructure characteristics summarized below:
Table 8 — Cost Drivers of Building Rail Transit Modes, adapted with author's observations and (Vuchic, 2005)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Intensity (lowest to highest cost)</th>
<th>Streetcar/Tramway</th>
<th>Light Rail Transit (LRT)</th>
<th>Rapid Transit (subway, heavy rail)</th>
<th>Regional Rail (commuter, S-Bahn)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exclusive right-of-way (ROW)</td>
<td>The more a right-of-way is separated from traffic, the more expensive the train is to build, on scale of 0% to 100% separated.</td>
<td>0-40%</td>
<td>40%-90%</td>
<td>100%</td>
<td>100%*</td>
</tr>
<tr>
<td>Platform Height</td>
<td>Low = 0-6 in. Medium = approx 20 in. High = approx. 40 in. or more.</td>
<td>Low</td>
<td>Low, medium, or high</td>
<td>High</td>
<td>Low, medium, or high</td>
</tr>
<tr>
<td>Access Control</td>
<td>None or Full</td>
<td>None</td>
<td>None or Full (less common)</td>
<td>None or Full (most common)</td>
<td>None or Full (rarely)</td>
</tr>
<tr>
<td>Operating Speeds Maximum and Average</td>
<td>Most maximum and average speed makes system infrastructure more costly. Max Speed (kph) Average Speed (kph)</td>
<td>60-70 12-20</td>
<td>60-120 118-50</td>
<td>80-120 25-60</td>
<td>80-130 40-75</td>
</tr>
<tr>
<td>Maximum Headway (Frequency)</td>
<td>More trains, more often, is more costly. Max trains per hour</td>
<td>60-120</td>
<td>40-90</td>
<td>20-40</td>
<td>10-30</td>
</tr>
<tr>
<td>Capacity</td>
<td>More capacity is more costly Passengers per hour</td>
<td>4,000-15,000</td>
<td>6,000-20,000</td>
<td>10,000-60,000</td>
<td>8,000-45,000</td>
</tr>
<tr>
<td>Reliability</td>
<td>More reliable from a greater investment in infrastructure, largely dependent on factors above. Low-Medium***</td>
<td>Low-Medium***</td>
<td>High+</td>
<td>Very High++</td>
<td>Varies%</td>
</tr>
<tr>
<td>Station Spacing</td>
<td>More station spacing reduces costs from building stations</td>
<td>250-500</td>
<td>350-1600</td>
<td>500-2000</td>
<td>1200-7000</td>
</tr>
</tbody>
</table>
Table 8 (con't) — Notes From Previous Table

* May have signal protected grade-crossings
** Automatic Train Protection — System to prevent train drivers from ignoring signals, automatically stopping their train if there is a danger of collision with another train.
*** Largely a function of automobile traffic.
+ So long as train is on its right-of-way with 100% priority over cars
++ Same as LRT but usually with better signaling systems
% In the U.S. commuter rail can face slowdowns from conflicts with freight traffic and lack of track capacity (only one track).
~ In the U.S. and Europe efforts are underway to "harmonize" and improve signaling systems. In the Europe the EU directive is for the implementation European Rail Traffic Management System (ERTMS, www.ertms.net) in the U.S. it is called Positive Train Control (PTC)

It is tempting to provide a table listing the per-mile or kilometer price of streetcar, light-rail, and rail rapid transit. However construction, land acquisition, labor, and materials cost very greatly by geography. Vuchic quotes the following common figures — $10 to $30 million per km for LRT, $60 to $100 million per km for rail rapid transit, and as little as $1 to $4 million per km for regional rail build on existing, upgraded tracks (Vuchic, 2007). However, extreme values abound, notably New York's Second Avenue Subway at approximately $2 billion per mile (Fink, 2012) or Washington Metrorail's Silver Line Phase 1 at more than $580 billion (Holeywell, 2012).

Rail Transit Infrastructure and a CBA (Cost Benefit Analysis)

The cost benefit analysis rose to prominence across public finance last century and transportation programs were no exception. The most prominent example in the U.S. was the publication in 1952 of the "Red Book" by the American Association of State Highway Officials explaining the CBA process (American Association of State Highway Officials (AASHO), 1960). Vuchic derives the following useful CBA formula for rail transit infrastructure:
\[ NPV = BPV - CPV = \sum_{t=0}^{L} \frac{B_t - C_t}{(1 + r)^t} + \frac{S_b - S_c}{(1 + r)^L} \]

Where:

NPV is Net Present Value

BPV is Benefits Present Value

CPV is Cost Present Value

r is the Discount Rate

L is the Years of Project Life

\( B_t \) and \( C_t \) are the Benefits and Costs in each year \( t \), respectively

\( S_b \) and \( S_c \) are the Salvage Benefits and the Salvage Costs

If you can assign monetary values to all relevant factors going into \( B \), \( C \), and \( S \) and you know the discount rate and the years in the project life you can find the Net Present Value of the project. If \( NPV > 0 \) than you should do the project. Otherwise do not build the project (Vuchic, 2005, p. 521).

The weaknesses of the CBA framework are well known. In common with other uses of the CBA there is: 1) The problem of what and how much to monetarily quantify qualitative non-market elements, 2) Small changes in the presumed discount rate leads to greatly different potential conclusions, and 3) Policy makers getting "blinded" by the attraction of a single, supposedly quantitative measure (Vuchic, 2005).

Vuchic finds a most serious problem however is that rail transit is typically a very long term investment and in fact can have permanent worth (2005).\(^\text{13}\) The CBA formula is unable to take account for permanent worth. Vuchic has particular concern that international financial institutions are rejecting rail projects in the developing world precisely because of what he would

\(^{13}\) While some 19th or early 20th century rail systems have been dismantled. Many more, including subway tunnels, are well over 100 years old and show no sign of being decommissioned.
calls a misuse of a CBA (2005). Often the transformative effects of rail transit on land-use are not taken into account, let alone even less monetarily quantifiable quality-of-life factors.

Attempts to compare projects with similar characteristics and life lengths are probably the most fruitful use of a CBA for rail transit in Vuchic's view (2005).

CONCLUSIONS

Quantitative Research

From my quantitative OLS work I have a high degree of confidence that decreasing headway (that is, increasing the numbers of trains per hour) is positively correlated with higher ridership. This is true across a variety of transit systems in terms of size (track-miles), ridership, and operating characteristics (such as light-rail or subway). My natural experiment approach using the Los Angeles Metro Rail system also returned positive headway coefficients that are very similar to the OLS regression approach.

Both quantitative approaches could have their econometric validity improved through better design and better data. I summarize those possible approaches below:
Table 9 — Deficiencies and Possible Remedies in Empirical Approaches

<table>
<thead>
<tr>
<th>OLS Regression</th>
<th>Natural Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Problem</strong></td>
<td><strong>Possible Remedy</strong></td>
</tr>
<tr>
<td>- Small sample size and missing data.</td>
<td>- Increase sample size by collecting data on individual transit lines.</td>
</tr>
<tr>
<td>- Minor discrepancies in headway definition</td>
<td>- Collect line-by-line data from each transit agency.</td>
</tr>
<tr>
<td>- Causality concern (does ridership cause more trains).</td>
<td>- Do more well designed longitudinal studies with panel data of ridership.</td>
</tr>
</tbody>
</table>

**Policy Implications — More Trains on Existing Systems**

I proposed a general mathematical formula, based on Vuchic's work, to determine whether a headway reduction is efficient. A significant factor in that analysis is a qualitative factor. In fact, it can amount to essentially a "saving" clause for headway reductions that do not "pencil-out" by "the numbers". The qualitative factor can be used to capture community desire for more trains, attempts to integrate more rail transit into land-use planning, provide service to late night workers, or other factors.

**Policy Implications — Building Systems Capable of More Frequent Trains**

A cost benefit analysis (CBA) framework can be useful to evaluate between competing modes or train types in a "alternatives analysis" of what to build (or not to build) so long as their characteristics have many similarities. However, a CBA's other failings are well known and they have analogues in passenger rail investment. The CBA cannot capture quality-of-life, land-use
developments related to rail transit, and participation in a global climate change mitigation regime, among a few examples.
Bibliography


