Transportation and Economic Development: The Link Between Highway Investment and Economic Development—A Time-Series Investigation

Final Report: Appendix IV
The report written by Prof. Yorgos J. Stephanedes contains nine volumes. Copies of the report may be obtained in its entirety or by separate volume. The title of each volume is as follows:

1. TRANSPORTATION AND ECONOMIC DEVELOPMENT
   Final Report - Executive Summary

2. TRANSPORTATION AND ECONOMIC DEVELOPMENT
   Final Report

3. TRANSPORTATION AND ECONOMIC DEVELOPMENT: THE GEOGRAPHICAL LITERATURE
   Final Report - Appendix I

4. TRANSPORTATION AND ECONOMIC DEVELOPMENT: TRANSPORTATION AND THE MINNESOTA ECONOMY; TRANSPORTATION/ECONOMY LITERATURE
   Final Report - Appendix II

5. TRANSPORTATION AND ECONOMIC DEVELOPMENT: EVALUATING CRITERIA FOR HIGHWAY PROJECT SELECTION
   Final Report - Appendix III

6. TRANSPORTATION AND ECONOMIC DEVELOPMENT: THE LINK BETWEEN HIGHWAY INVESTMENT AND ECONOMIC DEVELOPMENT - A TIME-SERIES INVESTIGATION
   Final Report - Appendix IV

7. TRANSPORTATION AND ECONOMIC DEVELOPMENT: THE LINK BETWEEN HIGHWAY INVESTMENT AND ECONOMIC DEVELOPMENT: SPECIFIC ECONOMIC SECTORS
   Final Report - Appendix V

8. TRANSPORTATION AND ECONOMIC DEVELOPMENT: HEURISTIC DECISION FRAMEWORK FOR UPGRADING HIGHWAY WEIGHT LIMITS
   Final Report - Appendix VI

9. TRANSPORTATION AND ECONOMIC DEVELOPMENT: SIMULATION OF HIGHWAY INVESTMENT IMPACTS ON THE FORESTRY SECTOR IN NORTHEAST MINNESOTA
   Final Report - Appendix VII

Additional copies may be obtained by writing to:

Minnesota Department of Transportation
Research Administration & Development Section
Materials & Research Laboratory
1400 Gervais Ave.
Maplewood, MN 55109
A time series methodology is developed that differentiates the effects of highways on development from the effects of development on highways. This methodology uses pooled time-series and cross-sectional data on highway expenditures and county employment for the 87 Minnesota counties and all 9 economic sectors over the 25-year period 1957-1982 and includes classification of counties based on access, demographic and socioeconomic features. Results from vector autoregressions are tested against modern causality tests of Granger-Sims type. In the wholesale and natural-resource-based service sectors (e.g., tourism), increased highway expenditures result in long-term employment increases. While regionally very substantial, the impacts are distributional, i.e., the statewide impact is negligible. Government role is mostly reactive, increasing funding to counties whose economy is increasing, except in rural areas where government also attempts to stimulate declining economies. Funding decisions are highly sensitive to changes in the economy, especially in rural areas, and (as our evaluation of the Minnesota Department of Transportation [Mn/DOT] project selection process indicates) are primarily influenced by the District recommendation. Further, a new B/C project selection process is developed and tested on highway weight restriction policies in Northeast Minnesota. Both simulation with large I/O model and comparison with actual funding decisions made independently by Mn/DOT indicate agreement with our results. An extensive literature review and 175 references are included.

This report consists of nine separate publications: an executive summary, the final report and seven appendices. The publications are listed on the following page.
INTRODUCTION

Possible economic effects of highways influence highway funding decisions either directly through stated objectives or indirectly through the political arena. For example, DOTs in 36 states explicitly consider regional economic development in their highway program selection (Forkenbrock and Plazah, 1986). This paper seeks to determine whether highway projects have a definite and foreseeable effect on economic development, increasing jobs above the normal trend of the economy. If highway projects lack such an effect, then some funds may be inefficiently allocated. On the other hand, if highway projects significantly affect economic development and if the aim is to stimulate the economy, more use of highway funds for economic development purposes may be justified.

Although transportation historically has had undeniable effects on economic development by "opening up the frontier," some studies indicate that now that our highway system is mature, additional highway improvements in transportation have little, if any, effect on economic development. Unfortunately, the existing empirical evidence uses cross-sectional or input-output analysis and is mixed and inconclusive. The purpose of our research is to address the causality issue more directly by using time-series statistical techniques instead of cross-sectional techniques.

Our analysis uses annual data on highway expenditures and employment for all 87 Minnesota counties from 1964 to 1982. While we pool cross-sectional data with time-series data, we analyze the time-series aspects of the data. By first
paneling the data, we remove the cross-sectional element for each county. Causality tests (Granger-Sims type) are used to test whether highway expenditures affect employment levels and whether employment levels affect highway expenditures. Vector autoregressions supplement these causality tests by quantifying the dynamics associated with these relationships.

We conclude that the evidence strongly indicates that higher employment levels lead to higher highway expenditures but that causality in the opposite direction is weak. However, when we separately analyze the counties that are economic centers of the state (see Regional Center counties defined below), the evidence indicates that higher highway expenditures in these counties lead to a statistically significant increase in employment levels, above the normal trend of the economy.

LITERATURE REVIEW

Historically, transportation had undeniable effects on economic development. The location of communities was often determined by the location of transportation be it a river or railroad. However, today's mature highway network provides a high degree of accessibility relative to what existed a hundred years ago. Thus, today's highway projects may lack the stimulative economic effects experienced as our country was being developed.

Possible ways that highways may be able to affect economic development include:

(a) **Residential Location:** In response to changes in the transportation infrastructure, people may change their residences to take advantage of the new transportation facilities. In urban areas this effect has been well studied and its existence verified (Mudge, 1974).
(b) **Work Place Location**: A new transportation facility may enable people to work far from where they reside (Stephanedes & Eagle, 1982).

(c) **Enterprise Location resulting from change in labor supply**: Stephanedes and Eagle (1983) argue that if new transportation facilities allow people to participate in the labor market of an area to which they previously lacked accessibility, then that area's labor supply may increase. The increased labor supply may attract new industry to the area.

(d) **Enterprise Location resulting from decreased transportation costs**: An improvement in transportation often will decrease the transportation costs of companies in the area served by the transportation facility. These decreased costs may attract new firms to the area (Gamble, et al., 1978).

The work we do in this paper tests the validity of points (c) and (d) - whether changes in the transportation network will affect enterprise location or expansion. We will, however, measure enterprise location or expansion with employment levels in the counties where highway changes take place.

Several investigators have studied the effects of interstate highways on population and employment growth (Humphrey & Sell, 1975; Lichter & Fuguitt, 1980; Briggs, 1981). These investigators have found that counties with interstate highways do have an advantage over other counties with regard to population and employment growth but only in counties within 25 miles of a metropolitan area. The effects on employment are primarily related to industries servicing those using the highways (e.g. service stations, restaurants, motels, etc.) and are not related to manufacturing or wholesale operations. Research in the Atlantic Region of Canada (which includes New
Brunswick, Nova Scotia, Prince Edward Island, and Newfoundland) concluded that increased investment in transportation infrastructure and freight subsidies would attract few industries because "a reasonably mature transportation system [is] properly in place and maintained" (Wilson, et.al., 1982). Similarly, a study of the region around the Ozark Plateau in Arkansas concluded that little correlation exists between highways and economic development (Kuehn & West, 1971).

Other sources have found significant relationships existing between highways and economic growth. Expressway investments in North England were found to lead to greater regional employment growth (Dodgson, 1974), although this greater growth was minimal. In Connecticut, manufacturing employment and population increased more in towns close to the new Turnpike than in towns further away (McKain, 1965; Gaegler, et.al., 1979).

Regional economic forecasting and policy analysis using large-scale regional models (such as Harris, 1973), often based on the input-output method (Roberts, & Kresge, 1968; Amano & Fujita, 1970; Sakashita, 1974; Polenske & Levy, 1975; Maki, et.al., 1978; Stevens, et.al, 1982), have indicated that some economic effects do result from changes in transportation. The implications of these models, however, often depend critically on the users' assumptions. An important variable in input-output models is market share, the amount of the total final demand that is produced locally. In some models, such as SIMLAB (Maki, et.al., 1978), the user determines market share. Other models, such as the Amherst Model (Stevens, et.al., 1982), estimate market share using an equation of other variables of the model.

When models estimate market share, the parameters of the estimation are often based on cross-sectional statistical studies and therefore merely represent correlation. However, by their very existence within the large-scale
models, these parameters are used as if they did represent a causal relationship. The direction and nature of that relationship often follows from an ad hoc model structure. For example, modeler A may write market share as a function of the number of highways in the region relative to the rest of the country. Modeler B may write the number of highways in the region as a function of the market share. Both modelers use the same data and cross sectional analysis. The correlation they obtain is therefore the same, but the two modelers will interpret that correlation as different causal relationships.

Most of the studies we have discussed use cross-sectional techniques. However, time-series techniques address the issue of causality more directly than do cross-sectional techniques. Time-series techniques can test whether changes in one series (such as highway expenditures) statistically precede changes in another series (such as employment levels). Thus, we have undertaken a time-series investigation into the causal links between highway investment and economic development.

DATA

This section briefly describes the data, the groupings of counties we wish to analyze separately, and our normalization of the data to filter out the effects of inflation, regional or national trends, and other factors common to each particular grouping. The data consist of annual observations of state highway system construction expenditures and employment for all 87 Minnesota counties. The expenditure data, provided by the Minnesota Department of Transportation, are broken down by county for the fiscal years 1957-1982. The employment data from the County Business Patterns (U.S. Dept. of Commerce) for the years 1964-1982 represent the employment in the middle of March of each year. (These data do not include self-employed people, railroad workers, or
governmental employees.) Because the majority of each fiscal year's highway budget (a fiscal year is from July 1 of one year to June 30 of the next) is spent before March, we view the expenditure data for each year as preceding the employment data.

The length of each county's employment time series is 19 years; the length of each county's highway-expenditures time series is 26 years. Traditional time-series analysis could do very little with such short series. However, by pooling the cross-sectional data with the time-series data, our data elements will increase from 19 to 1653 (19 years x 87 counties) for employment and from 26 to 2262 (26 years x 87 counties) for highway expenditures.

Prior to the analysis, we wish to filter out changes reflecting regional or national trends, inflation, and other effects that are common to the grouping of counties under study. To accomplish this filtering, we define our variables for the statistical analysis as follows: Let $x_{i,t}$ be the basic variable (such as expenditures or employment) for county $i$ in year $t$. Then,

$$^\wedge x_{i,t} = \frac{x_{i,t}}{\sum_{j \text{ in } G} x_{j,t}}$$

where $G$ is the grouping of counties we are considering. The newly defined $x_{i,t}$ variables compare each county to the total of the counties in the grouping.

The groupings of counties that we use are defined in figure 1. These groupings are:

(a) **Statewide**: All 87 Minnesota counties
FIGURE 1. COUNTY GROUPINGS

REGIONAL CENTERS = □□

URBAN = □ + □□

NEXT TO REGIONAL CENTERS = □□ + □□

NEXT TO URBAN = □□

RURAL = □

Source: Upper Midwest Council
FIGURE 2. MINNESOTA COUNTIES

Source: Upper Midwest Council
(b) **Urban**: Counties in the Twin Cities seven-county metropolitan area and counties containing a city of population 28,000 or larger. (We have followed this definition strictly. Thus, even St. Louis county which includes the City of Duluth and a very large rural area is classified as an Urban county.)

(c) **Next to Urban**: Counties bordering the counties in (b)

(d) **Regional Center counties**: In the Twin Cities metropolitan area, these counties include Hennepin and Ramsey counties, which include the cities of Minneapolis and St. Paul. Outside the metropolitan area, counties are included if they contain a city the size of Mankato (located in south Minnesota and having a population of 28,000) or larger (Alternatively: counties in (b) minus the Twin Cities metropolitan counties other than Hennepin and Ramsey). It is noted that these counties are the economic centers of the state as they employ two thirds of the state workers and contain approximately half of the population.

(e) **Next To Regional Centers**: Counties in (c) plus the Twin Cities metropolitan counties other than Hennepin and Ramsey

(f) **Rural**: Counties not included in categories (b) and (c).

Naturally, other types of groupings are possible, e.g. counties whose economy is agriculturally-based, light manufacturing-based, border counties, etc. However, we have not yet considered such other groupings.

**EQUATIONS**

We hypothesize that the variables $y_{i,t}$ and $x_{i,t}$ (e.g. employment and highway expenditures) are stationary stochastic processes having the following form for some $q$ and constants $a_i$ for each county $i$ in the grouping:
\[ y_{i,t} = \alpha_i + a_1 y_{i,t-1} + a_2 y_{i,t-2} + \ldots + a_q y_{i,t-q} \]
\[ + b_1 x_{i,t-1} + b_2 x_{i,t-2} + \ldots + b_q x_{i,t-q} + \mu_{i,t} \]

where \( \mu_{i,t} \) is the error term assumed to be serially uncorrelated.

That the a's and b's are the same across counties is a crucial assumption of this formulation and implies that the processes behave similarly across counties. However, the a's do reflect differences among the counties. Although a joint estimation of all the coefficients in (1) would be most efficient, a two-step procedure enables the estimation of the a's, b's and the a's in a manner that reduces the statistical efficiency slightly but greatly saves computer time.

The first step of the two-step procedure subtracts out the sample mean of each variable over time. This subtraction panels the data, forming the following new variables, \( \bar{x}_{i,t} \), where k and m are respectively the beginning and the last years of the data:

\[
\bar{x}_{i,t} = x_{i,t} - \frac{\sum_{s=k}^{m} x_{i,t}}{m - k + 1}.
\]

Since the sample mean is the estimate of the true mean of each variable, the a's are then eliminated from the equation. Therefore, we can rewrite equation (1) as:

\[
\bar{y}_{i,t} = a_1 \bar{y}_{i,t-1} + a_2 \bar{y}_{i,t-2} + \ldots + a_q \bar{y}_{i,t-q} \]
\[ + b_1 \bar{x}_{i,t-1} + b_2 \bar{x}_{i,t-2} + \ldots + b_q \bar{x}_{i,t-q} + \mu_{i,t} \]

We would expect the absolute variation of employment and highway
expenditures to be greater in large counties than in small counties. If our statistical methods did not adjust for this difference in variation, the largest two counties containing the Twin Cities would dominate giving biased results. To eliminate this bias, we consider a county's $\mu_{i,t}$ to be the sum of $n$ independent and identical random variables; then the variance of $\mu_{i,t}$ equals $n$ times the variance of one of the individual random variables. Let us next assume that the number of the random variables in a county is proportional to its total employment, $E_i$, then $\varepsilon_{i,t} = \mu_{i,t}/E_i$ is serially uncorrelated and has variance $\sigma^2$, which is independent of the county. We have found that this specification is reasonably consistent with the data.

Our final transformation filters out the effects of county size on data fluctuations by dividing both sides of equation (2) by $E_i$, yielding:

$$y_{i,t} = a_1 y_{i,t-1} + a_2 y_{i,t-2} + \ldots + a_q y_{i,t-q} + b_1 x_{i,t-1} + b_2 x_{i,t-2} + \ldots + b_q x_{i,t-q} + \varepsilon_{i,t}$$

where $y_{i,t} = \bar{y}_{i,t}/E_i$ and $x_{i,t} = \bar{x}_{i,t}/E_i$. Because this two-step procedure is not perfectly efficient, we add to (3) a constant term, independent of the county, yielding a standard regression form:

$$y_{i,t} = \gamma + a_1 y_{i,t-1} + a_2 y_{i,t-2} + \ldots + a_q y_{i,t-q} + b_1 x_{i,t-1} + b_2 x_{i,t-2} + \ldots + b_q x_{i,t-q} + \varepsilon_{i,t}$$

Equation (4) is the process we estimate.

METHODOLOGY

We use three methods to investigate the link between transportation and economic development: (a) vector autoregressions, (b) Granger-causality tests, and (c) structural plots.
Vector Autoregressions

A Vector AutoRegression (VAR) estimates a system of equations having forms similar to (4). For our analysis, the vector consists of two equations - one explaining highway expenditures and one explaining employment. The form of this VAR is given below:

\[
\begin{align*}
H_{i,t} &= \gamma_1 + a_{11} H_{i,t-1} + a_{12} H_{i,t-2} + \ldots + a_{1q} H_{i,t-q} \\
&\quad + b_{11} E_{i,t-1} + b_{12} E_{i,t-2} + \ldots + b_{1q} E_{i,t-q} + \epsilon_{i,t} \\
E_{i,t} &= \gamma_1 + a_{21} H_{i,t-1} + a_{22} H_{i,t-2} + \ldots + a_{2q} H_{i,t-q} \\
&\quad + b_{21} E_{i,t-1} + b_{22} E_{i,t-2} + \ldots + b_{2q} E_{i,t-q} + \eta_{i,t}
\end{align*}
\]

where \( H_{i,t} \) is the highway construction expenditures in county \( i \) during year \( t \), \( E_{i,t} \) is the employment in county \( i \) during year \( t \), the \( a \)'s, \( b \)'s, and \( \gamma \)'s are coefficients, and \( \epsilon \) and \( \eta \) are the error terms. When equations (5a) and (5b) are estimated, the system of equations (5a & 5b) are simulated, first, to reflect a change in highway expenditures and, second, to reflect a change in employment. The simulations reflecting these changes are then plotted through time enabling one to interpret the dynamics of the relationships between employment and highways. We note that each change is one orthogonalized standard deviation of each variable. Using one standard deviation insures that the simulated deviation is reasonable given the data; the orthogonalization insures that the change is typical of the data reflecting correlations with other variables. For examples of plots of these simulations, see figures 3 and 4.

Granger Causality

The direct Sims test of whether a variable \( x \) Granger-causes a variable \( y \) first formulates the null hypothesis that \( x \) does not Granger-cause \( y \). Then, \( x \) is regressed on past, present, and future values of \( y \):
VECTOR AUTOREGRESSIONS

HWY. EXPENDITURES
TOTAL EMPLOYMENT

HIGHWAY ($1000s)
EXPENDITURES
EMPLOYMENT

0
5
10
15
years

0
-30
+30

0
-30
+60

0
-300
+300

0
-60
+60

0
-600
+600

0
-200
+200

0
-10
+10

0
-0.4
+0.4

0
-0.08
+0.08

100.00
99.98
100.00
99.98

100.00
99.964
100.00
99.98

100.16
100.00
100.00

100.03
100.00
100.00

100.003
100.000

100.21
100.00
100.00

100.03
100.00
100.00

110
100
0
5
10
15

FIGURE 3. Effect of a change in highway expenditures on employment.
VECTOR AUTOREGRESSIONS

HWY. EXPENDITURES
TOTAL EMPLOYMENT

0 5 10 15 years

HIGHWAY ($1000s) (jobs)
EXPENDITURES EMPLOYMENT

+ 4 +8
0 0
- 4 -8
+80 +200
0 0
-80 -200
+0.3 +1.0
0.0 0.0
-0.3 -1.0
+60 +200
0 0
-60 -200
+2 +4
0 0
-2 -4
+0.6 +0.6
0.0 0.0
-0.6 -0.6

FIGURE 4. Effect of a change in employment on highway expenditures.
\[(6) \quad x_{i,t} = \gamma + a_1 x_{i,t-1} + a_2 x_{i,t-2} + \ldots + a_q x_{i,t-q} + b_0 y_{i,t} + b_1 y_{i,t-1} + \ldots + b_q y_{i,t-q} + c_1 y_{i,t+1} + c_2 y_{i,t+2} + \ldots + c_k y_{i,t+k} + \epsilon_{i,t}\]

for some integers \(q\) and \(k\).

Under the null hypothesis of no causality, all future coefficients of \(y\) should be zero, i.e. \(c_h = 0\) for \(h = 1, 2, \ldots, k\). An F-test is used to test whether these coefficients are zero. If the F-test indicates the observed data is unlikely to have occurred if all the future coefficients of \(y\) were zero, then we would reject the null hypothesis and conclude that \(x\) does Granger-cause \(y\).

**Structural Plots**

The VAR system (5) is a reduced-form system, not a structural one. Reduced-form systems cannot be meaningfully analyzed equation by equation. However, we can analyze the individual equations of a structural system. To estimate a structural system of highway expenditures and employment, we first hypothesize the following structural form explaining highway expenditures for some integers \(q\) and \(k\):

\[(7) \quad H_{i,t} = \gamma_1 + a_{11} H_{i,t-1} + a_{12} H_{i,t-2} + \ldots + a_{1q} H_{i,t-q} + b_{11} E_{i,t-1} + b_{12} E_{i,t-2} + \ldots + b_{1k} E_{i,t+k} + \epsilon_{i,t}\]

A simultaneity bias would exist in equation (7) if current employment, \(E_{i,t}\), were included. However, because all the variables on the right side of (7) are predetermined, (7) is free from any simultaneity bias. We exclude \(E_{i,t}\) from (7) because we know from discussions with MnDOT that funding for highways is determined some time before the money is expended - resurfacing projects are selected usually two years ahead of time and construction projects are selected usually four or more years ahead of time. Thus, it is reasonable for us to exclude current employment from (7).
Another justification for excluding current $E_{i,t}$ from (7) is that the highway expenditures data are by fiscal year, from July 1 of one year to June 30 of the next, whereas each year's employment data represent employment in the second week of March. Since the vast majority of the highway expenditures have already been expended (and certainly appropriated) by the second week in March, it is unlikely that $E_{i,t}$ has a significant impact on $H_{i,t}$.

We hypothesize that the structural equation for employment is, for some $q$ and $k$:

\begin{equation}
E_{i,t} = \gamma_1 + a_{20} H_{i,t} + a_{21} H_{i,t-1} + \ldots + a_{2q} H_{i,t-q} \\
+ b_{21} E_{i,t-1} + b_{22} E_{i,t-2} + \ldots + b_{2q} E_{i,t-k} + \epsilon_{i,t}
\end{equation}

We argue that all variables on the right-hand side of (8) are predetermined, and thus the structural equation (8) is identified. The only variable that one may view as not being predetermined is $H_{i,t}$. However, as we discussed in the previous paragraph, the vast majority of $H_{i,t}$ occurs before $E_{i,t}$, i.e. the vast majority of $H_{i,t}$ is predetermined when $E_{i,t}$ occurs.

To interpret the structural equations, we simulate an exogenous change in employment for (7) and highway expenditures for (8). For examples of these structural plots see figures 3 and 4. The structural plot of equation (8), which explains employment, exogenously increases highway expenditures 10% for one period and then returns the expenditures back to its original level. The structural plot of equation (7), which explains highway expenditures, exogenously increases employment 10% and keeps employment at that level. The reason the change in highway expenditures for equation (8) is temporary is that we are interested in the effects of highways after construction, not the effects from the construction of the highway. On the other hand, for equation (7), we are primarily interested in how highway expenditures respond to a permanent employment increase, not a temporary one.
While we have argued that (7) and (8) can be viewed as the structural equations representing highway expenditures and employment respectively, we only have limited confidence in them. Clearly, highway expenditures is not the only variable that affects employment, and employment is not the only variable that affects highways. Thus, we expect that the two equations suffer from some misspecification. We also note that part of our justification of no simultaneity bias in equations (7) and (8) stems from properties of the data rather than from true structural properties. This limited confidence in these structural equations is why we also look at the VARs. Sims (1980) considers the use of vector autoregressions as preferable to the use of structural systems, which often rely on weak assumptions to deal with the simultaneity bias. Thus, the plots of the structural equations should be considered as a supplement to, not a replacement of, the VARs.

RESULTS

We use the methods of VARs, causality tests, and structural plots to study the relationships between highway expenditures and employment for the following groupings of counties:

(a) Statewide
(b) Urban
(c) Next to Urban
(d) Regional Centers
(e) Next to Regional Centers
(f) Rural

Because Box and Jenkins (1976) indicate that a lag structure of three to five lags will usually capture most of the dynamics of a system, we use an autoregressive structure of five lags for the VARs and structural plots (i.e. q=5 in equations [5a] and [5b], and q=5 and k=5 in equations [7] and [8]). However, because of leads in equation (6), the causality tests require more data for a given autoregressive structure than do the VARs and structural plots.
Therefore, since the length of our time series is only 19 years and each additional lag decreases our degrees of freedom by the number of counties, we chose to use three lags in the causality tests (i.e. $q=3$ in equation [6]).

Corresponding to the three-lag autoregressive structure, we initially used three leads in the causality tests, i.e. $k=3$ in equation (6). However, the major effects of highways on economic development may occur beyond three years into the future. Thus, we also did the causality tests for six leads. Tables 1 and 2 present both the three-lead and the six-lead causality tests. A low significance level in the three-lead tests indicates the existence of a short-term effect, whereas a low significance level in the six-lead tests indicates the existence of a long-term effect.

The results of the VARs, causality tests, and structural plots are summarized in tables 1 and 2, and the VAR plots and structural plots are presented in figures 3 and 4. Table 1 and figure 3 present how changes in highway expenditures affect employment. Table 2 and figure 4 present how changes in employment affect highway expenditures.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3 leads</td>
<td>6 leads</td>
</tr>
<tr>
<td></td>
<td>F-statistic (sign. level)</td>
<td>F-statistic (sign. level)</td>
</tr>
<tr>
<td>StateWide</td>
<td>0.43 ( &gt; 20%)</td>
<td>0.37 ( &gt; 20%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.8 (0.005)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.5 (0.001)</td>
</tr>
<tr>
<td>Urban</td>
<td>1.26 ( &gt; 20%)</td>
<td>1.66 (12.86%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>44.8 (0.042)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>52.0 (0.008)</td>
</tr>
<tr>
<td>Next to Urban</td>
<td>0.49 ( &gt; 20%)</td>
<td>0.47 (12.86%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-4.9 ( -0.003)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-14.0 ( -0.002)</td>
</tr>
<tr>
<td>Regional Centers</td>
<td>2.51 (4.82%)**</td>
<td>3.31 (0.60%)***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>137.0 (0.086)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>107.6 (0.013)</td>
</tr>
<tr>
<td>Next to Regional Centers</td>
<td>0.24 ( &gt; 20%)</td>
<td>0.77 ( &gt; 20%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2.6 ( -0.003)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.1 ( -0.001)</td>
</tr>
<tr>
<td>Rural</td>
<td>2.10 (7.83%)*</td>
<td>1.88 (9.9%)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.8 (0.0034)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.3 (0.001)</td>
</tr>
</tbody>
</table>

* significant at the 10% level
** significant at the 5% level
***significant at the 1% level

+ The number of jobs represents the average annual increase of jobs over the base year's employment for a typical county in each grouping. In the VARs, the $1 million were expended over a five-year period rather than over a one-year period as with the structural plots. Thus, in the VARs, highway expenditures increased on average by $200 thousand per year for each of those five years; whereas, in the structural plot, highway expenditures increased by $1 million dollars in the first year only. Since elasticity computations are based on the average percent increase of highway expenditures over the period during which expenditures increase, the elasticity of the VARs should be expected to be five times greater than the elasticity of the structural plots.

++ The number of jobs represents the average annual increase of jobs over the base year's employment for a typical county in each grouping. The structural plots simulate the direct effects on employment of a one-year impulse in highway expenditures. Thereafter, highway expenditures are exogenously set at its base level. However, after the first period, the VAR treats highway expenditures as endogenous; thus the VAR reflects feedback effects in addition to the direct effects.
Effect of Highway Expenditures on Employment:

Based on the results from the VARs, causality tests, and structural plots; table 1 summarizes, for each grouping, how highway expenditures lead to employment. For the causality tests, the lower the significance level, the greater the indication of causality. We consider a significance level around 1% or less as very strong indication of causality. A 5% significance level may indicate causality, but such a significance level would have about a 50% chance (.95^12) of occurring for at least one of the groupings if no causality existed. A 10% significance level provides a small indication of causality, but such a significance level would have over a 70% chance of occurring if no causality exists. Significance levels greater than 10% provide very little indication of causality.

The VAR plots and the structural plots are presented in figure 3. For the structural plots, highway expenditures are temporarily increased by 10% as indicated in the top right graph of figure 3; the effect on employment is seen in the other graphs on the right side of figure 3.

The VAR plots differ from the structural plots since, after the change is made in highway expenditures, the VAR lets highway expenditures change endogenously according to equation (5a) whereas the structural plot exogenously returns highway expenditures to its original level and keeps it there. Thus, in the VAR, Fig. 3 indicates that the increase in highway expenditures is spread over about five years; whereas, in the structural plot, the increase occurs completely within one year. As a result, even though the absolute effect of a $1 million highway expenditure increase is similar between the two methods, the elasticity for the VARs will generally be about five times greater than the elasticity for the structural plots.
For all groupings, the behavior of highway expenditures in the VARs is such that the expenditures die out slowly, taking between 5 and 6 years to return to the equilibrium level. This behavior, which is consistent in all the VARs, may reflect that work on a road is spread out over several projects on different years.

The statewide VAR plot on the left side of figure 3 shows that, for the typical Minnesota county, an increase of highway expenditures of $77,600 (0.57% of the base value, where b.v. = $2,740,000/yr * 5 yrs) over a five-year period will lead to an increase of 0.45 jobs (0.003%) on average over the ten years following the first change in highway expenditures. (This calculation is based on measuring the area under each curve.) This implies, as summarized in table 1, that 5.8 new jobs follow a $1 million increase in highway construction expenditures (an elasticity of 0.0048).

The statewide structural plot on the right side of figure 3 shows that, for the typical Minnesota county, a one-year increase of highway expenditures of $274,000 (10%) leads to an annual average increase of two jobs (0.01%). This implies, as summarized in table 1, that 7.5 new jobs follow a $1 million increase in highway construction expenditures (an elasticity of 0.001). (Note that the elasticity for the VAR is about five times greater than the structural plot elasticity as we would expect it to be.)

As indicated by the causality tests in table 1, the effects of highway expenditures on employment, depicted in the VAR and structural plots, are statistically insignificant; neither the three-lead nor the six-lead tests indicate any evidence of highways Granger-causing employment. This insignificance is due to the small magnitude of these effects.
In the evaluation of the statewide data set, the model assumes that all counties behave the same. If, in fact, they behave differently, then the results of the previous section may only be true on average, but not for every type of county. To isolate differences in behavior, we have analyzed different groupings of counties according to their urbanization.

The causality tests indicate strong causality only for the Regional Centers, although a small degree of causality is indicated for Rural counties. For the Regional Centers, the six-lead significance level of 0.60% strongly indicates a long-term employment effect of highways on employment. This long-term effect is brought out by the VAR plot for the Regional Centers on the left side of figure 3. After highway expenditures increase, employment increases gradually, reaching its highest level after eight years.

The structural plot for Regional Centers in figure 3, indicates two effects occurring. The first effect is the construction effect which lasts between two to three years. The construction effect on the economy is due to the road being constructed - e.g., construction jobs are created, the workers spend some of their earnings in the county, and the construction companies make local purchases causing multiplier effects throughout the county's economy. The second effect is the longer-term employment effect resulting because the highway improvement exists. This latter effect is the more sustaining effect of highways and the one in which we have been primarily interested.

Some causality is also indicated for the Rural counties. The VAR plot for the Rural counties on the left side of figure 3 shows that employment closely follows the change in highway expenditures, indicating that the effects of highway expenditures on rural economies are primarily due to the
construction effect. That the construction effect and not a more long-term effect of highways takes place in rural economies is also indicated by the causality tests. The causality for three-leads (short-term) has a significance level of 7.83% but the significance level for the six-lead test is about 10%, i.e., there is little evidence of long-term causality.

While causality is not indicated for the other groupings, the VARs and structural plots do indicate that the effects of highways on employment are not always positive. For the Next-to-Urban counties, the VAR plot in figure 3 shows that a $16,000 (0.44%) increase in highway expenditures over four years is followed by an average 0.08 job (0.0014%) decrease over the next decade. As summarized in table 1, this effect amounts to 4.9 jobs per million dollars of new highway expenditures (an elasticity of -0.003). One explanation of this negative employment effect is that improved highways into a regional center will allow business activity to move from the Next-to-Urban counties to the Regional Centers. This explanation also may be the reason that employment in Regional Centers significantly increases following an increase in highway expenditures.

In summary, the statewide data set did not indicate any significant effect of highway expenditures on employment levels. Nevertheless, in Regional Centers, higher levels of expenditures did lead to significantly greater levels of employment. In counties next to urban areas, employment actually dropped following increases in expenditures, although this effect lacks statistical significance. A possible explanation of these results is that improved highways in or around urban areas cause business activity to be drawn into the regional centers from counties near the urban area.
Table 2: Effect of employment on highways

<table>
<thead>
<tr>
<th></th>
<th>3 lags</th>
<th>6 lags</th>
<th>100-Job Increase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F-test of Ho: EMPL $\rightarrow$ HWY</td>
<td>Effect of 100-Job Increase</td>
<td>VAR Effect</td>
</tr>
<tr>
<td></td>
<td>F-statistic (sign. level)</td>
<td>F-statistic (sign. level)</td>
<td>$1000s$ (elasticity)</td>
</tr>
<tr>
<td>Statewide</td>
<td>3.31 (0.35%)***</td>
<td>3.84 (1.1%)**</td>
<td>28.4 (1.71)</td>
</tr>
<tr>
<td>Urban</td>
<td>0.88 (&gt; 20%)</td>
<td>1.41 (&gt; 20%)</td>
<td>32.2 (2.08)</td>
</tr>
<tr>
<td>Next to Urban</td>
<td>0.58 (&gt; 20%)</td>
<td>0.93 (&gt; 20%)</td>
<td>-0.1 (-0.004)</td>
</tr>
<tr>
<td>Regional Centers</td>
<td>0.20 (&gt; 20%)</td>
<td>0.47 (&gt; 20%)</td>
<td>17.2 (1.38)</td>
</tr>
<tr>
<td>Next to Regional Centers</td>
<td>1.01 (&gt; 20%)</td>
<td>1.46 (&gt; 20%)</td>
<td>42.8 (1.75)</td>
</tr>
<tr>
<td>Rural</td>
<td>2.42 (2.61%)**</td>
<td>3.38 (1.80%)**</td>
<td>57.5 (2.37)</td>
</tr>
</tbody>
</table>

* significant at the 10% level
** significant at the 5% level
*** significant at the 1% level

* The reported increase in highway expenditures for both the VARs and the structural plots is the average annual increase over the first ten years for a typical county in each group. The structural plot simulates the direct effects on highway expenditures of a permanent and exogenous (step) increase in employment of 100 new jobs. In the VAR, employment is endogenous after the first period and will reflect feedback effects in addition to the direct effects of the new jobs.
Effect of Employment on Highway Expenditures

Based on the results from the VARs, causality tests, and structural plots; table 2 summarizes how employment leads to highway expenditures for each grouping. The VAR plots and the structural plots are presented in figure 4. For the structural plots, employment was permanently increased by 10% as indicated in the top right graph of figure 4; the effect on highway expenditures is seen in the other graphs on figure 4's right side. The VAR plots differ from the structural plots since after the change is made in employment, the VAR lets employment change endogenously according to equation (5b) whereas the structural plot exogenously sets employment at the increased level.

As indicated in table 2, increases in employment do lead to more highway expenditures. The VAR for the statewide data set, plotted on the left side of figure 4, demonstrates that a 7.2 employment increase (.04%) will lead to an additional $2,050 (.07%) being expended on average for highway construction over the next decade. As summarized in table 2, this is a $28,400 increase per 100 new jobs (an elasticity of 1.71). Since the significance levels of both the three-lead and six-lead causality tests are less than 1%, strong indication exists that highway expenditures follow changes in employment levels.

The analysis for the groupings of counties is generally consistent with the findings for the statewide data set. The VARs (left-side of figure 4) and the structural plots (right-side of figure 4) are both very similar among the groupings with the exception of the Next-to-Urban grouping. For the Next-to-Urban grouping, the VAR plot shows an increase in employment leading to an increase in highway expenditures in the next year followed by a stronger negative effect lasting for three years. Over
the ten years following the change in employment, highway expenditures experience a slight net decrease.

For all the groupings except the Next-to-Urban grouping, the response of highway expenditures to higher employment has two peaks in figure 4. The first peak is especially noticeable in the VARs and structural plots for the Statewide, Regional-Centers, Next-to-Regional-Centers, and Rural groupings. We interpret the first peak, occurring one year after the increase in employment, as a quick response of MnDOT to the immediate needs of the newly created employment. For example, if a new plant is built, MnDOT may react quickly modifying the road to improve access to the plant. We interpret the second and larger peak, occurring four years after the employment increase, as a response of MnDOT, through its regular programming selection, to meet the needs created by the increased level of employment. Since MnDOT usually selects the projects it funds two to four years ahead of time, it cannot fully react to an unanticipated employment change before this time. The structural plot in figure 4 of a permanent change in the employment for the statewide grouping shows that the highway program is relatively unchanged in the second and third years following the employment change.

The causality tests did not indicate significant causality for each subgroup even though, with the exception of the Next-to-Urban counties, the VARs showed strong effects, as evidenced by elasticities greater than 1. Only the rural counties experienced significant causality. Because the significance of the F-test for causality depends on the degree of data variation, the significant causality for the statewide group without statistical significance for most of the subgroupings indicates that little variation occurs within each group whereas more variation occurs between groups resulting with significant causality for the statewide F-test.
Conclusion

The time-series analysis presented here indicates that, in general, increases in total employment lead to increases in highway expenditures. However, the analysis indicates that increases in highway expenditures do not, in general, lead to increases in employment levels. Some previous observers have mistaken the high correlation between highway expenditures and employment as an indication that highway expenditures do have a substantial affect on economic development. However, our analysis indicates that the correlation stems from two other factors:

(a) Higher employment levels attract higher levels of expenditures, and

(b) During the year of construction, employment levels do increase. However, this effect is only temporary and disappears when the construction ends.

Thus, we conclude that total employment does Granger-cause highway expenditures, but that highway expenditures do not Granger-cause total employment. However, in counties that are economic centers of the state (defined as, "Regional Center counties"; these counties employ two thirds of the state workers and approximately half of the population), highway expenditures do have an effect on total employment, above the normal trend of the economy. In these counties a one-year $1 million increase in highway expenditures leads to between 108 to 137 new jobs.

Discussion

Although our analysis concluded that, in general, highway expenditures do not Granger-cause total employment, highway expenditures may Granger-
cause the employment within a specific economic sector. Other work we have been conducting indicates that, for some sectors, the Granger-causality of highway expenditures is significant even for groupings other than the Regional Centers. That work will be reported elsewhere.
autoregressive representation - a variable $y$ has an autoregressive representation if it can be expressed in terms of past values of itself and (past and current) noise.

causality tests - tests to determine if changes in variable $x$ consistently precede changes in variable $y$. The literature and our experience suggest that, in most cases, causality is an overly strict test. When employed with properly pretreated data, the test is not likely to indicate two variables are causally related if they actually are irrelevant. In our analysis, any such indication of causality had to pass two important tests, i.e., (a) careful examination of the time-series behavior of the data as revealed by the vector autoregression plots; and (b) a theoretical or empirical explanation. Overall, our analysis is not perfect but where it may have erred this has most likely been on the conservative side. In other words, there may exist additional causalities between variables, which strict application of our tools -- causality tests, direct examination of time-series plots, theoretical and empirical expectations -- may have missed.

correlation coefficient - a statistic showing how the variables vary together. If the correlation coefficient is 1 then the two variables are perfectly correlated with each other meaning that, if one increases by a percentage, the other will increase by the same percentage. If the correlation coefficient is 0, then the two variables are uncorrelated meaning that the value of one variable is unrelated to changes in the other variable.
covariance - a statistic showing how variables vary together. The covariance equals the correlation coefficient times the two standard deviations (i.e., times the standard deviation of each of the two variables).

covariance function - for a random variable \( x \), the function \( C(t,s) \), which equals the covariance of \( x_t \) and \( x_s \), where \( t \) and \( s \) refer to two different periods of time.

covariance-stationary processes - a stochastic process is covariance-stationary if and only if the process' covariance function \( C(t,s) \) depends only on the difference \( t-s \).

cross-sectional analysis - statistical analysis that uses data, individual observations of which are from different areas, individuals, etc. at the same point in time

elasticity - the percent change in one variable for each percent change in another variable

dependent variable is endogenous to a model if that variable is determined within the model, i.e. there is a model equation where that variable is the dependent variable.

dependent variable is exogenous if it is not endogenous. The variable is determined outside of the model, i.e. there is no model equation where that variable is the dependent variable. Also, see statistical exogeneity.

exogeneity tests - the same as causality tests (See Sims, 1972, for proof that Granger-causality test is the same as exogeneity test)

Granger-causality - a variable \( x \) is said to Granger-cause \( y \) if the information on present and past \( x \) is useful to predicting \( y \).
input-output analysis - economic analysis using models that are based on the input-output relationships observed for a particular region or country. These models usually assume that these relationships are fixed and do not change endogenously.

independent random variables - two random variables are independent if the value of one of these variables has no effect on the value of the other variable.

identical random variables - two random variables are identical if they have the same probability distributions.

lag - refers to the number of periods we look back when using a particular variable. For example, \( x_{t-i} \) has a lag of \( i \). If \( t=1980 \) and the lag \( i \) is 10, then \( x_{t-10} \) refers to \( x \) at time 1970.

lead - refers to the number of periods we look into the future when using a particular variable. For example, \( x_{t+i} \) has a lead of \( i \). If \( t=1980 \) and the lead \( i \) is 5, then \( x_{t+5} \) refers to \( x \) at time 1985.

moving average representation - the representation of a stochastic process as a weighted sum of present and past white noise.

panel data - time-series data having its mean over time subtracted out.

probability distribution - a probability distribution is a description of the probability associated with each possible value of a random variable. For example, when tossing a coin, you will have a 50% probability of getting heads and a 50% probability of getting tails. The following table can be viewed as the probability distribution of flipping a coin:

<table>
<thead>
<tr>
<th>Side of coin</th>
<th>Heads</th>
<th>Tails</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>50%</td>
<td>50%</td>
</tr>
</tbody>
</table>
random variable - a variable whose value depends on some randomness

reduced-form equation - an equation that has all predetermined or exogenous
variables on the right side of the equation. In general such an
equation is not structural (See example at end of glossary).

serial correlation - a variable is serially correlated if $x_t$ and $x_s$ are
correlated for some $s$ not equal to $t$, where $t$ and $s$ refer to two
different periods of time.

stationary process - a stochastic process $\{x_t\}$ is strictly stationary of
order $k$ if the joint probability distribution of $x_t, x_s, \ldots, x_v$ is
the same as that of $x_{t+k}, x_{s+k}, \ldots, x_{v+k}$. (For instance, the mean
and variance of a stationary process are constant through time.)
A stochastic process $\{x_t\}$ is strictly stationary if, for any $k$,
it is strictly stationary of order $k$.

statistical exogeneity - a variable $x$ is statistically exogenous to $y$ if $x$
is uncorrelated with the error term of the structural equation
for $y$

stochastic process - a random variable defined over time, e.g.

$x_0, x_1, \ldots, x_n, x_{n+1}, \ldots$

structural equation - an equation that expresses the true behavior of a
variable in terms of the other variables (See example at end of
glossary)

time-series techniques - statistical techniques used for studying a time
series. These include causality tests, vector autoregressions,
ARIMA models, Box-Jenkins analysis, and others.

time series - a list of a stochastic process's values through time
white noise - a random variable that (a) has a mean of zero, (b) has a constant variance, and (c) is serially uncorrelated

Example of the difference between reduced-form and structural equations:

Suppose we have the following two reduced-form equations:

\begin{align}
(1) \quad x_t &= 10 + 2x_{t-1} + 3y_{t-1} + \epsilon_t. \\
(2) \quad y_t &= 5 + 3x_{t-1} + y_{t-1} + \eta_t. 
\end{align}

These are reduced-form equations because everything on the right-hand side of the equations is predetermined.

Reduced-form systems cannot in general be transformed to give the true structural system, because there is an infinite number of structural systems that correspond to (1) and (2). The lack of a unique structural system is called the identification problem. This identification problem occurs when \( y_t \) enters into the structural equation for \( x_t \) and \( x_t \) enters into the structural equation of \( y_t \). For example, if we add equation (1) to equation (2) and solve for \( x_t \), we get:

\begin{align}
(3) \quad x_t &= 15 - y_t + 5x_{t-1} + 4y_{t-1} + (\epsilon_t + \eta_t).
\end{align}

If we add 2 times equation (1) to equation (2) and solve for \( y_t \), we get:

\begin{align}
(4) \quad y_t &= 25 - 2x_t + 7x_{t-1} + 7y_{t-1} + (2\epsilon_t + \eta_t).
\end{align}

Equations (3) and (4), representing how \( x \) and \( y \) change, are equivalent to (1) and (2) and could be considered as the structural equations of (1) and (2). Then, if one analyzed equations (3) and (4) separately, one would conclude that, when \( y_t \) increases, \( x_t \) decreases; and when \( x_t \) increases, \( y_t \) decreases.
Now, if we subtract equation (2) from equation (1) and solve for $x_t$, we get:

$$x_t = 5 + y_t - x_{t-1} + 2y_{t-1} + (\epsilon_t - \eta_t)$$ (5)

If we subtract two times equation (1) from equation (2) and solve for $y_t$, we get:

$$y_t = -15 + 2x_t - x_{t-1} - 5y_{t-1} + (\eta_t - 2\epsilon_t)$$ (6)

Equations (5) and (6), representing how $x$ and $y$ change, are also equivalent to (1-2) and could also be considered as the structural equations of (1-2). Then, if one analyzed equations (5) and (6) separately, one would conclude that, when $y_t$ increases, $x_t$ also increases; and when $x_t$ increases, so does $y_t$. However, the opposite conclusions were reached when analyzing equations (3) and (4) separately. Clearly, the system of equations (3) and (4) differ from the system of equations (5) and (6). However, the statistical results from the regressions of equations (1) and (2) will not be able to identify which system of structural equations is the true system. In summary, we cannot derive a unique set of structural equations from the reduced-form set of equations (1-2).
BIBLIOGRAPHY


68. Smith, D.M. (????). Industrial Location, John Wiley.


87. U.S. Dept. of Commerce (1964-). County Business Patterns, Washington, D.C.


