

# Term Paper: Effects of Transit Spending on Unemployment

Eric Cheung

Department of Economics, Simon Fraser University

ECON 435: Econometrics Methods

Dr. Xiaoting Sun

December 2, 2025

## 1. Introduction

Public transit decreases traffic and is more environmentally friendly when compared to driving. It also helps non-car owning workers get to their jobs that are not in walking distance. But what is the actual effect on unemployment when access to public transit is increased? This paper will attempt to quantify this relationship between public transit and unemployment by using time-series data that tracks monthly infrastructure spending, transit ridership and other variables. A static time-series OLS regression model will be applied on log-differenced variables with heteroskedasticity robust standard errors reported. There will be an additional model that uses quarterly compiled data that will include real GDP as a control variable. The results suggest that transit infrastructure spending does not lead to a direct effect on unemployment however certain modes of transportation do seem to have some effect.

## 2. Literature Review

Current literature does not directly the effects of transit on unemployment. There is much debate on how much should be spent on transit infrastructure specifically and who it benefits. Evidence shows that public transit serves to decrease congestion and helps disabled individuals especially. Additionally, there is literature on how other non-transit infrastructure spending has a positive effect on decreasing unemployment.

Nelson et al. (2007) talk about the congestion-reducing effect of transit. Using Washington D.C. as a case study they found that transit spending does not directly translate to economic value since there are many residents who do not take public transit (Nelson et al., 2007). Their conclusion is that governments need to find the optimal amount of spending to gain the benefits of transit without the wasteful spending (Nelson et al., 2007).

Leigh and Neill (2011) explore the relationship between infrastructure spending on unemployment in Australia. They used IV regression to estimate the effects of a national infrastructure on unemployment. The infrastructure program helps repairs old roads and the data is on an election district level (Leigh & Neill, 2011). They found that spending on infrastructure projects do have the potential to reduce unemployment (Leigh & Neill, 2011).

Overall, there is a lack of literature that directly studies the effect of transit on unemployment. However, there are more studies that focus on infrastructure spending. The papers that investigate infrastructure expenditure uses cross-sectional data in their models, and it leads to examining shorter term effects. This paper will use a time-series approach which differs from others, but it will also focus on transit through infrastructure spending. More specifically the main variable used in this paper will be construction spending on mass transit projects.

## 3. Data Source & Collection

- Table 1, Summary statistics for main variables

Main Variables Monthly	Mean	SD
St. and Local Govt Construction Spending - Mass Transit (\$)	602,145,749	214,205,402
St. and Local Govt Construction Spending - Land Passenger Terminal (\$)	214,558,704	94,644,755
St. and Local Govt Construction Spending - Transportation (\$)	2,473,016,194	703,166,138
Ridership - Other Transit Modes (#)	15,229,448	2,864,711
Ridership - Bus (#)	383,379,422	86,594,309
Ridership - Urban Rail (#)	345,500,491	81,816,905
Transportation Employment (#)	444,644.1	51,303.87
Unemployment (Rate)	0.0577085	0.02127128

The data is obtained from the US Bureau of Transportation Statistics' website (*Monthly Transportation Statistics*, 2025). The website compiles variables collected various American governmental agencies including the US Department of Transportation, the US Census Bureau, and the US Bureau of Labor Statistics. All variables are country-wide measure of monthly time series with the total dataset containing over 900 periods of observations dating back to 1946. However, this paper will only use 247 periods of observations because data from before said period contains missing observations. The usable data begins from January of 2005 and ends in May of 2025.

The dataset contains three variables tracking state and local government construction spending. The three types of construction spending include spending on mass transit, transportation, and on land passenger terminals. The construction spending variables are recorded in US dollars and is defined as the dollar value of construction work done in that category (*Monthly Transportation Statistics*, 2025). Also included in the data are three variables that track ridership for bus, urban rail, and other modes of transit. There is also a variable that records the number of individuals that are employed to work in a transit compacity including those that work in transit or in ground passenger transportation (*Monthly Transportation Statistics*, 2025). Not included in the dataset, but there will be a binary variable included in the regression models. The binary variable is called covid, and it indicates the observation occurred during the covid-19 pandemic that started in March 2020 and ended on May 2023.

The main independent variable of interest is the amount of state and local government construction spending on mass transit projects. This variable will serve as the measure for transit infrastructure spending. The main dependant variable of interest will be the national unemployment rate measured as a ratio.

- Table 2, Augmented Dickey-Fuller Test for Stationarity

Variable	ADF_Statistic	P_Value
State and Local Government Construction Spending - Mass Transit	-1.81	0.65
State and Local Government Construction Spending - Land Passenger Terminal	-1.85	0.64
State and Local Government Construction Spending - Transportation	-4.39	0.01
Transit Ridership - Other Transit Modes	-2.35	0.43
Transit Ridership - Fixed Route Bus	-2.22	0.48
Transit Ridership - Urban Rail	-2.44	0.39
Transportation Employment	-3.43	0.05
Unemployment Rate	-2.40	0.41

Plots (see Appendix) of nearly all the variables show that there are significant trends in the data making it highly unlikely that the variables are stationary. Further testing using the Augmented Dickey-Fuller Test from the “tseries” package confirms all the variables are non-stationary except transit employment. Therefore, a log-difference procedure is applied to the all the variables (Wooldridge, 2019). The variables are first logged which will help interpretation since the construction spending variables are in the range of hundreds of millions of dollars. Then the variables are differenced which will detrend and transform them into stationary variables. The plots (see Appendix) of the variables after this log-difference process shows significantly less trending and another use of the Augmented Dickey-Fuller Test indicates that they variables are now stationary.

- Table 3, Summary Statistics for Quarterly Variables

Main Variables Quarterly	Mean	SD
Unemployment Rate	0.05758835	0.02117682
Bus Ridership	382,731,304	84,473,298
Rail Ridership	345,250,643	79,868,750
Other Transit Modes Ridership	15,254,066	2,754,037
Transit Employment	444,464.5	47,754.3
Mass Transit Spending	1,770,595,238	667,362,496
Transportation Spending	7,271,845,238	2,225,739,203
Land Passenger Terminal Spending	630,904,762	289,902,128
Real GDP	19,021,497,512,195	2,359,178,304,169

As a robustness check, another model will be estimated that will include real GDP that will control for macroeconomic conditions. However, real GDP is only reported on a quarterly basis. Therefore,

this separate model will have the construction spending variables recorded as a sum of all the months in a quarter and the ridership and rate variables as an average of those variables in a quarter.

#### 4. Econometric Methodology

The model used will be a static time-series because the variables available are all stochastic processes that have been reported monthly. The coefficients of this static time-series model will be estimated via OLS. Methods such as IV, fixed/random effects, and other more advanced time-series models will not be used in this case. For instrumental variable regression, there is not a suitable instrumental variable in the data that can affect construction spending on transit projects without affecting unemployment. As for panel data methods such as fixed/random effects model, the available does not involve observations from different countries/jurisdictions over multiple periods of time as it only records monthly observations from the whole of the US.

- Full Model Specification

$$\begin{aligned}\Delta \ln(\text{UnemployRate}) &= \beta_0 + \beta_1 \Delta \ln(\text{MassTransit}) + \beta_2 \Delta \ln(\text{LandPassTerm}) + \beta_3 \Delta \ln(\text{Transport}) \\ &+ \beta_4 \text{covid} + \Delta \beta_5 \ln(\text{OtherTransMod}) + \beta_6 \Delta \ln(\text{Bus}) + \beta_7 \Delta \ln(\text{Rail}) \\ &+ \beta_8 \Delta \ln(\text{TransEmpty})\end{aligned}$$

The interpretation of the model is made easy as the dependent variable unemployment rate is logged and most of the independent variables are logged as well. This can be interpreted as a 1% change in the independent variable will lead to a percent change in the unemployment equal to the result of the coefficient (Wooldridge, 2019).

There are assumptions for the static time-series model that when are correct will ensure that the estimator is consistent and normally distributed. One assumption is that the variables are stationary and ergodic (Sun, 2025). Stationary meaning that its joint distribution is the same across different time periods (Sun, 2025). And ergodic meaning that observations are less correlated as the time periods between them increases (Sun, 2025). The raw variables are non-stationary as confirmed by the earlier plots and by the Augmented Dickey-Fuller tests. However, after taking their log-differences they have been transformed into stationary variables. This is again confirmed by the plots and another round of the Augmented Dick-Fuller test. As for the ergodic assumption, it cannot be directly tested which means it can only be assumed to be ergodic. But it is known that if a time-series is stationary it can only help in the assumption of ergodicity.

- Table 4, Augmented Dickey-Fuller Test for Stationarity

Variable	ADF_Statistic	P_Value
$\Delta \ln(\text{MassTransit})$	-8.66	0.01
$\Delta \ln(\text{LanPassTerm})$	-6.89	0.01
$\Delta \ln(\text{Transportation})$	-13.34	0.01
$\Delta \ln(\text{OtherTransMods})$	-7.45	0.01
$\Delta \ln(\text{Bus})$	-6.01	0.01
$\Delta \ln(\text{Rail})$	-6.53	0.01
$\Delta \ln(\text{TransEmpty})$	-8.71	0.01
$\Delta \ln(\text{UnemployRate})$	-6.41	0.01

Another assumption for static time-series models is non-collinearity (Sun, 2025). This assumption means that the regressors have no perfect collinearity. Or in other words, independent variables should not be a singular value, not a transformation, or not a perfect linear combination of another variable (Wooldridge, 2019). Non-collinearity ensures that the matrix representing the data will be invertible (Sun, 2025). The covariance matrix calculated from the data showed that there is no perfect collinearity in the variables.

- Table 5, Covariance-Matrix, All Variables are Log-Differenced

	MassTransit	LanPassTerm	Transportation	OtherTranMods	BusRouts	UrbanRail	Tran_Employ	UnempRate
MassTransit	1.000	0.083	0.563	0.190	0.153	0.098	-0.020	-0.006
LanPassTerm	0.083	1.000	0.421	0.250	0.240	0.182	0.036	-0.034
Transportation	0.563	0.421	1.000	0.305	0.223	0.156	-0.072	0.039
OtherTranMods	0.190	0.250	0.305	1.000	0.848	0.864	-0.021	-0.617
BusRouts	0.153	0.240	0.223	0.848	1.000	0.893	0.369	-0.634
UrbanRail	0.098	0.182	0.156	0.864	0.893	1.000	0.363	-0.802
Tran_Employ	-0.020	0.036	-0.072	-0.021	0.369	0.363	1.000	-0.365
UnempRate	-0.006	-0.034	0.039	-0.617	-0.634	-0.802	-0.365	1.000

The next assumption for this static time-series model is predetermined regressors and no contemporaneous correlation (Sun, 2025). This means that the independent variables cannot be affected by shocks from the same period. Even though this assumption cannot be directly tested, it can be generally assumed that the variables in the data do not generally change immediately to shocks. This is especially true for the construction spending variables which are part of governmental budgets that must go through planning and proposal phases. Other variables such as ridership, transit employment, and labor force participation should not react to shocks so easily.

The last assumption for this model is that the product of the regressors and the error terms is a martingale difference sequence (Sun, 2025). This assumption also cannot be directly confirmed however, since it is safe to assume that past shocks cannot be used to predict current shocks, the product can be assumed to be a martingale difference sequence.

- Table 6, Breusch-Pagan Test for Heteroskedasticity

Model	BP_Statistic	P_Value
TS OLS 1	8.65	0.07
TS OLS 2	135.51	0.00
TS OLS 3	134.95	0.00

The assumptions stated above regarding the static time-series model ensures that the estimators are consistent and unbiased (Sun, 2025). However, the standard errors of the estimators can be biased if there is heteroskedasticity in the errors (Wooldridge, 2019). This can be tested formally with the Breusch-Pagan Test from the “lmtest” package (Wooldridge, 2019). The low p-values from the Breusch-Pagan indicates that the model contains heteroskedasticity. Since there is heteroskedasticity in the static time-series models, heteroskedastic corrected Newey-West standard errors will be reported instead.

## 5. Results and Interpretation

The first model contains the three construction spending variables which are mass transit, land passenger terminal, and general transportation. The coefficient for the main variable mass transit suggests that increase 1% increase in construction spending leads to a 0.044 %. decrease in unemployment. The transportation construction spending causes a slight increase in unemployment while the land passenger terminal decreases unemployment as well. This effect of mass transit spending decreasing unemployment aligns economically. However, the heteroskedastic robust standard error reported for this model suggests that none of the effects are significant.

- Table 7, Regression Results with Monthly Data

Main Models	TS OLS 1	TS OLS 2	TS OLS 3
(Intercept)	-0.002 (0.003)	-0.001 (0.002)	-0.001 (0.003)
$\Delta \ln(\text{MassTransit})$	-0.044 (0.039)	-0.025 (0.031)	-0.022 (0.030)

Main Models	TS OLS 1	TS OLS 2	TS OLS 3
$\Delta \ln(\text{Transportation})$	0.118 (0.094)	0.147** (0.055)	0.152** (0.058)
$\Delta \ln(\text{LandPassTerm})$	-0.051 (0.054)	0.022 (0.024)	0.025 (0.025)
covid	0.003 (0.033)	-0.003 (0.008)	-0.003 (0.008)
$\Delta \ln(\text{OtherTransMods})$		0.108 (0.068)	-0.035 (0.116)
$\Delta \ln(\text{Bus})$		0.295** (0.094)	0.372** (0.135)
$\Delta \ln(\text{Rail})$		-0.776*** (0.159)	-0.707*** (0.130)
$\Delta \ln(\text{TransEmpty})$			-0.118 (0.089)
Num.Obs.	246	246	246

• + p < 0.1, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

The second model adds the ridership variables which include, bus, urban, and other modes of transportation. The coefficient for the main variable suggests that a 1% increase in construction spending leads to a 0.025 % decrease in the unemployment. The new coefficients suggests that bus ridership increases unemployment, and urban rail decreases unemployment. The heteroskedastic robust errors suggests that the mass transit coefficient is not significant while the bus and rail ridership coefficients are significant. The effects of ridership do perhaps suggest further investigation is necessary.

The full model adds transit employment variable to the previously mentioned variables. This model suggests that a 1% increase in construction spending on mass transit leads to a 0.022% decrease in unemployment rate. The added variables suggests that transit employment decreases unemployment. However, the heteroskedastic robust errors suggests that only the labour force participation is significant. The new coefficients added in this model aligns economically.

- Table 8, Regression Results with Quarterly Data

Secondary Model	Quarterly TS OLS
(Intercept)	0.011 (0.019)
$\Delta \ln(\text{MassTransit\_Qtrly})$	0.099 (0.120)

Secondary Model	Quarterly TS OLS
$\Delta \ln(\text{Transportation\_Qtrly})$	0.141 (0.121)
$\Delta \ln(\text{LandPassTerm\_Qtrly})$	-0.044 (0.073)
$\Delta \ln(\text{Bus\_Qtrly})$	0.680+ (0.371)
$\Delta \ln(\text{Rail\_Qtrly})$	-0.867** (0.254)
$\Delta \ln(\text{OtherTransMod\_Qtrly})$	0.047 (0.249)
$\Delta \ln(\text{TransEmpty\_Qtrly})$	0.081 (0.225)
$\Delta \ln(\text{rGDP\_Qtrly})$	-2.867 (2.361)
covid	-0.004 (0.016)
Num.Obs.	80

• + p < 0.1, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

The static time-series model computer with quarterly variables with real GDP include suggest that a 1% change in mass transit construction spending increases unemployment by 0.099% however most of the variables are not reported as significant. Apart from urban ridership with does decrease unemployment in a significant manner.

The non-significant effect of mass transit construction spending on unemployment in all three models along with the quarterly model suggests that the results are robust. There is not enough evidence in national level data to signify that mass transit spending by itself has any major impact on affecting unemployment. However, this does not suggest that increase mass transit does not helpful in other ways that are not represented in this paper. Notably in two of the three models, construction spending on general transportation projects and bus ridership increases unemployment in a statistically significant manner. These results do not align with expected results from an economic perspective or because the data available does not allow the relationship to be properly modeled.

Further study of this topic conducted with data on specific state or municipality should yield more conclusive results. There are many factors such as geography, demographics, and transit funding structures that vary significantly across different states and municipalities.

## 6. Conclusion

From the time-series regression results, it appears that infrastructure spending on transportation does not have a significant effect on unemployment. Even with the control variables of covid and quarterly real GDP, the heteroskedasticity robust errors suggest that the effects are non-significant. The effect of rail ridership is significant across the quarterly and monthly models which suggest that one specific part of public transit does have a positive effect in decreasing unemployment.

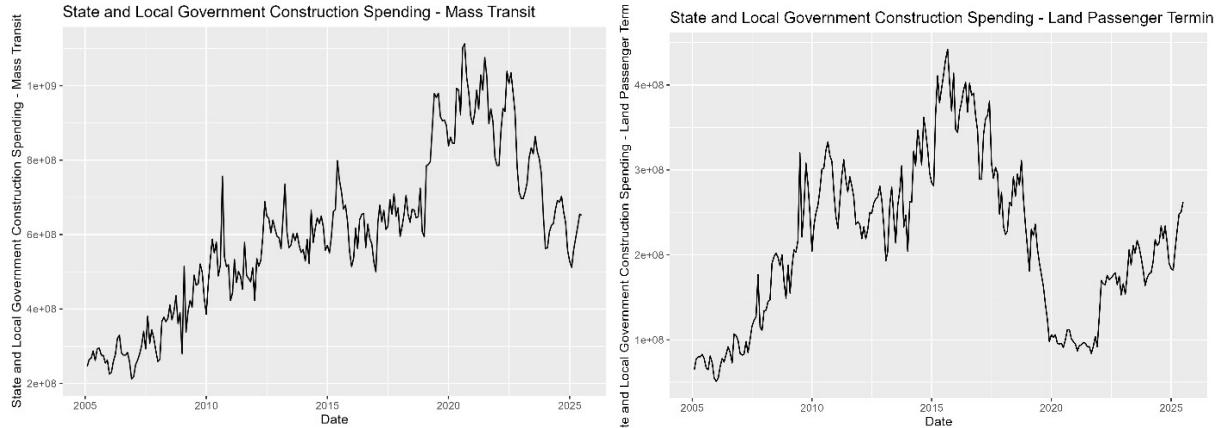
However, these results are derived from aggregated data from the whole of the United States. Further estimates with more granular data such as those at the state or municipal level may uncover different results. It could also be the result of omitted variable bias that is causing the results to be unreliable as there are many other immeasurable factors that affect unemployment.

## References

- Leigh, A., & Neill, C. (2011). Can national infrastructure spending reduce local unemployment? Evidence from an Australian roads program. *Economics Letters*, 113(2), 150–153.  
<https://doi.org/10.1016/j.econlet.2011.05.037>
- Monthly transportation statistics*. (2025). [Dataset]. Bureau of Transportation Statistics.  
[https://data.bts.gov/Research-and-Statistics/Monthly-Transportation-Statistics/crem-w557/about\\_data](https://data.bts.gov/Research-and-Statistics/Monthly-Transportation-Statistics/crem-w557/about_data)
- Nelson, P., Baglino, A., Harrington, W., Safirova, E., & Lipman, A. (2007). Transit in Washington, DC: Current benefits and optimal level of provision. *Journal of Urban Economics*, 62(2), 231–251. <https://doi.org/10.1016/j.jue.2007.02.001>
- Sun, X. (2025). *ECON 435 lecture Slides* [Slide show].
- Wooldridge, J. M. (2019). *Introductory Econometrics: a Modern Approach*. Cengage Learning.

## Appendix

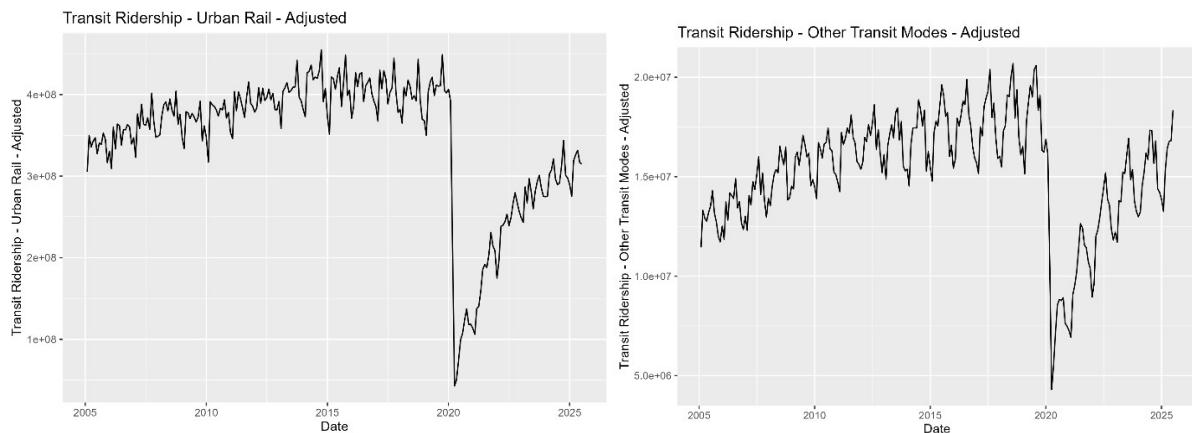
### Pre Log Differencing



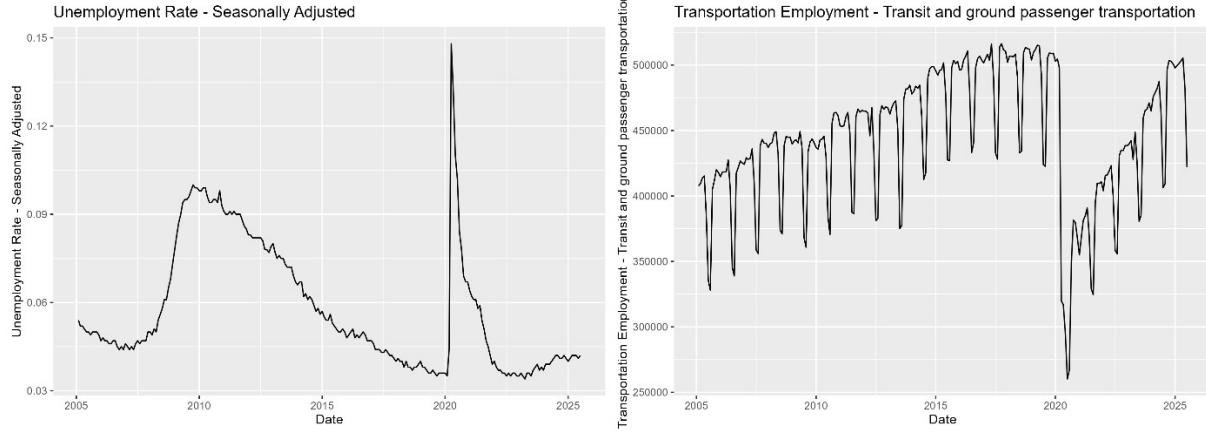
Figures A.1 & A.2



Figures A.3 & A.4

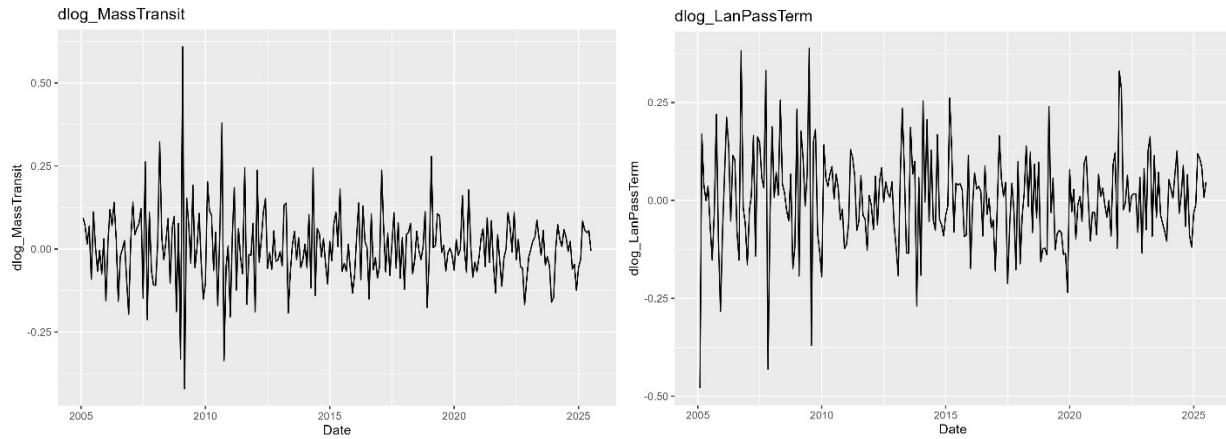


Figures A.5 & A.6

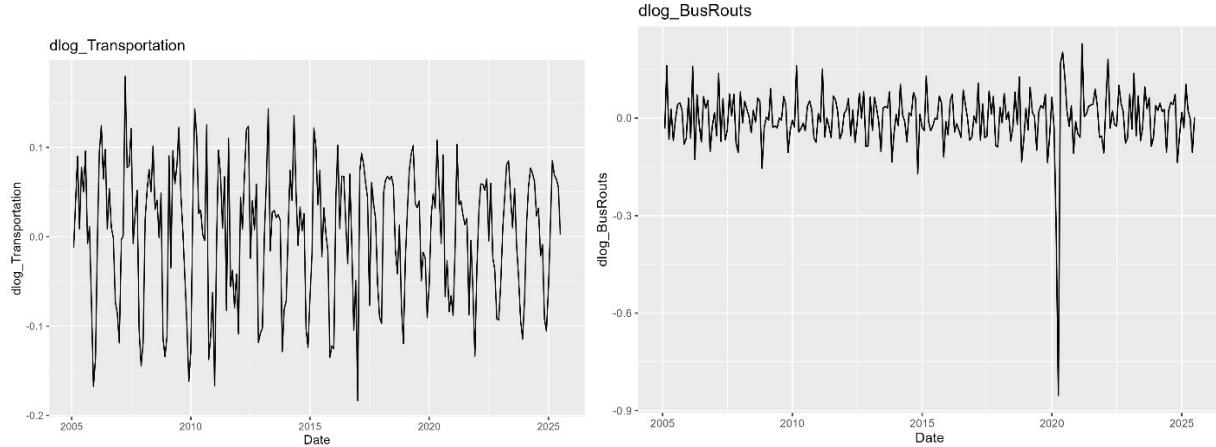


Figures A.7 &amp; A.8

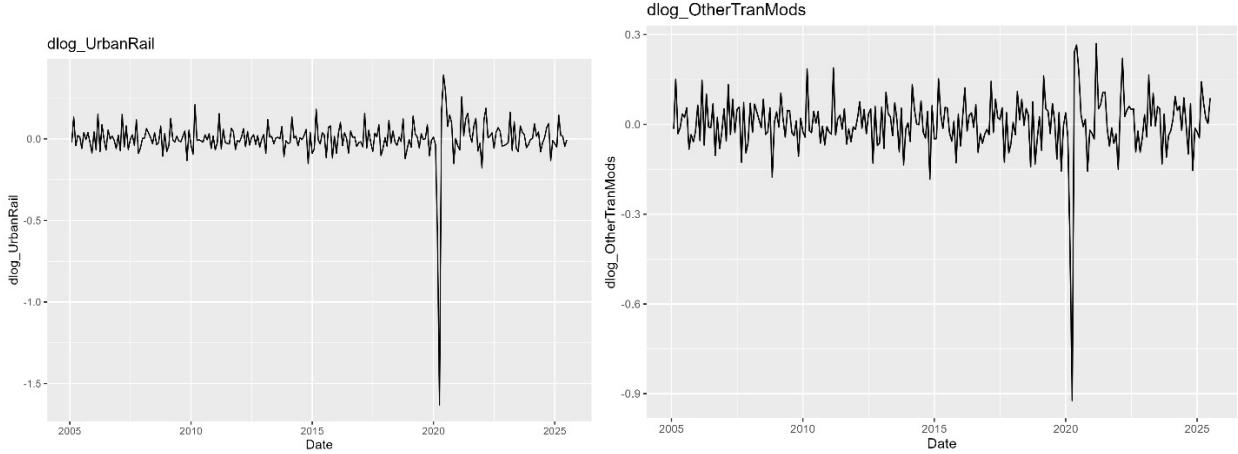
After Log-Differencing



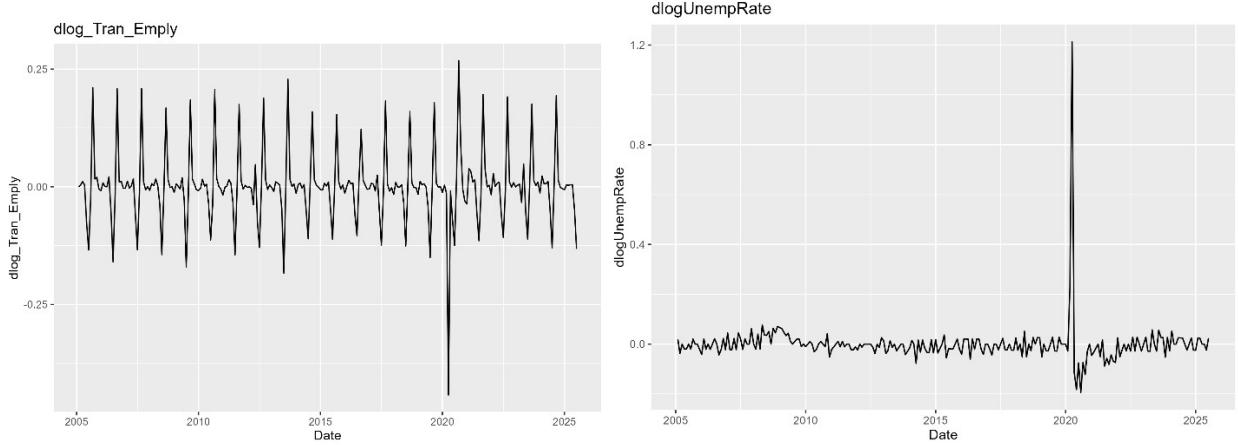
Figures A.9 &amp; A.10



Figures A.11 &amp; A.12



Figures A.13 & A.14



Figures A.15 & A.16