Literature Review

Transportation investments boost productivity by reducing travel times and shipping costs, benefiting businesses through lower production costs and increased competitiveness (Banister and Berechman, 2001). These cost reductions can range from 10% to 30%, varying by industry and region.

Large-scale transportation projects also have immediate positive effects on local employment. Weisbrod and Reno (2009) found that every $1 billion spent on transportation infrastructure creates about 47,500 jobs in the construction sector alone. The broader economic impact, including indirect job creation and increased economic activity, can be even more substantial. Additionally, regions with significant transportation investments experience a 16% higher income growth rate (Chatman and Noland, 2011).

In the long term, transportation infrastructure supports sustained economic growth by promoting agglomeration economies and improving overall economic efficiency. Aschauer (1989) noted that a 1% increase in infrastructure capital can lead to a 0.4% increase in GDP. Duranton and Turner (2012) found that a 10% increase in interstate highway density within metropolitan areas can result in a 1.5% increase in employment over a decade.

Methodology

Data Sources and Preparation

The first component of the Jupyter Notebook analysis utilizes four primary datasets obtained from Statista, the Census Bureau, and the U.S. Department of Transportation. The first dataset showed 2022 expenditure and revenue metrics aggregated by spending type for each state. The second data set includes the 2022 poverty rate for each state, along with statistical significance indicators for differences in poverty levels between states at the 90% confidence level. The third data set shows the 2022 Gini Coefficient for every state. The Gini Coefficient is a ratio that represents the level of income inequality within a society, with a coefficient of 1 representing perfect inequality and a coefficient of 0 representing perfect equality. The fourth data set is a data set that shows 2022 unemployment levels for each state.

The datasets were cleaned using the pandas library, visualized using matplotlib and seaborn, and analyzed using scikit.learn and statsmodels.api. The four above data sets were merged, with statistical significance indicators removed from the poverty level sub-data frame and all expenditure variables being removed except overall expenditures and different transportation expenditures.

The second component of the Jupyter Notebook analysis included looking at historical data sets from 2017-2021 for Real GDP and transportation expenditures by state.

Analytical Approach

First, scatter plots were created between the transportation spending levels and the different economic indicators. Least Squares Regression models were then used, with the four economic

indicators being regressed on the percentage of expenditures used on highways, public transit, and airports (2022 data). Then, r-squared metrics and coefficients were calculated to determine the fit of the models.

For the second part of the analysis, historical data from 2017, 2018, 2019, and 2021 was analyzed to determine a relationship between the proportion of transportation spending and Real GDP, due to that being the most promising relationship after running the first analysis. Then, more robust linear regression models were run, which split each data set into training and testing data sets with an 80-20 split. Afterwards, r-squared and root mean-squared error values were calculated for both the training and testing data.

Results

Analysis of Coefficients and R-squared Values

For the first four models, which used transportation spending metrics to predict Real GDP, inequality, poverty rates, and unemployment rates, a statistically significant F-Statistic was obtained for all except the poverty rate model. This indicated that transportation spending metrics aided in explaining the variation in Real GDP, inequality, and unemployment between states. The most significant coefficients in the models were the 1.099 * 10 ^5 coefficient on Percent_Transit for predicting Real GDP, the -0.2986 coefficient on Percent_Highways for predicting poverty rates, and the -0.1312 coefficient on Percent_Highways for predicting unemployment rates.

For the last four models, which used transportation spending metrics to predict Real GDP fro 2017, 2018, 2019, and 2021, moderately high r-squared values ranging from 0.4 - 0.42 were generated, indicating that that our original r-squared value of 0.56 was potentially an overestimate. However, the relationship remains significant. Furthermore, it is important to note that these r-squared values were obtained using a test-train split and stayed relatively constant when comparing metrics obtained with training data vs test data. This underscores how it is unlikely that this relationship is influenced heavily by outliers and is a sign that our model isn't overfitting our data.

Discussion

How does transportation contribute to increased productivity and competitiveness?

The first part of the analysis showed a significant relationship between transportation spending and Real GDP, with a particularly high coefficient for Percent_Transit. The extended analysis with historical data (2017-2021) further validates this relationship, though with slightly lower r-squared values. The significant positive impact of transportation spending, especially on transit, on Real GDP suggests that such investments enhance productivity and competitiveness, likely by reducing travel times and costs, as indicated in the literature (Banister and Berechman, 2001).

How do large-scale transportation projects influence local employment rates and income levels, and what are the long-term effects on economic growth?

The analysis aligns with Weisbrod and Reno (2009), indicating that transportation investments create jobs and stimulate income growth. Negative coefficients on Percent_Highways for poverty rates and unemployment rates indicate that higher spending on highways may contribute to reducing these metrics. This suggests that such investments could have broader social benefits by improving access to employment opportunities and stimulating economic activities in less affluent areas.