Elton Cheung
47306162
April 19, 2021

FRE 527 Final Paper

Intro:

Transportation infrastructure is an important element to the economic growth of the world. These infrastructures include roads that span coast to coast, tunnels that burrow through mountains and even bridges that connect countries. Countries all over the world invest significant amounts into transportation infrastructure resulting in diverse effects on the environment and economy.

Specifically, transportation infrastructure increases the movement of workers, who can get to places faster and find employment opportunities, and goods, which increases economic activity. Thus, it is hypothesized that economic growth is caused by building more transportation infrastructure. This paper aims to evaluate the effect of the increase of interstate highways on MSA trucking and warehouse employment in US cities. The main regression method to estimate this relationship is through instrument variables. By estimating this relationship, we can find how much employment is created through highways. This in turn allows us to determine whether or not creating new transportation infrastructures is a viable option to jumpstart places with low economic performance or growth.

Problems of Endogeneity:

Many previous studies that examined the impact of transportation infrastructure (specifically highways) on employment recognized the potential problem of endogeneity where reverse causality or simultaneity might occur. Where instead of highways creating more employment, it could be that places with high levels of employment lead to the development of highways. The reasoning for this could be that places with these high employment levels could indicate economic growth. Which in turn means these places would have the necessary tax revenue to improve or build more highways. Additionally, places with poor economic growth may see an increase in transportation infrastructure spending (i.e new highways) as a way to stimulate the economy (Jiwattanakulpaisarn et al., 2009, pg. 135).

All previous literature that was reviewed applied ordinary least squares (OLS) regression for their analysis. However, many applied additional methods that allowed them to deal with the problem of simultaneity and estimate a causal effect. Jiwattanakulpaisarn et al (2009) applies a Granger causality test for highway infrastructure on state employment growth using a panel data set. The Granger causality test essentially uses the lagged levels of the endogenous variable (in this case interstate highways) from previous time periods as instruments. However, when using this method, the problem of spatial autocorrelation arises if one fails to include “relevant variables that include neighbouring state road densities” (Jiwattanakulpaisarn et al., 2009, pg. 143). Due to the lack of geographical aggregation in the data, Jiwattanakulpaisarn et al (2009) found multiple problems of spatial autocorrelation that made causal effects insignificant.

On the other hand, Kim and Han (2016) uses a difference in differences (DID) technique to analyze the effects of building two new highways in Korea. With DID they estimated the change in population and employment growth before and after these new highways were built (treatment).

The third method that Duranton and Turner (2012), Michaels (2008) and Hymel (2009) used to deal with endogeneity was through instrumental variables estimation. What was interesting about these three papers is that they all used the US interstate highway plans (ranging from different years) measured in either miles or kilometers as an instrument for their endogenous regressor. Where the endogenous regressor is the length of the highway by the time the highway plan was almost completed (either in 1975 or 1983).

Comparing Results of Other Papers:

In regards to the results of the literature reviewed; all roughly came to the same conclusion. Where the creation of interstate highways leads to increased economic activities in the areas that the highways went through, which in turn increased employment growth. This was found by Duranton and Turner (2012), Michaels (2008), Kim et al (2012), Jiwattanakulpaisarn et al (2009) and Kim and Han (2016). On the other hand, since Hymel (2009) did his paper on the effects of traffic congestion on employment growth within the United States (U.S.) he found different results. Being that a 10% increase in congestion would reduce long-run employment growth by 4%. This conclusion could support the result of the other papers, as creating highways could reduce congestion and thus lead to an increase in employment growth. However, this assumption does contradict Downs (1962) "fundamental law of highway congestion", where building more roads does not decrease traffic congestion.

Data:

Since this paper is a replication study of Duranton and Turner's paper, "*The Fundamental Law of Road Congestion: Evidence from US Cities*" all the data is sourced from their study. To measure the effect of highways on employment growth (in both urban and non-urbanized areas), MSA-year panel data will be used. Where metropolitan statistical areas (MSAs) within the United States will be the main observations being studied overtime (Duranton and Turner, 2011, pg. 2618). An MSA represents a geographical region that groups together counties and cities based on similar economic and social factors. Furthermore, a large portion of an MSA's land area is sparsely populated due to MSAs being organized around "urbanized areas", meaning MSAs can be further divided up into urban and non-urban areas (Duranton and Turner, 2011, pg. 2621).

In order to measure employment, data from County Business Patterns was used (Duranton and Turner, 2011, pg. 2639). Data collected for years 1983, 1993 and 2003, provided information of county level employment in freight transportation (trucking) and warehousing. This provides the dependant variable to measure the MSA trucking and warehousing employment levels

Using data from the Highway Performance Monitoring System (HPMS), collected by the US Federal Highway Administration, the total length of interstate highways in kilometres (km) within each MSA can be collected for years 1983, 1993 and 2003 (Duranton and Turner, 2011, pg. 2621). Furthermore, this data also allowed the interstates within MSAs to be divided into urban and non-urban interstates.

Table 1 is the summary statistics for key variables that will be used within the econometric analysis section. Which includes the dependant variable, the main explanatory variables and the instruments that will be used. The total number of MSAs included in the study is 228, but since we are using three different points in time for each MSA it comes to a total of 684 observations.

Table 1: Summary Statistics of Key Variables

Variables	Obs	Mean	Std. Dev	Min	Max
*IH Length (km), 1983	684	1140.266	1647.345	8.304	12611.23
IH Length (km), 1993	684	1208.158	1726.898	111.345	13628.85
IH Length (km), 2003	684	1279.748	1854.857	118.195	14582.28
*IHU Length (km), 1983	684	510.154	1102.776	0	7961.15
IHU Length (km), 1993	684	621.989	1253.814	0	8867.764
IHU Length (km), 2003	684	719.508	1408.709	0	10048.43
*IHNU Length (km), 1983	684	630.112	654.339	0	5643.273
IHNU Length (km), 1993	684	586.169	581.299	16.093	5242.151
IHNU Length (km), 2003	684	560.239	557.866	0	5494.711
Log Truck/Warehouse Employment	684	7.781	1.351	4.094	11.916
1898 Railroad Route (km)	684	284.922	297.683	0	2033.64
1947 Highway Plan (km)	684	117.485	127.604	0	1010.7
1835 Exploration Routes	684	806.377	1050.679	0	5407
MSA Population, 1983	684	754000	1700000	49866	1.73E+07
MSA Population, 1993	684	834000	1860000	61226	1.78E+07
MSA Population, 2003	684	950000	2070000	66533	1.94E+07

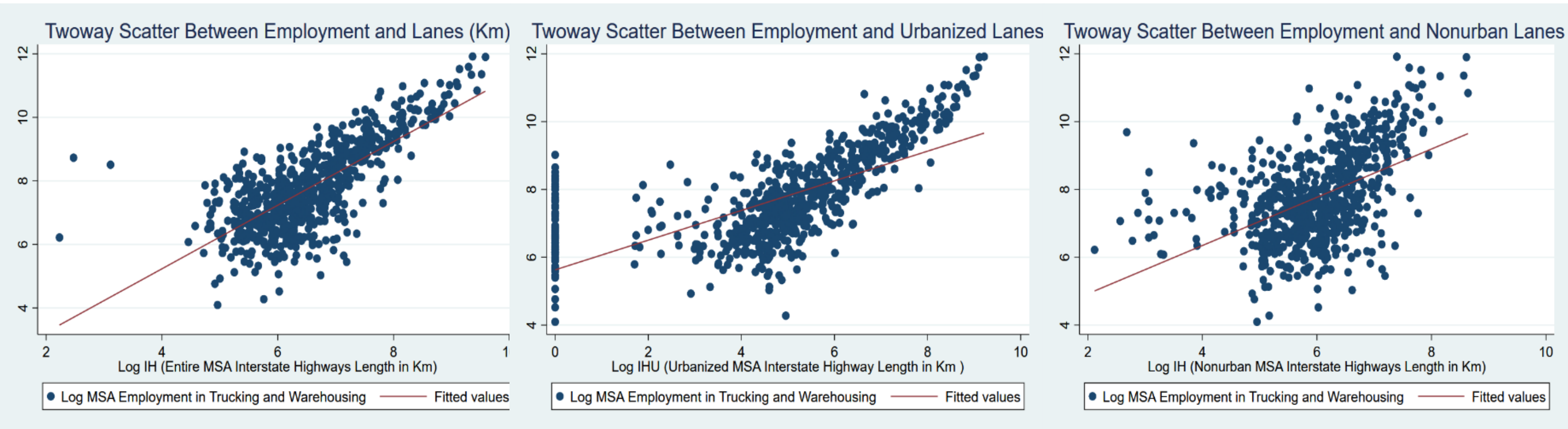
*Note: IH stands for interstate highway, IHU for interstate highway urban, IHNU for interstate highway non-urban

Econometric Theory/Defining Key Variables:

In this study the Log-Log model will be utilized, and most variables are transformed through the natural logarithm. The exceptions being some of the variables that control for geography, census divisions and socioeconomic characteristics. This model allows easier interpretation of the results as the regression coefficients represents the dependant variable's elasticity with respect to the explanatory variables. This section will discuss what variables will be used in the analysis and the underlying economic theory for including them.

To evaluate the effect of the increase of interstate highways on MSA trucking and warehouse employment in the US, the dependant variable will be the log count of MSA employment in trucking and warehousing. On the other hand, the main explanatory variable is the log of the entire MSA interstate highway (IH) length in kilometers. Additional explanatory variables include both logs urbanized (IHU) and non-urbanized (IHNU) MSA interstate highway lane kilometers. The underlying economic theory between the dependant and interstate highway variables is that highway infrastructure increases the flow of goods (which creates jobs) and people's ability to access job opportunities. From Figure 1, we see a positive trend between the dependant and highway variables. Thus, providing evidence for this economic theory.

Figure 1: Relationship Between Trucking/Warehouse Employment and Highways



Furthermore, since population density is also a determining factor for employment levels in trucking and warehousing, the log of population will be included as another explanatory variable in the model. The economic theory is that a higher population density indicates an urbanized area, which in turn provides better job opportunities.

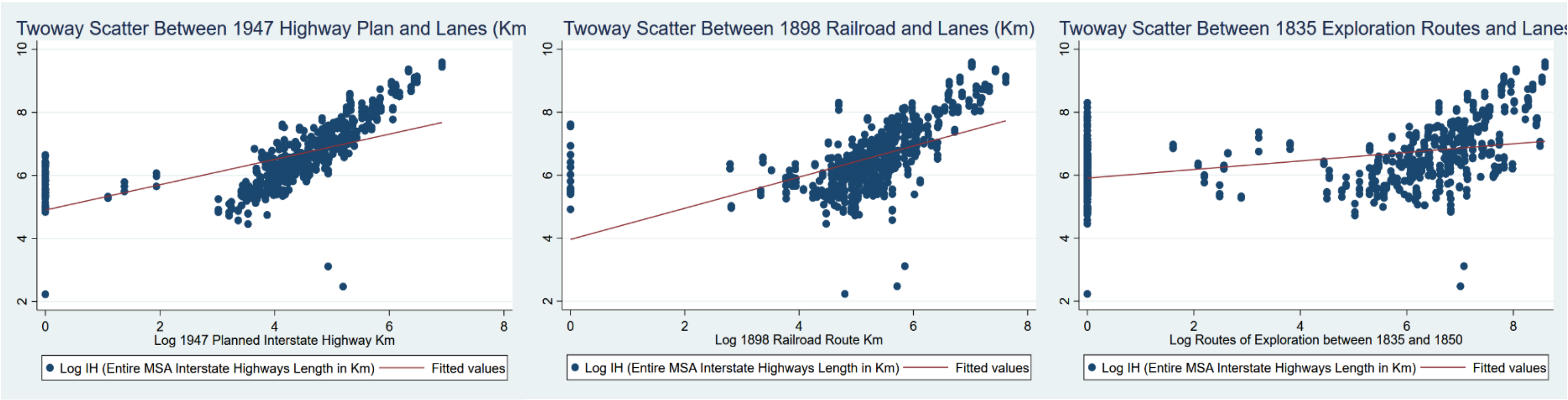
Three sets of control variables will also be included within the model to mitigate omitted variable bias arising from differences among different MSAs. Firstly, the elevation range, terrain ruggedness, two climate variables, and the distribution of development within the MSA were measures used to control for geography (Duranton and Turner, 2011, pg. 2624). Secondly, nine dummy variables where each represented a census division within the US. Thirdly, socioeconomic controls focusing around demographic and population characteristics. Where the level of segregation, manufacturing employment, low income, mean income and some college education, measures as a share of the population, are variables included within demographics. Population characteristics controlled the MSAs long-run population growth using “decennial population variables from 1920 to 1980” (Duranton and Turner, 2011, pg. 2625).

In order to deal with the problem of endogeneity (specifically simultaneity), three instruments will be used for log interstate highway (IH), the main endogenous regressor. The first instrument is the 1947 plan of the interstate highway system. The plan focused on placing highways in strategic places that connected places with large populations, economic activities, interregional traffic demand and military establishments (Duranton and Turner, 2012, pg. 9). This instrument is measured by the log planned kilometres of interstate highway that passed through each MSA. While, the second instrument is simply the log kilometres of major railroads that passed through each MSA in 1898. On the other hand, the third instrument uses the major expedition routes from 1835 to 1850 and measures the intensity of exploration in each MSA (Duranton and Turner, 2011, pg. 2649).

Instrument Validity:

For an instrument to be valid it must satisfy three conditions. Firstly, it must have relevance, where the instrument is strongly correlated with the endogenous variable. Most highways from the 1947 highway plan were completed by 1980 and are still being used today. Therefore, we expect log 1947 highway plan to be highly correlated with log interstate highway (IH). Due to technological advancements in transportation and railroad tracks offering perfect conditions for automobile roads (in terms of levelling and grading), many railroad tracks were replaced by highways. Therefore, we expect log 1898 railroad routes to also be highly relevant. Additionally, most expedition routes now have modern highways following that same route (Duranton and Turner, 2011, pg. 2629). Thus, log 1835 exploration routes are also expected to be highly relevant. The relevancy of these instruments can be illustrated in Figure 2, but a more formal test will be conduced in the results section.

Figure 2: Relationship Between Interstate Highways (IH) and Instruments



The second condition for validity is exogeneity. Where the instrument must be as good as random, meaning it is uncorrelated with the error term (like omitted variable bias). Therefore, all three instruments could satisfy this condition if enough control variables are included. This will be addressed in the econometric model and result sections.

The third condition for validity is exclusion. Where the instrument can only affect the dependant variable through its effect on the endogenous regressor. It is difficult imagining that these three instruments have a direct effect on the count of MSA employment in trucking and warehousing. The main reason being that all these instruments occurred a long time ago and the purpose of these transportation networks was not to anticipate modern day trucking and warehousing employment levels (Duranton and Turner, 2011, pg. 2631). Therefore, these three instruments most likely fulfill the exclusion condition.

Econometric Model:

To estimate the relationship between all interstate highways and employment we will be using three different regression methods. Firstly, using ordinary least squares (OLS) where the estimating regression equation is shown below:

$$LogEmployment_{it} = \beta_1 + \beta_2 LogHighway_{it} + \beta_3 LogPopulation_{it} + \beta_x Controls_{it} + \varepsilon_{it}$$

Where *LogEmployment* represents the log of the entire MSA interstate highway (IH) while *LogHighway* represents the log count of MSA employment in trucking and warehousing. *Controls* is the three sets of control variables (geography, census divisions and socioeconomic). Furthermore, *i* and *t* represent indexing by MSAs and years respectively. The main problems with OLS are that coefficients will be biased if there is not enough control variables or presence simultaneity. Therefore, three OLS regressions where sets of controls will be slowly added in to mitigate omitted variable bias. However, our second regression method as shown below, uses fixed-effects that will be better at dealing with this bias:

$$LogEmployment_{it} = \beta_1 + \beta_2 LogHighway_{it} + \beta_3 LogPopulation_{it} + \beta_x Controls_{it} + \delta_i + \eta_{it}$$

Here δ_i represents MSA fixed-effects that will remove time-invariant city effects (Duranton and Turner, 2011, pg. 2620). However, this does not eliminate the MSA-year variation, η_{it}. While fixed-effects has its benefits over OLS, both methods fail to deal with simultaneity or reverse causality (identification issue). This is where the third method of using instrumental variable (IV) comes in to play as shown below:

$$LogHighway_{it} = \beta_1 + \beta_2 Log1947Plan_{it} + \beta_3 1898Railroad_{it} + \beta_4 Log1835Exploration_{it} + \beta_5 LogPopulation_{it} + \beta_x Controls_{it} + \varepsilon_{it}$$

$$LogEmployment_{it} = \beta_1 + \beta_2 Log\widehat{High}way_{it} + \beta_3 LogPopulation_{it} + \beta_x Controls_{it} + \varepsilon_{it}$$

Here a two staged regression is used where *Log1947Plan*, *1898Railroad*, and *Log1835Exploration* are the three instruments for *LogHighway*. The first stage creates $Log\widehat{High}way$ which represents the predicted log lane kilometers of interstate highways (IH). As mentioned, these instruments require relevancy, exogeneity and exclusion in order to be valid and will be tested in the results section.

Furthermore, we will be using both logs urbanized (IHU) and non-urbanized (IHNU) as our main explanatory variables (replacing all interstate highways IH) and will utilize the same OLS and fixed-effect specifications. However, IV will not be used as instruments cannot be applied for the two endogenous variables simultaneously.

Results:

Tables 2 shows the regression results using the three methods as described in the previous section. In Panel A, the main explanatory variable is log of the entire MSA interstate highway (IH). Columns 1 to 3 are the OLS estimates and we find significant positive relationships (all at the 99% confidence level). However, the OLS elasticity of trucking and warehousing employment relative to interstate highways (IH) seems rather low. Where a 10% increase in interstate highways (IH) leads to a 1.65% to 1.31% increase in trucking and warehousing employment. From Column 4, we see that using fixed effects resulted in a non-significant relationship. Columns 5 and 6 are the IV estimates and we also find significant positive relationships between employment and highways (at the 95% and 90% confidence levels). However, these estimates are larger than the OLS estimates. Where a 10% increase in interstate highways (IH) leads to a 2.82% to 2.5% increase in trucking and warehousing employment. Additionally, we that the F-statistic for both IVs is greater than 10. This indicates that the instruments are strong and therefore have relevance. The IV estimate in Column 5 have an overidentification p-value of 0.03. This means that at least one of the instruments is not exogenous and therefore fails the exogeneity condition. However, the IV estimate in Column 6 passes the overidentification test and thus the instruments here meet the requirement of exogeneity.

Table 2: Trucking/Warehousing Employment as a Function of Lane Km

Variable	(1) OLS	(2) OLS	(3) OLS	(4) Fixed Effects	(5) 2SLS	(6) 2SLS
Panel A. Dependant variable is log MSA trucking/warehousing employment as a function of all MSA interstate highways (IH)						
LogHighway(IH)	0.165***	0.156***	0.131***	0.00854	0.282**	0.250*
	(0.046)	(0.048)	(0.048)	(0.040)	(0.140)	(0.136)
LogPopulation	0.923***	0.942***	1.057***	0.751***	0.842***	0.867***
	(0.039)	(0.051)	(0.149)	(0.141)	(0.099)	(0.114)
Geography		Yes	Yes			Yes
Census divisions		Yes	Yes			Yes
Socioeconomic			Yes			Yes
MSA fixed effects				Yes		
R squared	0.84	0.86	0.89	0.73		
Overidentification P-value					0.03	0.34
First Stage Statistics					16.5	11.8

Panel B. Dependant variable is log MSA trucking/warehousing employment as a function of urban (IHU) and non-urban (IHNU) interstate highways				
LogUrbanHighway(IHU)	0.15***	0.084**	0.094**	0.031
	(0.040)	(0.038)	(0.042)	(0.057)
LogNonUrbanHighway(IHNU)	0.092**	0.12***	0.099**	0.065
	(0.039)	(0.04)	(0.038)	(0.058)
LogPopulation	0.80***	0.86***	1.16***	0.79***
	(0.05)	(0.06)	(0.18)	(0.16)
Geography		Yes	Yes	
Census divisions		Yes	Yes	
Socioeconomic			Yes	
MSA fixed effects				Yes
R squared	0.85	0.86	0.89	0.72

*, **, and *** denote statistical significance at 10%, 5%, and 1% levels

From Panel B, the main explanatory variables are logs urbanized (IHU) and non-urbanized (IHNU) MSA interstate highway lane kilometers. Columns 1 to 3 are the OLS estimates (similar to Panel A) and we find significant positive relationships (at the 99% and 95% confidence levels). We see that the OLS elasticity for both urbanized and non-urbanized highways are smaller when compared to Panel A. Furthermore, the fixed-effects from Column 4 are also statistically insignificant.

Conclusion:

In conclusion, we find from our main IV estimate (Table 2: Column 6) that transportation infrastructure does promote employment growth. Specifically, a 10% increase in interstate highways (IH) leads to a 2.5% increase in trucking and warehousing employment. This IV estimate meets the relevancy and exogeneity requirement for a valid instrument. However, one severe limitation of this paper comes from the third requirement for a valid instrument. While in theory the three instruments used should directly not affect the present-day trucking and warehousing employment levels and meet the exclusion requirement, this can not be actually tested. Therefore, there is still the potential of endogeneity (omitted variables, simultaneity, reverse causality) which will cause biases in the coefficients. Aside from this limitation it does seem that building highways is a viable solution to improving places with low economic performance or growth. However, further study can be done based on this result and topic. For example, rather than using elasticity, a cost-benefit analysis can also be done in terms of its net contribution to the economy, but also the environmental impacts arising from actually building the highways and emissions from the increased economic activities.

References

Downs, Anthony. 1962. "The Law of Peak-Hour Expressway Congestion." *Traffic Quarterly*, 16(3): 393–409.

Duranton, G., & Turner, M. A. (2011). The fundamental law of road congestion: Evidence from US cities. *American Economic Review*, *101*(6), 2616-52.

Duranton, G., & Turner, M. A. (2012). Urban growth and transportation. *Review of Economic Studies*, *79*(4), 1407-1440.

Hymel, K. (2009). Does traffic congestion reduce employment growth?. *Journal of Urban Economics*, *65*(2), 127-135.

Jiwattanakulpaisarn, P., Noland, R. B., Graham, D. J., & Polak, J. W. (2009). Highway infrastructure and state-level employment: A causal spatial analysis. *Papers in Regional Science*, *88*(1), 133-159.

Kim, H., Lee, D. H., Koo, J. D., Park, H. S., & Lee, J. G. (2012). The direct employment impact analysis of highway construction investments. *KSCE Journal of Civil Engineering*, *16*(6), 958-966.

Kim, J. Y., & Han, J. H. (2016). Straw effects of new highway construction on local population and employment growth. *Habitat International*, *53*, 123-132.

Michaels, G. (2008). The effect of trade on the demand for skill: Evidence from the interstate highway system. *The Review of Economics and Statistics*, *90*(4), 683-701.