ANALYSIS OF FACTORS AFFECT HIGH SPEED TRAIN RIDERSHIP IN THE UNITED STATES —THE ACELA EXPRESS CASE STUDY

Zhenhua Chen, PhD student

zchen7@gmu.edu

703-835-1683

School of Public Policy

George Mason University

ABSTRACT

In this study, we focus on the Acela Express, and try to find out how selected internal and external factors affect the Acela Express's ridership. A two-stage least square regression model is introduced in order to eliminate the endogeneity problem caused by price and ridership. Also the Cochrane-Orcutt Procedure is adopted to solve autocorrelation. The result shows that ticket price and train on-time performances, which are used to being thought as important factors affect ridership become insignificant, while other factors like employment of business and professional in the Northeast Corridor areas have higher influence on high speed train ridership. The broader objective of this research is to provide policy suggestions for building of an efficient high-speed rail network that can both be profitable and solve practical problems that the contemporary transportation system faces.

Key words: High-speed Train, Acela Express, Ridership, Two-Stage Least Square

(I) INTRODUCTION

When President Barack Obama was sworn into White House, he set a priority on developing a high-speed rail plan for the country, which is thought to be one of the solutions for addressing the increasing traffic congestion and improving the environment. There are, however, significant differences between the United States and other countries, such as Japan, France and Germany, that have developed successful high-speed rail projects. It is still not sure whether high-speed rail will truly allure U.S. citizens to get out of their vehicles or off air planes and choose high speed rail as an alternative. The only high-speed rail in the United States, the Amtrak's Acela Express, which serves the Northeast Corridor (NEC) —from Boston via New York, Philadelphia, and Baltimore, to Washington, D.C., has attained positive revenue because of its steadily increasing ridership since it was put into service, and it has become the most shining service of Amtrak (See Figure 1).

Given these facts, it is not clear what factors affect Acela Express's ridership, or which factors contribute to the growth of ridership performance. Before a nationwide high-speed rail project construction, an empirical analysis of the current high-speed rail ridership, becomes necessary in order to understand the relative factors that affect high speed rail projects' success.

In this study, we will focus on the Acela Express, and try to find out how several selected factors affect the Acela Express's ridership using multivariable regression analysis. The broader objective of this research is to provide policy suggestions for building of an efficient high-speed rail network that can both be profitable and solve practical problems that the contemporary transportation system faces.

Figure 1 Amtrak Northeast Corridor

Source: Amtrak's monthly performance reports

Based on the daily experience, there are many factors affecting rail ridership, such as population density, levels of private vehicle ownership, topography, service frequency, fares, system reliability, and cleanliness. Studies also show that ridership increases with increased income, whether for business, personal, or leisure travel.[i/ii] One notable study, *Review and Analysis of the Ridership Literature* by Brian D. Taylor and Camille N.Y. Fink (2003), categorizes all these factors into two groups: (1) external factors, which are largely exogenous to the system and its managers, such as service area population and employment; and (2) internal factors, on the other hand, are those over which managers exercise some control, such as fares and service level. This study also categorizes studies of transit ridership factors into two general groups: (1) research that focuses on traveler attitudes and perceptions tends to use the descriptive approach; and (2) studies that examine the environmental, system, and behavioral characteristics associated with transit ridership, and tend to be structured as causal analyses. The significance of this study is that it examines both principal findings and methodological weaknesses for the two analytical approaches respectively. This study helps clarify the appropriate methodology for ridership research.

(II) DATA SOURCES AND METHODOLOGY

According to the literature on ridership studies, rail ridership attributes to many factors. Some of data such as rail ridership, ticket revenue, fare, are easy to obtain, while data of some other factors are not easy to obtain, such as topography, system reliability, cleanliness, etc. Due to the constraints of data accessibility, in this study, we intend to make a comprehensive factor analysis which is based on three internal factors----fare of both the Acela Express (high speed train) and Acela Regional Train (normal train) and on-time performance of the Acela Express, and nine external factors, including gasoline price and GDP, personal income and factors reflect employment of different careers in NEC. Meanwhile, in order to test how seasonal factor affects the ridership of the Acela Express, twelve monthly dummy variables are also included in this analysis. Therefore, the null hypothesis should be:

There is no relationship between selected factors and the Acela Express's ridership.

Dependent Variable

In this study, the dependent variable is Acela Express's ridership. In fact, the Acela Express was in service since December, 11, 2000. Unfortunately, we can only obtain its monthly ridership performance from its monthly report ranging from January, 2004 to July, 2009 released on Amtrak's website. In their reports, they have separated statistics of the Acela Express's monthly performance. Totally, we obtained 79 months data from January, 2003 to July, 2009.

Independent Variable

Acela Express Average Fare: This variable represents the ticket price of the Acela Express. Since the actual ticket price of the Acela Express varies in terms of different seats and departure time, here we only obtained an average price from the monthly performance report, using monthly ticket revenue divided by monthly ridership.

The Acela Regional Train's Average Fare: Acela Regional Train is a normal train which runs on the same track between Washington. D.C and Boston via New York. The Regional train targets general passengers who travel in NEC while the high speed Acela Express targets a premium passenger market. Compared with the high speed train, the regional train's ticket price is much lower. In our analysis, we are going to check whether the regional train's price do have influence on the high speed train's ridership. The data also comes from Amtrak's monthly performance reports.

On Time Performance: This is only a percentage number indicates the monthly on time performance of the Acela Express high speed train. The data also comes from Amtrak's monthly performance reports.

Gasoline Price: Gasoline price are normally treated as a key factor affect public's decision of choosing outdoor transportation tools. It is assumed that higher gasoline price contributes to less usage of private vehicle and more usage of public transport. This analysis will also test how this variable affects the usage of high speed train in the United States. The data of U.S. Regular All Formulations Retail Gasoline Price from the Energy Information Administration (EPA) is used as the gasoline price variable in this analysis.[iii]

GDP: The monthly national real GDP data is represented by Macroeconomic Advisers' index of Monthly GDP (MGDP), which is a monthly indicator of real aggregate output that is conceptually consistent with real Gross Domestic Product (GDP) in the NIPA's. The consistency is derived from two sources. First, MGDP is calculated using much of the same underlying monthly source data that is used in the calculation of GDP. Second, the method of aggregation to arrive at MGDP is similar to that for official GDP. Growth of MGDP at the monthly frequency is determined primarily by movements in the underlying monthly source data, and growth of MGDP at the quarterly frequency is nearly identical to growth of real GDP.[iv]

Disposable Personal Income: Prior research indicates personal income also affects the decision making of different transportation modes among general public. It is understandable that higher disposable personal income may increase people's preference on choosing premium transportation mode. In this analysis, we assume the disposable personal income has influence on the ridership of high speed train in NEC.

Employment of different careers in NEC: As one of the important demographic characteristics that affect regional transportation demand pattern, employment is assumed to be an essential factor that influences ridership of high speed train. In order to test whether employment of different careers in the Northeast Corridor Metropolitan Areas have different influence on the usage of high speed train, we introduce seven employment variables in terms of career type into our analysis. The seven employment careers are: professional and business, government, information, civilian labor force, manufacturing, and other service employment. The data unit of each career employment is thousands of persons, aggregated by employments of two major metropolitan statistical areas (MSA).[v]

Month Dummy Variable: Twelve monthly dummy variables are created to test how during different month, how does ridership of high speed train change. These variables can also help us to understand the characteristic of ridership fluctuation of the Acela Express high speed train.

(III) Descriptive Analysis

Table 1 shows the descriptive statistics of an array of preliminary variables. In this analysis, the primary goal is to find which of these variables have influence on the ridership of the Acela Express.

Table1 Descriptive Statistics of Preliminary Variables

Independent Variable (variable name)	Obs	Mean	Std. Dev.	Min	Max
Average Fare (p)	79	123.62	10.05	106	143
On time Performance (otp)	79	80.99	8.12	61.8	93.2
Acela Regional (nr)	79	565622.2	58650.16	418351	673628
Gasoline Price (gp)	79	236.3177	65.44	145.8	406.2
Real GDP (rgdp)	79	12730.94	524.803	11608.63	13507.66
Unemployment Rate (unem)	79	3.69	.86	2.7	6.5
Disposable Personal Income (dpi)	79	9720.48	900.11	8134	11236.3
Pro & Business Employment (bs)	79	1414.57	46.90024	1329	1496
Gov Employment (gov)	79	1519.68	23.27483	1457.1	1579
Information Employment (info)	79	309.8228	6.830849	291	324
Leisure & Hosp Employment (lh)	79	684.6456	41.77352	600	784
Civilian Labor Force (cl)	79	12195.32	272.3396	11745	12839
Manufacturing Employment (ma)	79	530.8405	38.55293	455.1	592.1
Other Services Employment (other)	79	422.8873	13.39579	394.6	447.3

In order to get a better result, three major relevant problems must be solved. The first problem is multicollinear problems among these preliminary variables. In this situation of highly correlated variables existing together, the coefficient estimates may change erratically in response to small changes in the model or the data. Table 2 shows the correlation among preliminary variables. As it clearly shows, the correlated coefficient among some variable pairs are higher than 0.8. For example, the variable *logp* and *logf* has a correlation coefficient at 0.8860, which indicates the log transformation of the high speed train's price variable and normal train's price variable are highly correlated, thus when doing regression, it is better to include only one of the two correlated variable so as to avoid multicollinearity.

Table 2 Correlation among preliminary variables

	logp	logf	otp	loggp	logbs	loggov	loginfo	logcl	loglh	logma	logoth
logp	1.0000										
logf	0.8860	1.0000									
otp	0.5946	0.7107	1.0000								
loggp	0.5994	0.7071	0.5684	1.0000							
logbs	0.7254	0.8666	0.6041	0.8507	1.0000						
loggov	0.6463	0.5917	0.4465	0.3557	0.4444	1.0000					
loginfo	-0.6283	-0.5280	-0.5083	-0.4149	-0.4048	-0.5592	1.0000				
logcl	0.7974	0.8738	0.6554	0.7344	0.8221	0.5890	-0.7261	1.0000			
loglh	0.5487	0.6793	0.4891	0.7283	0.8497	0.3766	-0.5387	0.8421	1.0000		
logma	-0.8468	-0.8296	-0.6729	-0.5970	-0.6918	-0.6400	0.8879	-0.9233	-0.7115	1.0000	
logother	0.7366	0.8021	0.6430	0.7589	0.8235	0.6552	-0.7575	0.9052	0.8598	-0.8917	1.0000

The second problem is endogeneity problem of analyzing the relationship of high speed train ridership and its fare. The initial reason that made us feel there should have an endogeneity problem is that, as the scatter plot of both price and ridership and log transformation of price and log transformation of ridership indicates, the price has a positive relationship with the ridership, which is different from our normal cognition on transportation price (Figure 2). Thus we doubt a much complicated relationship may exist between ridership of the Acela Express high speed train and its price. Also, according to some other public transit ridership scholar, the ridership of intercity passenger rail is affected by both rail supply and demand, as well as changes of price.[vi] While some of the factor variables are apparent exogenous, others are not. In fact, like the air transport, the Acela Express high speed train is a premium transportation mode which specifically focuses on premium passengers. Its price is not fixed, but varies at different travel time. For example, according to Amtrak's online ticket reservation website, ticket price of the Acela Express is offered at a lower price in non rush hour period, but is normally sold at a higher price during rush hour period. As a result, ridership and price can affect each other at the same time and thus may have a simultaneous relationship which inevitably causes an endogeneity problem.

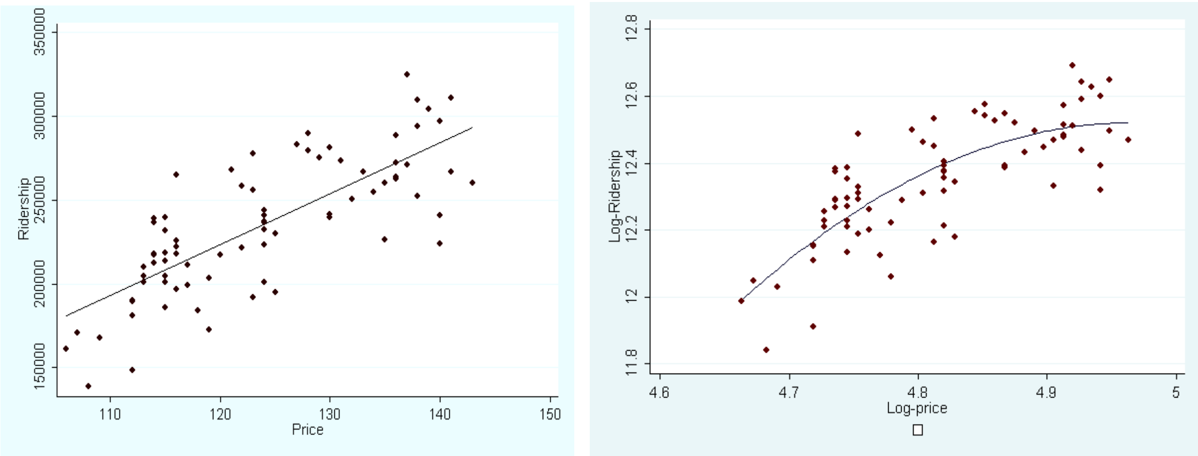

Figure 2 Scatter Plot of price and ridership

The third problem originated from the dataset itself. Since in our dataset, the 79 observations represents 79 successive monthly data concerning each variable. These adjacent observations are too similar than those that would be expected under independence. As a result, autocorrelation may happen which can make independent variables more significant than they may really be through smaller s.e for the beta coefficient.

In summary, in our case of the Acela Express ridership factor research, multicollinearity, endogeneity, as well as autocorrelation may happen simultaneously that requires a better solution to explain. Figure 3 shows the analysis structure, including regression problems and the related solutions in this analysis.

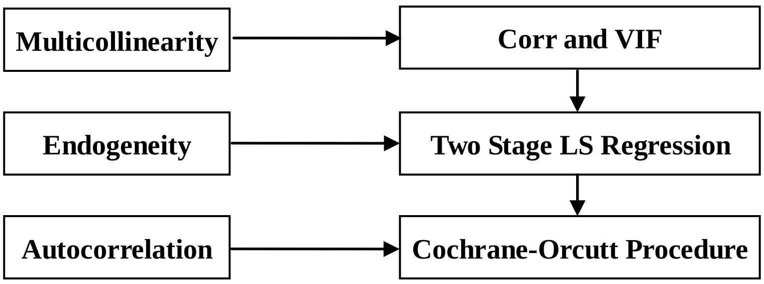

Figure 3 Analysis Structure

(IV) REGRESSION ANALYSIS

In the following discussion, we will show three steps of how our find model has been achieved, with solving multicollinearity, endogeneity and autocorrelation problems step by step.

Step One: Multicollinearity

Without considering endogeneity and autocorrelation, we begin our analysis by running two stepwise regressions using all of the preliminary variables to determine which ones have a statistically significant impact on the two dependent variables—ridership and price. As table 3 and 4 exhibit, the regressions respectively shows seven and six variables are statistically significant under 0.05 significant level.

Table 3 Result of Stepwise Regression on *Logp (logprice)*

Variable	Coef.	t-value	95% Conf.	Interval
logma	-1.589892	-12.66**	-1.840229	-1.339554
logrider	.1613656	7.52**	.1185857	.2041455
loginfo	2.20797	7.22**	1.597971	2.817969
loglh	-.4593681	-5.65**	-.6214109	-.2973252
otp	-.0014974	-2.91**	-.0025215	-.0004733
loggp	.0370914	2.16*	.0028707	.0713121
_cons	3.04647	2.36*	.4732928	5.619646

Note: Adj R-squared = **0.9031**, F = 0. † $p < .1$, * $p < .05$, ** $p < .01$.

Table 4 Result of Stepwise Regression on *Logrider*

Variable	Coef.	t-value	95% Conf.	Interval
logp	1.332894	5.81**	.8752803	1.790508
logbs	4.484596	7.04**	3.214228	5.754965
logcl	-6.649926	-5.45**	-9.082454	-4.217398
loginfo	-2.482033	-3.29**	-3.98524	-.978825
mar	.1123463	3.24**	.0432578	.1814347
dec	-.0832957	-2.21*	-.1584166	-.0081747
aug	-.0626187	-1.55	-.143036	.0177986
oct	.0814448	2.19*	.0073673	.1555222
jun	.065393	1.83†	-.005777	.136563
_cons	50.19376	4.25**	26.60767	73.77985

Note: Adj R-squared = **0.8496**, F = 0. † $p < .1$, * $p < .05$, ** $p < .01$.

Then we start to check VIF of the two regression equations and variable pairs with a correlated coefficient higher than 0.8, using command *corr* in STATA. Although the results of VIF of the two models are 2.87 and 4.32, which are not bad, the correlation check, as shown in Table 5, found some variables of employment among different careers are highly correlated. For example, variable *logcl* and log*p* and with *logbs*, *logma* with *logp* and *logma* with *logcl*, *loggp* with *logbs*. Therefore, we then try manual stepwise regressions on ridership as dependent variable and price as dependent variable, introducing statistical significant variables that have correlated coefficient lower than 0.8 with other variables step by step.

Table 5 Correlated Coefficients among Statistical Significant Variables

	logrider	logp	logbs	logcl	loginfo	logma	loglh	loggp	otp
logrider	1.0000								
logp	0.7629	1.0000							
logbs	0.6909	0.7254	1.0000						
logcl	0.5617	0.7974	0.8221	1.0000					
loginfo	-0.4519	-0.6283	-0.4048	-0.7261	1.0000				
logma	-0.5938	-0.8468	-0.6918	-0.9233	0.8879	1.0000			
loglh	0.4942	0.5487	0.8497	0.8421	-0.5387	-0.7115	1.0000		
loggp	0.5664	0.5994	0.8507	0.7344	-0.4149	-0.5970	0.7283	1.0000	
otp	0.5014	0.5946	0.6041	0.6554	-0.5083	-0.6729	0.4891	0.5684	1.0000

After controlled the correlation among variables, we got the following results with solving multicollinearity problem (See table 6 and 7). The Adj R-squared of regression model on price is 0.8496, which has a 0.0535 decrease from the prior model. All independent variables are statistically significant at 0.05 significant level except the dummy variable of November. On the other hand, the Adj R-squared of regression model on ridership is 0.6890 which represents a 69.90% explanatory ability of the real fact, has a 0.0854 decrease compared with the prior model on ridership. The independent variables are almost statistical significant at 0.05 significant level except the variables of gasoline price (*loggp*), on time performance (*otp*), as well as the dummy variable of September.

Table 6

Result of Regression on *Logp (logprice)* after eliminating Multicollinearity

Variable	Coef.	t-value	95% Conf.	Interval
logma	-.7894616	-10.53**	-.9388884	-.6400348
logrider	.1642355	6.01**	.1097168	.2187542
loglh	-.3646891	-3.69**	-.5616487	-.1677296
dec	.0357103	2.63**	.0086588	.0627619
loggp	.0557288	2.65**	.0138546	.097603
nov	.0246787	1.78†	-.0028946	.052252
_cons	9.810254	9.31**	7.709495	11.91101

Note: Adj R-squared = **0.8496**, F = 0. † p < .1, * p < .05, ** p < .01.

Table 7

Result of Stepwise Regression on *Logrider* after eliminating Multicollinearity

Variable	Coef.	t-value	95% Conf.	Interval
Logp	.6943342	2.89**	.215847	1.172821
loggp	-.1311333	-1.49	-.3071941	.0449274
otp	-.00093	-0.50	-.0046633	.0028033
logbs	3.921168	4.46**	2.169286	5.67305
jul	-.1463235	-3.35**	-.2335567	-.0590902
aug	-.1707968	-3.60**	-.2654453	-.0761483
sep	-.0293055	-0.67	-.1163288	.0577178
dec	-.1508018	-3.15**	-.2461594	-.0554443
_cons	-18.60999	-3.42**	-29.45425	-7.76573

Note: Adj R-squared = **0.6890**, F = 0. † p < .1, * p < .05, ** p < .01.

Step Two: Endogeneity

The second step is to tackle endogeneity. As we mentioned earlier, ridership is the result of a mutual effect from rail supply, demand, as well as price. Price can affect the ridership, but at the same time, ridership also affect price. As a result, a normal multiple variable regression becomes insufficient to explain this simultaneous relationship, since the dependent variable ridership has a endogenous problem with explanatory variable price. Thus, we introduce a two stage least square regression model. We follow steps listed below and finally get the result, as shown in Table 8.

- OLS regression on price only, with all exogenous variables which are statistically significant from previous test.
- Predict the estimates of price.
- OLS regression on ridership, with predicted price and other exogenous variables.

The Adj-R of the first stage model on price is 0.8029, which is 0.0467 lower than the prior model. All the independent variable is still statistically significant at 0.05 level except the variable of employment of leisure and hospitality (*loglh*). The Adj-R of the second stage model on ridership is 0.7404, which is 0.0514 higher than the previous model. Meanwhile, the variable of price, which is a predicted value with solving endogeneity, becomes statistically insignificant. The insignificant variables in this model also include on time performance of the train (*otp*), together with gasoline price (*loggp*).

Table 8

Result of Two-Stage LS Regression on *Logrider* after eliminating Endogeneity

Variable	Coef.	t-value	95% Conf.	Interval
logrider				
logp	.3808711	1.31	-.1932616	.9550037
otp	-.001459	-0.79	-.0050921	.002174
logbs	4.94085	5.24**	3.076054	6.805645
loggp	-.16749	-1.95†	-.3375815	.0026015
mar	.1119462	2.80**	.0328014	.191091
jul	-.1547896	-3.63**	-.2390823	-.070497
aug	-.184476	-3.92**	-.2774635	-.0914885
dec	-.1514875	-3.29**	-.2425227	-.0604523
_cons	-24.26789	-4.28**	-35.47994	-13.05584
logp				
loggp	.0761247	3.23**	.029511	.1227383
logma	-.8583523	-9.57**	-1.035683	-.6810217
loglh	-.15980091	-1.15	-.4334135	.1138117
jul	-.0411186	-2.42*	-.0747251	-.0075122
aug	-.0426847	-2.45*	-.0771193	-.0082501
nov	.033527	2.08*	.0016691	.065385
_cons	10.83152	8.46**	8.301267	13.36176

Note: Equation logrider: Adj R-squared = **0.7404**, F = 0.

Equation logp: Adj R-squared = **0.8029**, F = 0.

† p < .1, * p < .05, ** p < .01.

Endogenous variables: logrider logp

Exogenous variables: otp logbs loggp mar jul aug dec logma loglh nov

Step Three: Autocorrelation

Since our dataset is based on time series data, each observation which reflects performance of each successive month becomes correlated with other adjacent observation. Consequently, the observation makes independent variables more significant than they may really be through smaller s.e for the beta coefficient. After Durbin-Watson (DW) test of our previous model confirms that autocorrelation indeed exists (table 9 and 10).

In this study, we follow Cochrane-Orcutt Iterative Procedure to solve autocorrelation. The procedure is implemented by the Prais-Winsten transformation in STATA. The procedure of Prais-Winsten transformation is shown in Figure 4.

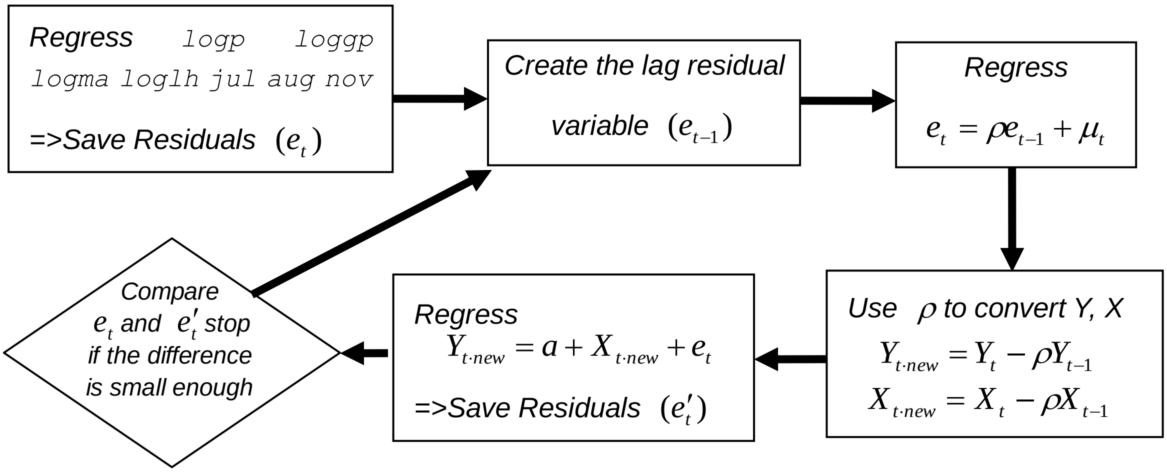

Figure 4 the procedure of Prais-Winsten transformation

The final results of the Cochrane-Orcutt Iterative Procedure are shown is Table 9 and 10. As we compared the DW value for the price regression on the first stage, it increased from the original 0.307812 to the transformed 1.44775. And the DW value for the ridership regression on the second stage increased from original 0.811855 to the transformed 2.000870, which indicates the autocorrelation has already been eliminated.

Table 9

Result of Regression on *Logp* after eliminating Autocorrelation

Variable	Coef.	t-value	95% Conf.	Interval
Loggp	.0329455	1.22	-.0208598	.0867508
Logma	-.9939553	-4.56**	-1.429193	-.5587171
Loglh	-.3021769	-2.23*	-.573126	-.0312279
Jul	-.0303305	-4.51**	-.0437706	-.0168904
Aug	-.0369788	-5.51**	-.050391	-.0235666
Nov	.0147482	2.62*	.003485	.0260113
_cons	12.86703	7.55**	9.464997	16.26907

Note: Adj R-squared = **0.4561**, F = 0. † p < .1, * p < .05, ** p < .01.

Durbin-Watson statistic (original) 0.307812

Durbin-Watson statistic (transformed) 1.444775

Table 10

Result of Regression on *Logrider* after eliminating Autocorrelation

Variable	Coef.	t-value	95% Conf.	Interval
Plogp	.4002818	0.78	-.627224	1.427788
Otp	.0018952	1.07	-.001637	.0054274
Logbs	3.646159	2.80**	1.048342	6.243975
Loggp	-.0837126	-0.73	-.3123146	.1448895
Mar	.0910242	3.68**	.0415303	.1405181
Jul	-.1183893	-3.78**	-.1809125	-.0558662
Aug	-.1470117	-4.05**	-.2195342	-.0744892
Dec	-.1374485	-3.75**	-.2106515	-.0642455
_cons	-15.69346	-1.91†	-32.10835	.7214255

Note: Adj R-squared = **0.5086**, F = 0. † p < .1, * p < .05, ** p < .01.

Durbin-Watson statistic (original) 0.811855

Durbin-Watson statistic (transformed) 2.000870

Result Interpretation

The null hypothesis is: there is no relationship between the selected factor variables (shown in table 1) and the Acela Express's ridership.

According to the final result of our two-stage least square model, we found a couple of variables have statistically significant relationship with the dependent variable.

The first stage regression (Table 9) identifies what exogenous variables have influence on the Acela Express's price. The final result shows employment of manufacture in NEC, employment of leisure and hospitality, and monthly dummy variables of July, August and November are statistically significant. Factors including on-time performance and gasoline price are not significant. Part of the reason is due to the multicollinearity among different variables. The Adj-R is 0.4561, which means this model have 45.61% of explanation for the real fact. The detailed interpretations of each variable are described as follow.

- *Logma*: 1% increase in the employment of manufacture is associated with around 1% decrease in the price of the Acela Express high speed train.
- *Loglh*: 1% increase in the employment of manufacture is associated with around 3% increase in the price of the Acela Express high speed train.
- *Jul*: the price of the Acela Express is associated with 0.30% decrease when it is July
- *Aug*: the price of the Acela Express is associated with 0.36% increase when it is August
- *Nov*: the price of the Acela Express is associated with 0.01% increase when it is November

The second stage regression (Table 10) identifies what exogenous variables affect on the ridership of the Acela Express. The final result shows employment of business and professional in NEC, monthly dummy variables of July, August and December are statistically significant. Factors such as ticket price, on-time performance, gasoline price are not statistically significant. The Adj-R is 0.5086, which means this model have 50.86% of explanation for the real fact. The detailed interpretations of each variable are described as follow.

- *Logbs*: 1% increase in the employment of business and professionals in the Northeast Corridor region is associated with 3.65% increase of the ridership of the Acela

Express high speed train

- *Mar*: the ridership of the Acela Express is associated with 0.09% increase when it is March
- *Jul*: the ridership of the Acela Express is associated with 0.12% decrease when it is July
- *Aug*: the ridership of the Acela Express is associated with 0.15% decrease when it is August
- *Dec*: the ridership of the Acela Express is associated with 0.14% decrease when it is December.

CONCLUSION

The result of this analysis above shows that the Amtrak's Acela Express strategy is quite penetrated on a premium passenger market. As some media reported, since the debut of the high speed train in NEC, air shuttle services which also dominated by premium passenger have suffered a huge slump. In our study, through a two-stage least square model, we have three major findings:

First, the study demonstrates that a higher business and professional employment contributes to higher usage of the Acela Express. This has revealed that in the United States, the high speed intercity train service is currently serving only as a premium transportation competing with air transport.

Second, both the high speed train price and ridership are strongly affected by seasonal factors. According to the result, generally speaking, the ticket price decreases in the third quarter and so does the similar trend on the ridership performance. In the fourth quarter, the performance of the high speed train seems a little uncertain because of the increasing price and decreasing ridership. One possible explanation for this phenomenon might be that since a business dominance marketing strategy, the ridership of the Acela Express is naturally associated with business performance in the Northeast Corridor. Normally, in the beginning and end of one year, business activity is much prosperous than in the middle of that year.

Third, the study shows that current riders of the high speed train are not sensitive to price and on-time performance, which used to be considered as key factors affect ridership. Part of the reason for the insignificance of price is that premium passengers pay more attention to travel amenity and service quality than price. And this also can be found from the goodness fit of the final model, because the model only shows roughly 50% of explanation for the real fact, which implies that other factors also play important role on ridership of the high speed train service. As for the insignificance of on time performance, two possible reasons may have influences. The first reason might be that compared with Europeans and Japanese, American has lower requirements on the on time performance of transportation. It is known that Japanese society has much strict punctuality requirement than American society, especially for people who is involved in business field. Facts can be seemed at the on time performance of Japanese bullet train Shinkansen. The average on time performance is almost 99%, and the figures are similar in Ave in Spain and French SNCF[vii]. The other reason might be current on time performance satisfied riders of the high speed train compared with other modes. The bad on time performance of American air shuttle service has been blamed for a long time. Compared with air, an average 81% of the on time performance might have already met riders' expectation.

In the United States, because of the dominance of air transport and the advanced national highway system, intercity passenger train has long been neglected. However, with the emergence of problems such as congestion, environment protection, as well as energy saving, high speed intercity train that has already demonstrated a great success in other countries starts to gain new attention in the United States. In this study, we tried to explore how the current high speed train performs, and it has demonstrated that it serves only for a premium passenger market. However, as President Obama said in his HSR speech on April 16, 2009, HSR will be a smart national transportation system equal to the needs of the 21st century, it is necessary to get the general public into consideration in the process of the HSR projects planning, designing, implementation and operation. Increasing the ridership should not be the ultimate goal. Rather, how to achieve the equilibrium of public transportation utility maximization and appropriate allocation of public transportation fund should be further explored in order to build a better high speed train in the United States in the near future.

Reference

Brian D. Taylor, and Camille N.Y. Fink. 2003. "The Factors Influencing Transit Ridership: A Review and Analysis of Ridership Literature." *UCLA Department of Urban Planning Working Paper*. http:// www.uctc.net/papers/681.pdf

Brian D. Taylor, Douglas Miller, Hiroyuki Iseki and Camille N.Y. Fink. 2003. "Analyzing the Determinnants of Transit Ridership Using a Two-Stage Least Squares Regression on a National Sample of Urbanized Areas." *2004 Annual Meeting of the Transportation Research Board*

Ewing, R. and Cervero, R., 2001, "Travel and the Built Environment: A Synthesis." *Transportation Research Record 1780.* Washington, D.C.: Transportation Research Board, National Research Council.

GAO, *High-speed Passenger Rail: Future Development Will Depend on Addressing Financial and Other Challenges and Establishing a Clear Federal Role*, CAO-09-317 (Washington, D.C.: March, 2009)

Gorsuch, R. L. *Factor Analysis*, 2nd ed. Lawrence Earlbaum Associates, Hillsdale, N.J., 1983.

Johanna Zmud, Mark Bradley, Frank Douma, and Chris Simek. Attitudes and Willingness to Pay for Tolled Facilities: A Panel Survey Evaluation. *Transportation Research Record: Journal of the Transportation Research Board, No. 1996*, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 58–65.

Maren L. Outwater, Steve Castleberry, Yoram Shiftan, Moshe Ben-Akiva, Yushang Zhou, Arun Kuppam. Attitudinal Market Segmentation Approach to Mode Choice and Ridership Forecasting: Structural Equation Modeling. *Transportation Research Record 1854*, paper No. 03-4481, pp. 32-42.

Passenger Rail Working Group. 2007. *Vision for the Future: U.S. intercity passenger rail network through 2050*.

Polzin, S. E., 2004, "Relationship Between Land Use, Urban Form And Vehicle Miles Of Travel: The State Of Knowledge And Implications For Transportation Planning," Tampa: University of South Florida, Florida Department of Transportation, Federal Highway Administration.

Sapleton, C.D. Basic Concepts and Procedures of Confirmatory Factor Analysis. Working Paper. Texas A&M University, Jan. 1997, pp. 34-48.

Yoav Hagler. 2009. "Where High-Speed Rail Works Best". America2050 04-05

[i] Polzin, S. E., 2004, "Relationship Between Land Use, Urban Form And Vehicle Miles Of Travel: The State Of Knowledge And Implications For Transportation Planning," Tampa: University of South Florida, Florida Department of Transportation, Federal Highway Administration. http://www.cutr.usf.edu/pubs/Trans-LU%20White%20Paper%20Final.pdf

[ii] Ewing, R. and Cervero, R., 2001, "Travel and the Built Environment: A Synthesis." *Transportation Research Record 1780.* Washington, D.C.: Transportation Research Board, National Research Council. http://depts.washington.edu/trac/concurrency/lit_review/trr1780.pdf

[iii] http://www.eia.doe.gov/

[iv] http://www.macroadvisers.com/content/MA_Monthly_GDP_Index.xls

[v] The two MSA are Washington-Arlington-Alexandria, DC-VA-MD-WV (MSA) and New York-Northern New Jersey-Long Island, NY-NJ-PA (MSA). Data is from U.S. Department of Labor: Bureau of Labor Statistics, released on http://research.stlouisfed.org/fred2/categories/27281

[vi] We greatly appreciate recommendations of Professor Brian D Taylor in UCLA and Professor Hiroyuki Iseki in NOU.

[vii] Central Japan Railway Company Data Book 2009. http://english.jr-central.co.jp/company/company/others/data-book/_pdf/2009.pdf