A predictive model of highway interchange land use development: A decision tool for planning

by
RAYMOND W. EYERLY
RICHARD D. TWARK
RICHARD B. NASSI
The Pennsylvania State University, Pennsylvania

INTRODUCTION

The most extensive highway construction program ever attempted in the United States of America is the National System of Interstate and Defense Highway which, when completed, will comprise over 42,000 miles of multi-lane divided highway. This system will link together more than 90 percent of the cities of the United States having populations of 50,000 or more.[1]

Incorporated into the program (Federal-Aid Highway Act of 1956) authorizing the construction of the Interstate System are provisions which prohibit roadside development of user oriented facilities such as gasoline stations, restaurants and motels within the rights-of-way of the Interstate System. The controls, in effect, force users to exit the system for services. In addition, the Act also limits the number of access and egress points to grade-separated interchanges with connecting highways. Interchanges are designed with approach and exit ramps so as to channel traffic to and from the Interstate System, without hindering the free and save movement of traffic upon the system or the connecting highways. It is estimated that there could be as many as 14,000 of these interchanges on the completed Interstate System.[2]

The fact that traffic to and from the Interstate System must be channeled through an interchange makes the neighborhood area uniquely favorable for economic development. Users traveling long distances do not ordinarily desire to go far from an interchange in search of gasoline, food, or lodging, and hence, the vicinity of an interchange is advantageous for the construction of highway oriented establishments such as service stations, restaurants and motels.

Just as many present trade centers, towns, and cities can trace their origins to the existence of transportation crossroads (e.g., river junctions or railroad connections) which provided improved linkages between local land uses and distant land uses, the interchange area also offers opportunity for community development by reason of improved access between land uses. The development of commercial or industrial establishments, recreational facilities, residential units, and the above mentioned highway oriented establishments at an interchange stimulates general business activity, creates new jobs, increases income, and expands the tax base of the community.

Many of the completed interchanges have already become focal points for economic development and are likely to become more so in the future. However, not all interchange areas have the same potential for economic development. This is supported by the fact that some interchanges have grown considerably in a very short period while others have shown no development even after a lapse of several years. A problem of economic importance is the determination of the reasons for such differences in development.

OBJECTIVE OF THE STUDY

The primary concern of this investigation is to develop a method for predicting the economic development that is likely to occur at a given interchange site. The goal is to identify the interrelationships among the important factors leading to interchange development and to provide planners with a guide to estimating the potential development of non-urban interchange areas. Knowing the probable level of interchange area development, State and local planners can then prepare a reasonable land use plan for the interchange. Design engineers can, in turn, proceed to determine the highway capacity requirements, geometric configurations and traffic control needs according to the land use plan framework. With an accurate development forecast, the highway design and land use plan can be coordinated to encourage desirable and efficient development, while at the same time facilitating both traffic flow and safety and reducing the probability of premature obsolescence. In cases where premature obsolescence of the interchange facility has occurred, knowledge of future growth potential can be incorporated in the redesign and reconstruction of the interchange.

Many problems are involved in attempting to develop a means of predicting economic development at interchange areas. Some of the initial points to be established involve proper definition of

1) the interchange area or "interchange community," and

2) the economic development or "economic growth" of an interchange community. Another type of problem arises in quantifying the various factors which cause differences in the economic growth of interchange communities.

AN ECONOMETRIC GROWTH MODEL

The spontaneous economic development of a community is a complex phenomenon which involves balancing the rates of growth of the various segments that together form the community. There is, however, considerable variety in community types (residential, commercial, industrial etc.), and the balanced growth of one type of community will differ from that of another.

The pattern of development for a particular community will often depend upon circumstances that are unique to that community. For example, the speculative withholding of land from the market may delay the growth of the community, whereas the establishment of a new industrial plant close to a residential area may substantially promote its growth.

Since the economic growth of a community is such a complex phenomenon, to describe the state or level of development for a given community must involve measurement of the many characteristics of that community. The population of the community, number of persons employed in the community, capital investment in service facilities, number of housing units and the like, are all examples of variables that may be used to describe the level of economic development for a given community.

In order to predict the development of the community as a whole, each variable (which measures a particular characteristic or segment of the community) would have to be predicted. On account of the interrelationships of these variables, each one affects the "growth" of the community, and is, in turn, affected by the growth that takes place. For example, as a residential community grows, a need for service facilities such as grocery stores, gasoline stations for automobiles, variety and drug stores etc. develops. Consequently, this development encourages further residential expansion.

Similarly, a large flow of traffic on the crossroad at an interchange area may lead to more intense development of service facilities such as service stations, restaurants and motels at the interchange. As a result, this development may cause additional traffic on the crossroad.

Factors of this type which affect one or more other factors and, in turn, are influenced by them, are called *endogenous* variables. In short, these are the variables whose values the model seeks to explain.

In addition to the endogenous variables described, there are other factors that affect the level of economic development or growth of an interchange community; these factors, however, are generally not affected by the growth that takes place. These factors are called *exogenous* variables. For example, the variables which describe the geographical or physical environment of an interchange community are generally not influenced by economic development, although they do affect the degree of development that may be expected to occur.

Exogenous factors affecting the growth of an interchange community may include distances from nearby urban centers, topographic characteristics of the interchange area, access, zoning etc. Briefly, these are the variables which are determined apart from development at an interchange area. They are predetermined and may be regarded as given for purposes of explaining the values of the endogenous variables.

The classification of a variable as exogenous or endogenous depends upon the nature and objectives of the model. A variable classified as exogenous in one model may be considered as endogenous in another. The criterion to be used in classifying a variable is to regard as exogenous to a specified system those variables which influence the remaining (endogenous) variables but are not, as a result, affected.[3]

ASSUMPTIONS OF THE MODEL

It will be assumed that the entire model can be presented as a system of simultaneous linear equations [4] where each equation describes the way in which a particular aspect of economic development (endogenous variable) is determined by other relevant endogenous and exogenous variables. Such an equation describes a particular "structure" of the economic community and is called a structural equation of the model.

Any one of the structural equations in the model will have the same mathematical appearance as that of an ordinary multiple regression equation. However, the parameters of a structural equation in a system of equations cannot (in general) be derived by ordinary regression techniques. Other suitable methods have been developed.[5] One of these should be used.

In addition to the endogenous and exogenous variables in the model, it is assumed that each equation contains a stochastic variable called a disturbance (or shock) term which represents the aggregate effects of the unspecified variables for each relation.

An exceedingly large number of variables may influence the economic development of an interchange community. Many of these may have very slight effects, others are not quantifiable. Still, others may be unique for a given area. It is not practical to include all of these variables in a model. Hence, the model specifies only the variables thought to be most important. The net effect of the excluded variables for each equation is then represented by a disturbance term.

It will be assumed, for purposes of estimating the parameters of the model, that each of the disturbance terms behaves as a random variable having an expectation of zero, constant variance, and zero covariance with each other disturbance term.

Besides assuming linear structural equations for the model, we shall assume that no errors of observation have been made in measuring the endogenous and exogenous variables.

LIMITATIONS OF THE MODEL

It is assumed that the entire model can be presented as a system of simultaneous linear equations, in which a separate equation is used to describe the level of economic development of each of the segments of the interchange community in terms of the levels of development of the other segments (endogenous variables) and in terms of the characteristics of the environment itself (exogenous variables). The actual relationships among the variables of the model, however, may be more complex than the linear form assumed, in which case misleading estimates of the influence of the variables may result.

There may be unique or peculiar characteristics associated with specific interchanges which the model (being of a general nature) will not take into account, except as they may be represented as a part of the stochastic term. For example, the speculative withholding of land from the market may delay the growth of an interchange area. Similarly, intangible factors (e.g., aesthetic considerations) or other non-economic factors which might have a profound influence upon the degree to which a given interchange develops may be unaccounted for by the model.

The model developed is of a static nature. That is, it describes the equilibrium state of economic development in the neighborhood of a given interchange. It cannot, however, predict the *course* of economic development.

Since the model will be designed to predict economic development at interchange areas, the problem of properly defining the interchange area or "interchange community" arises. The area to be included or the boundary of the interchange community is not easily determined.

Some authors have used the term "interchange area" rather loosely to cover the entire vicinity in which the existence of the interchange may stimulate intensive uses of land that would not, otherwise have located there.[6] Others have used the term "area of influence" to mean the area within the vicinity of the interchange that is affected by the facility.[7]

Various studies have found that for non-urban interchanges, the majority of new economic development occurs within one-half mile of the interchange.[8] For this reason, the model will consider the "interchange community" as the area located within one-half mile of the interchange. Some arbitrary limitation of the interchange community is necessary and the one chosen will have some adverse effects on the applicability of the model to some interchanges, where exceptional geographic or topographic conditions lead to important developments more than one-half mile from the interchange.

Another very practical as well as important consideration is that the model be designed such that it can be implemented easily using secondary data. Although a model that requires an extensive primary data collection effort may have greater predictability for a specific location, it would have very limited general value.

DEVELOPMENT OF THE SIMULTANEOUS EQUATIONS MODEL

The development of a particular structure of the general model that reasonably may be expected to describe and predict development of an interchange community is itself a complex matter. Firstly, the relevant variables have to be identified and measured. These variables must be classified into the endogenous and exogenous categories.

The next step is the construction of the equations of the model. Involved in this is the problem of determining the form of the equations that best describes the nature of the relationships among the variables. Also there is the problem of determining which of the many variables that belong to the complete system should be included in (or excluded from) a particular equation. An incorrect inclusion or exclusion of a variable constitutes what is known as an error of specification and may have serious effects on the estimates of the parameters of the model.[9]

After the variables of the model have been defined and the equations constructed, the next step is the statistical estimation of the parameters of the structural equations of the system. An essential prerequisite to the statistical estimation of the parameters of the model is to verify that each equation in the model is unique in the sense that it is not a linear combination of the other equations in the model.

The method of estimation to be used in this research endeavor is a form of the so-called two-stage least squares technique. [10] The set of equations that constitutes the structure contain both exogenous and endogeous variables as mentioned earlier. There are as many endogenous variables as there are equations - since each equation "explains" the level of a particular measure of development (an endogenous variable) in terms of the levels of all other variables (endogenous and exogenous). This set of equations may be solved so as to obtain for each endogenous variable an expression relating the level of that endogenous variable to the values of all the exogenous variables.

For example, the first of the structural equations will be of the form $y_1 = f_1(y_2,...,y_n, x_1,...,x_m)$ where the y's are endogenous variables and the x's are exogenous variables. There will be "n" such structural equations. This set of equations may then be solved for the y's in terms of the x's. The solution will yield a set of equations, the first of which will be of the form: $y_1 = g_1(x_1,...,x_m)$. There will be "n" of these latter equations. This derived set of equations is called the reduced form of the structure.

Each of the reduced form equations explains the level of one factor of development of an interchange (one of the y's) at a given time in terms of levels of the exogenous variables (the x's) alone. Further, the exogenous variables (x's) are predetermined in the sense that they are either fixed characteristics of the site and neighborhood of the interchange or they are determined by influences outside the structure. It follows that for a given interchange, the levels of the relevant x's may be known and used to predict the values of the y's. In other words the levels of development that may be expected to occur at any given interchange may be estimated by means of the reduced form equations. It follows from the above that the estimates derived from the reduced form equations may be interpreted as the "potential" for economic development of the given interchange.

The two-stage least squares technique involves obtaining the estimated values of the y's from the reduced form equations and substituting these estimates for the actual y's on the right hand side of the structural equations. The y to the left of the equal sign is then regressed on the estimated y's and the relevant x's. The first structural equation in this way modified to the following form:

$$y_1 = f_1(\hat{y}_2,...,\hat{y}_n; x_1,..., x_m)$$

where $\hat{y}_i$ is the estimated value of y_i obtained from the estimated reduced form equation:

$$\hat{y}_i = \hat{g}_i(x_1,...,x_m).$$

The estimates of the parameters of the structural equations obtained by this technique are asymptotically unbiased.

The importance of the structural equations may be illustrated by their ability to estimate the marginal effect on different types of development at an interchange when the level of some other type of development has been determined at a level different from that predicted by the reduced form equation. For example, the reduced form equations may estimate the economic "potential" of some particular interchange to be very small. However, if for any reason the site should be selected for a large motel, this fact would change the "potential" of the interchange and the structural equations could be used to estimate the effect of this on the further development of the interchange, e.g., the likely appearance of service stations and restaurants.

By appropriate definitions the *system* of structural equations may be represented by the following matrix form:

$$YA = XB + e$$

The solution of this system of equations would be:

$$Y = XBa^{-1} + eA^{-1}$$

which may be written:

$$Y = XC + u$$

where $C = BA^{-1}$ and $u = eA^{-1}$

The above solution is the matrix representation of the "reduced" form of the structure.

PRELIMINARY DATA ANALYSIS

Data describing traffic flow, economic activity, and geographic and demographic characteristics for 144 non-urban Pennsylvania interchanges have been collected annually by The Pennsylvania State University. The approach will be to use these data to design a model for forecasting interchange development.

The analysis presented in this section is of a preliminary nature. The purpose of the analysis is to provide insight for the construction of a specific structure of the model for predicting economic growth at non-urban interchange areas and to investigate whether the observed relationships among selected variables are in agreement with *a priori* considerations.

The data collected for each of the 144 interchanges in the sample were analyzed by means of appropriate classifications as well as simple and multiple correlation analysis. Some of the more important findings are presented here.

As a first step in the analysis, the various types of new economic development observed at the interchanges were classified as highway or non-highway oriented commercial, industrial, residential, or public. New economic developments were defined to be those developments that had taken place since the opening of the interchange. The interchanges studied were divided into two basic categories, "Complete access interchanges", which provided access to abutting lands along the cross route from all directions and "Incomplete access interchanges," which limited or totally prohibited access to one or more quadrants of the interchange.

Interchange area land development and interchange highway design are highly related. As the highway design restricts access to adjacent land, the potential for certain types of land development is sharply reduced.

For example, Table 1 shows the impact of the highway design upon land use. Of all the interchanges where access to adjoining land was limited, 44 percent of the interchange sites had no new development from their opening date to 1975. In contrast, only 4 percent of the interchanges which provided for complete access to adjoining land failed to attract any new development. The comparable figures for 1970 were 63 and 25 percent, respectively.

Commercial development was the most common form of new development occurring at 91 percent of the complete interchanges and 31 percent of the incomplete access interchanges by 1975. Perhaps of greater significance is the fact that 89 percent of the complete interchanges had highway oriented commercial development, while only 6 percent of the incomplete access interchanges had such development.

Some new industrial development occurred at 13 percent of the complete interchanges and 0 percent of the incomplete interchanges by 1970 while increasing to 25 and 6 percent, respectively by 1975. On a relative basis,

Table 1 - Types of new economic development and percentage of interchanges with each type

Type	N= 128 Complete Access Interchanges (%) 1970*	N= 128 Complete Access Interchanges (%) 1975**	N= 16 Incomplete Access Interchanges (%) 1970*	N= 16 Incomplete Access Interchanges (%) 1975**
NO NEW DEVELOPMENT	25	4	63	44
SOME NEW DEVELOPMENT	75	96	38	56
EXCLUDING RESIDENTIAL	73	94	38	50
COMMERCIAL DEVELOPMENT	70	91	31	31
HIGHWAY ORIENTED	67	89	12	6
NON-HIGHWAY ORIENTED	30	59	31	31
RESIDENTIAL DEVELOPMENT	19	35	6	19
SCATTERED	13	24	0	6
PLANNED	6	11	6	13
INDUSTRIAL DEVELOPMENT	13	25	0	6
PUBLIC DEVELOPMENT	9	13	13	25

* The ages of the interchanges in 1970 ranged from 1 to 15 years with an average age of 7.9 years for the complete and 8.0 for the incomplete access interchanges.

** The ages of the interchanges in 1975 ranged from 6 to 20 years with an average age of 12.9 years for the complete and 13.0 for the incomplete access interchanges.

Table 2 - Endogenous variables and their 1975 mean level of intensity per interchange, standard deviation and range at the 128 complete non-urban interchanges

Endogenous Variable		Mean	Standard Deviation	Range
SERVICE STATION DEVELOPMENTS				
Y_1	Number of Stations	2.1	1.5	0 to 7
Y_2	Number of Gas Pumps	13.5	11.4	0 to 54
RESTAURANT DEVELOPMENTS				
Y_3	Number of Restaurants	0.9	1.0	0 to 5
Y_4	Number of Seats	154	270	0 to 1434
MOTEL DEVELOPMENTS				
Y_5	Number of Motels	0.5	0.8	0 to 4
Y_6	Number of Rooms	38	79	0 to 455
TOTAL HIGHWAY ORIENTED SERVICES				
Y_7	Highway Oriented Developments	3.8	3.2	0 to 16
NON-HIGHWAY ORIENTED DEVELOPMENT				
Y_8	Public Developments	0.2	0.4	0 to 2
Y_9	Industrial Developments	0.5	1.1	0 to 6
Y_{10}	Other Commercial Establishments	1.6	2.5	0 to 15
TOTAL NON-HIGHWAY ORIENTED DEVELOPMENT				
Y_{11}	Non-Highway Developments (Sum of Y_8, Y_9, Y_{10})	2.3	3.1	0 to 18
TOTAL DEVELOPMENT				
Y_{12}	Highway and Non-Highway Oriented (Sum of Y_7 and Y_{11})	6.1	5.0	0 to 26

Table 3 - Exogenous variables and their 1975 mean level of intensity per interchange, standard deviation and range at the 128 complete non-urban interchanges

	Exogenous Variable	Mean	Standard Deviation	Range
X_1	Average Daily Traffic Cross Route (vehicles)	3.366	2,596	133 to 13,304
X_2	Average Daily Traffic Interstate (vehicles)	12,550	4,623	5,400 to 29,690
X_3	Community Zoning	0.59	0.49	0 or 1
X_4	Age of Interchange (years)	13	4	6 to 20 years
X_5	Distance to Nearest Urban Area (miles)	12.3	7.8	1 to 43 miles
X_6	Population Density-Local Community	157	232	10 to 2,010
X_7	Population Density Change – Local Community	24	69	–330 to 523
X_8	Market Value of Real Estate Per Capita – Local Community	$4,410	$3,710	$1,513 to $30,657
X_9	Change in Market Value Per Sq. Mile – Local Community (in 000's of dollars)	$423	$1,584	$–170 to $3,815
X_{10}	County Population Density	228	153	25 to 734
X_{11}	County Population Density Change	12	18	–17 to 78
X_{12}	County Market Value of Real Estate Per Sq. Mile (in 000's of dollars)	$890	$627	$73 to $2,977

public development occurred more frequently at the incomplete than complete interchanges. This may be due, in part, to the fact that highway oriented commercial development which usually competes for land around interchanges was not present to any appreciable degree at the incomplete access interchanges.

Due to the small number of incomplete access interchanges and the vast differences between developments at the complete and incomplete interchanges, further preliminary analysis is limited to the 128 complete interchanges.

Table 2 gives an indication of the endogenous variables considered for inclusion in a forecasting model as well as the average intensity and range in development at the 128 complete interchanges for year 1975.

Average total development (excluding residential) was six establishments per interchange and ranged from no development at a few of the interchanges to 26 units at one of the interchanges. The average number of highway oriented establishments (e.g., service stations, restaurants, motels, truck garages) was 3.8 units, while non-highway development such as industrial, public and other commercial establishments averaged 2.3 units per interchange.

As expected at the non-urban interchanges, gas stations were the most frequent form of new highway oriented development followed by restaurants and motels, respectively. For the non-highway oriented category, other commercial establishments such as those found in small shopping centers were most common, followed by industrial and public developments.

The exogenous variables considered for inclusion in the model together with their mean, standard deviation and range per interchange for the 128 complete interchanges are given in Table 3. The first two variables consisted of average daily traffic (ADT) volumes on the interstate highway and intersecting cross routes for the year 1970. Traffic data were obtained from traffic log books supplied by the Pennsylvania Department of Transportation.

The population per square mile of the local township or borough containing the interchange was also for the year 1970 and is simply referred to as population density – local community. In instances where more than one township contained the interchange, the populations of both townships were combined and an average density figure was computed.

Population density data for the county containing the interchange and all market value of real estate data were also for the year 1970, while all changes in population and market value from 1960 to 1970.

It should be noted that some of the variables that were classified as exogenous in this particular study are often classified as endogenous variables. For example, the population of an interchange community is usually of an endogenous nature in that it affects the growth of the community and is, in turn, affected by the growth that takes place. Annual population data are not readily available, hence, the population variables, being of a lagg-
vailable, hence, the population variables, being of a lagged or predetermined nature, are classified as exogenous

Similarly, the volume of traffic on the crossroad at the interchange will tend to influence the degree of development of service facilities such as gasoline stations, restaurants and the like. Consequently, these developments will create additional traffic. However, variation in the ADT will be related to the characteristics of a much larger geographic area than the immediate neighborhood of the interchange itself. The factors that determine the character of the larger area are often dominant in the explanation of the traffic flow. These factors may not be included in, or related to the exogenous variables explicit in the model, but they tend to describe the peculiar characteristics of the interchange area. Since variations in the ADT may largely be due to characteristics peculiar to the various interchanges, the ADT can be considered exogenous for it and serve as an aggregate measure of the unique characteristics of the interchange otherwise not included in the model.

It can also be noted that some of the endogenous variables in Table 2 are different measures of the same type of development. For example, service station development per interchange is measured by the actual number of gasoline stations and also by the number of gasoline pumps. In addition, service station development is included in the number of highway oriented establishments which is the sum of the numbers of gasoline stations, truck garages, restaurants, and motels at the interchange. Similarly, restaurants and motel development are measured in different ways. These different measures can give rise to several different models.

Simple correlations for each of the 12 exogenous variables with each of the 12 endogenous variables were obtained in order to get a "feel for the data." Table 4 shows the simple correlations with Y_7 through Y_{12}.

As can be seen from Table 4, among all 12 exogenous variables, ADT on the cross route correlated highest with highway oriented development (.327) and total development (.421), while population density change and change in market value of real estate of the local community had the two highest correlations with non-highway oriented development (.446 and .459, respectively). These coefficients are all statistically significant at the one percent level.

Table 4 - Simple correlation coefficients for each of the 12 exogenous variables with each of the six endogenous variables (Y_7, Y_8, Y_9, Y_{10}, Y_{11}, Y_{12}), 1975 data level

	Exogenous Variable	Y_7 Highway Oriented	Y_8 Public	Y_9 Industrial	Y_{10} Other Commercial	Y_{11} Non-Highway oriented	Y_{12} Total Development
X_1	Average Daily Traffic - Cross Route	.327**	.172	.253**	.294**	.348**	.421**
X_2	Average Daily Traffic – Interstate	.016	.069	.453**	.340**	.443**	.282**
X_3	Community Zoning	.101	−.006	.089	.075	.091	.120
X_4	Age of Interchange	−.087	.060	.320**	.223*	.300**	.129
X_5	Distance to Nearest Urban Area	−.100	−.093	−.167	−.210*	−.245	−.211*
X_6	Population Density – Local Community	−.021	.061	.297**	.199*	.273**	.154
X_7	Population Density Change – Local Community	.047	.127	.359**	.374**	.446**	.303**
X_8	Market Value of Real Estate Per Capita – Local Community	.070	−.021	.024	.049	.046	.073
X_9	Change in Market Value Per Sq. Mi. – Local Community	.081	.138	.402**	.370**	.459**	.333**
X_{10}	County Population Density	−.111	.079	.313**	.128	.224*	.066
X_{11}	County Population Density Change	−.168	.214*	.174*	.125	.189**	.009
X_{12}	County Market Value of Real Estate Per Sq. Mi.	−.083	.105	.344*	.206*	.301*	.132

* Coefficient is statistically significant at the 5% level.
** Coefficient is statistically significant at the 1% level.

It is also interesting to note that the age of interchange had virtually no correlation with highway oriented development, but did correlate significantly with non-highway oriented development (.300). Perhaps this is explainable by the fact that the minimum age of the interchanges using the 1975 data is six years with much of the highway oriented development occurring sooner and more rapidly than the public, industrial and other commercial developments. If recently constructed interchanges had been part of the study, one normally would have expected the age of the interchange to also correlate with highway oriented developments.

The highest and only significant correlation with public development was county population density change. Industrial development correlated significantly with 9 of the 12 exogenous variables, the two highest of which were average daily traffic on the interstate (.453) and market value of real estate per capita of the local community (.402). Other significant variables included population density change - local community age of interchange, county market value and county population density. Other commercial development was significantly correlated with 8 of the 12 variables; the highest correlation was with population density change in the local community (.374). As expected, the only negative correlation involved distance to the nearest urban center. This coefficient, (−.210), was not very high but did indicate an inverse relationship between the two variables and was statistically significant at the five percent level.

Tables similar to Table 4 were also constructed for the highway oriented endogenous variables $Y_1, Y_2, \ldots, Y_6$ but are not presented here for the sake of brevity. The results closely followed those of Y_7 (total highway oriented development) with ADT on the cross route being the most highly significant variable in each case.

For each endogenous variable in Table 4, the 12 simple correlation coefficients were ranked in order of magnitude. The results are presented in Table 5. Some variables ranked consistently high or low depending upon whether highway orientated or non-highway oriented development was considered (e.g. population density change and market value per square mile of the local community, community zoning, and county population density).

Table 5 - The simple correlation coefficients in Table 4 ranked for each endogenous variable, 1975 data levels

	Exogenous Variable	Y_7 Highway Oriented	Y_8 Public	Y_9 Industrial	Y_{10} Other Commercial	Y_{11} Non-Highway Oriented	Y_{12} Total Development
X_1	Average Daily Traffic – Cross Route	1	2	8	4	4	1
X_2	Average Daily Traffic – Interstate	12	8	1	3	3	4
X_3	Community Zoning	4	12	11	11	11	9
X_4	Age of Interchange	6	10	5	5	6	8
X^5	Distance to Nearest Urban Area	5	6	10	6	8	5
X_6	Population Density – Local Community	11	9	7	8	7	6
X_7	Population Density Change – Local Community	10	4	3	1	2	3
X_8	Market Value of Real Estate Per Capita – Local Community	9	11	12	12	12	10
X_9	Change in Market Value Per Sq. Mi. – Local Community	8	3	2	2	1	2
X_{10}	County Population Density	3	7	6	9	9	11
X_{11}	County Population Density Change	2	1	9	10	10	12
X_{12}	County Market Value of Real Estate Per Sq. mi.	7	5	4	7	5	7

Table 6 - Simple correlation coefficients for each of 3 exogenous variables with each of the 6 endogenous variables 1970 and 1975 data levels

	Endogenous Variable	Exogenous Variable X_1 ADT-Cross route 1975	1970	X_2 ADT-Interstate 1975	1970	X_4 Age of Interchange 1975	1970
Y_7	Highway Oriented	.33**	.25**	.02	.22**	–.09	.40**
Y_8	Public	.17*	.09	.07	.09	.06	.14
Y_9	Industrial	.25**	.17*	.45**	.41**	.32**	.35**
Y_{10}	Other Commercial	.29**	.23**	.34**	.34**	.22*	.33**
Y_{11}	Total Non-Highway	.35**	.26**	.44**	.47**	.30**	.43**
Y_{12}	Total Development	.42**	.31**	.28**	.40**	.13	.50**

* Coefficient is statistically significant at the 5% level.
** Coefficient is statistically significant at the 1% level.

A few variables ranked relatively consistent across all categories of development (e.g. county market value of real estate square mile, distance to the nearest urban center, and market value per capita of the local community). Still other variables such as ADT on the cross route and ADT on the interstate highway were less predictable. While it was not surprising to find that average daily traffic on the cross route correlated highest with highway oriented development, it was somewhat unexpected to find that average daily traffic on the interstate correlated the least among the 12 exogenous variables.

In order to see the extent to which the simple correlation coefficients and subsequent rankings would differ over time, additional preliminary analysis was conducted involving data levels for the year 1970. Remarkably, most of the coefficients and rankings remained consistent from 1970 to 1975. Three notable exceptions were the age of interchange, ADT on the interstate, and ADT on the cross route. Table 6 gives the simple correlations of these exogenous variables with each of the 6 endogenous variables (Y_7 through Y_{12}) for 1970 and 1975 data levels.

As can be seen from Table 6, all correlations with average daily traffic on the cross route were less in 1970 than in 1975. This caused the rankings for this variable to slip somewhat using 1970 data (see Table 7). The most notable changes were a drop in rank from 1 to 2 in the correlation with highway oriented development, and from 1 to 4 when correlated with total development.

Average daily traffic on the interstate highway, on the other hand, improved its position on both an absolute and relative basis when compared to 1970 data. The correlation of this variable with highway oriented development in 1970 was $r = .22$ as compared with $r = .02$ in 1975, thus moving from its rank of 12 in 1975 to a rank of 3 using 1970 data levels. It is interesting to note also that when comparing the simple correlations between the two ADT variables and total development Y_{12}, the coefficients in 1970 for the two variables are almost the reverse of those in 1975.

These results suggest that the ADT on the interstate is more significant in determining the level of development around an interchange in its earlier stages, but that as an interchange grows and matures, the ADT on the cross route is a more significant factor in determining the degree of economic development.

As can also be seen in Table 6 using the 1970 data levels for the 128 complete interchanges, the age of the interchange is significantly correlated with all endogenous variables (except the one measuring public development).

Using the 1975 data levels at these same interchanges resulted in a drop in all correlations, with total development and highway oriented development showing very little correlation. These results are also reflected in the rankings shown in Table 7.

This suggests that for interchanges which have not been opened for an extended period of time, the age of the interchange might be an important factor in explaining the degree of development. For older interchanges, age becomes less important, expecially for explaining highway oriented service development.

As a next step in getting a "feel for the data" and to aid in the construction of a forecasting model, a multiple linear regression analysis was conducted by regressing each endogenous variable against all 12 exogenous variables. This was done for three different time periods or data levels, 1970, 1973, and 1975, in order to see what the resulting effect would be in the proportion of explained variations, number of statistically significant variables, and regression coefficients.

Table 7 - Comparisons of the rankings of the simple correlation coefficients of the three exogenous variables using 1970 and 1975 data levels

	Endogenous Variable	Exogenous Variable X_1 ADT-Cross Route 1975 Rank	1970 Rank	X_2 ADT-Interstate 1975 Rank	1970 Rank	X_4 Age of Interchange 1975 Rank	1970 Rank
Y_7	Highway Oriented	1	2	12	3	6	1
Y_8	Public	2	5	8	4	10	2
Y_9	Industrial	8	7	1	1	5	2
Y_{10}	Other Commercial	4	6	3	3	5	4
Y_{11}	Total Non-Highway	4	6	3	1	6	2
Y_{12}	Total Development	1	4	4	2	8	1

Table 8 - Statistical summary of 12 multiple linear regressions,
$Y_i = f(X_1, X_2, ..., X_{12})$ i = 1, 2,..., 12 for 1970, 1973, 1975 data levels

	Endogenous Variable	Proportion of Explained Variance R^2			Number Statistically Significant Variables**		
		1970	1973	1975	1970	1973	1975
SERVICE STATION DEVELOPMENTS							
Y_1	Number of Stations	.220	.202	.212	5	5	8
Y_2	Number of Gas Pumps	.200	.199	.204	9	4	9
RESTAURANT DEVELOPMENTS							
Y_3	Number of Restaurants	.284	.158	.153	8	6	5
Y_4	Number of Seats	.351	.231	.191	8	9	7
MOTEL DEVELOPMENTS							
Y_5	Number of Motels	.221	.231	.225	6	5	5
Y_6	Number of Rooms	.386	.269	.245	7	4	5
TOTAL HIGHWAY ORIENTED SERVICES							
Y_7	Highway Oriented Developments	.297	.214	.203	10	7	5
NON-HIGHWAY ORIENTED DEVELOPMENT							
Y_8	Public Developments	.097	.082	.093	3	2	2
Y_9	Industrial Developments	.233	.230	.288	5	3	4
Y_{10}	Other Commercial Establishments	.215	.309	.260	4	9	10
TOTAL NON-HIGHWAY ORIENTED DEVELOPMENT							
Y_{11}	Non-Highway Oriented Developments (sum of Y_8, Y_9, Y_{10})	.301	.348	.354	4	10	9
TOTAL DEVELOPMENT							
Y_{12}	Non-Highway and Highway Oriented (sum of Y_7 and Y_{11})	.356	.287	.262	7	6	6

* Obtained from multiple linear regression of the respective endogenous variable on all 12 exogenous variables.
** 10% level of significance.

Some of the results are summarized in Table 8 which gives each endogenous variable, the proportion of explained variation R^2, and the number of statistically significant variables when all 12 exogenous variables were used in each linear equation. The results were not especially encouraging since, in most cases, only 20 to 30 percent of the variations in the endogenous variables could be explained by all the exogenous variables. The lowest proportion of explained variation occurred with public development. This had been somewhat anticipated on the basis of the simple correlation analysis.

For the 1973 and 1975 data levels, the number of statistically exogenous variables in the multiple regressions ranged from a low of two for public development to a high of ten of the possible 12 for the non-highway oriented (other) commercial establishments in 1975, and total non-highway oriented development in 1973. For the 1970 data, the range was a low of 3 for public development to a high of 10 for highway oriented development. As can also be seen from Table 8, the proportion of explained variation for total highway oriented developments for the years 1970, 1973, and 1975 were .297, .214, and .203, respectively; the corresponding figures for total non-highway oriented developments were .301, .348, and .354.

Regression coefficients associated with the exogenous variables are not presented here for the sake of brevity, but for the most part, also varied in direct proportion to the variation in degree of explained variation. That is, as more variability occurred in the degree of explained variation from 1970 to 1975, greater differences occurred among the corresponding regression coefficients.

These results suggested that if the current variables were to be used in constructing a static growth model and in estimating its structural parameters, great care would have to be given to defining the population of interchanges for which the model might be applicable, especially with regard to age. Even so, it appears that such a model would not yield very great predictive ability on the basis of the explained variations in the preliminary ordinary multiple regression analysis.

Hence, other attempts at obtaining a higher degree of explained variation are being made before attempting to specify an entire structural model. One of these attempts involves the transformation of existing variables. For example, actual distance to the nearest urban center may have better been measured as the logarithm of distance or as the square root of distance. These as well as other logarithmic transformations of the variables are being conducted, along with the addition of some new variables.

Realizing the many pitfalls in arbitrary transformations of data and in using stepwise regression analysis to find significant variables, a prime consideration in the addition of a new variable or modification of existing variables is that a logical theoretical basis be established for such changes. In this regard, a new variable, which at this preliminary stage has shown promise in significantly increasing the proportion of explained variation, is population of the nearest urban center divided by the distance to the nearest urban center.

Another consideration, in the addition of new variables, is that either secondary data sources are readily available for their measurement or primary data collection is relatively easy, as for example, in measuring distances to nearby urban centers.

Further refinements and analyses are still in progress; hence, final results of the preliminary analysis cannot be presented at this time. However, the work thus far is encouraging and it is felt that a model can be developed that will be usable by planners and highway designers.

SUMMARY AND CONCLUSIONS

This paper presents the results of an attempt to develop a method for forecasting the type and intensity of economic development at non-urban interstate highway interchanges. Such information could be used by planners and highway designers to avoid premature obsolescence of highway facilities.

Some of the difficulties encountered in developing such a model involve proper definition of the interchange area or "interchange community" and the "economic growth" of the interchange community. Another problem involves the identification and quantification of certain factors that cause differences in the economic growth of interchange communities.

Still another very practical and important consideration is that the model be designed such that it can be implemented easily using secondary data. Although a model that requires an extensive primary data collection effort may have greater predictability for a specific location, it would have very limited general value.

The "interchange community" was considered as that area located within one-half mile of the interchange. The "economic growth" of an interchange community was considered to be determined by two types of factors

1) endogenous variables, i.e., those variables that describe the state or levels of economic development for the given community, and

2) exogenous variables, i.e., those variables that affect the level of economic development of the interchange community but which are not, in general, affected by the growth that takes place.

A general theoretical simultaneous equation model is discussed together with its assumptions and limitations.

In order to aid in the development of a particular structure of the general model, a preliminary analysis was conducted utilizing a sample of 144 non-urban interchanges on Pennsylvania's interstate highways. The preliminary analysis involved classification of developments into various types and performing simple and multiple linear correlation analyses. Twelve endogenous and twelve exogenous variables were analyzed.

In particular the analysis revealed that the exogenous variables differ in their influence over time as well as across various types of economic developments. Data levels for the years 1970, 1973, and 1975 were utilized in a multiple linear regression to see how much fluctuation would occur in the proportion of explained variations and the regression coefficients over time. In some instances, the results revealed strong differences between the earlier and later time periods.

Because of the low levels of explained variation, further preliminary analyses involving transformation of existing variables and creation of some new variables is currently being conducted. Some preliminary analysis of these further refinements have shown considerable improvement.

Once the refined static model has been tested empirically, it is hoped that for short run purposes, it may be able to provide a "rough" approximation for forecasting potential development. In the long run, however, a dynamic model, even a somewhat crude one, is probably most appropriate. The complexity involved in designing and determining the proper structure for such a dynamic model will be greatly aided by the analysis conducted in this on-going investigation.

REFERENCES

[1] United States Congress. 1956. **Federal-Aid Highway Act of 1956.** Public Law 627, Title I Section 108, 84th Congress, 2nd Session. Washington: U.S. Government Printer.

[2] Levin, D.R. 1961. "The Highway Interchange Land-Use Problem." **Land Use and Development at Highway Interchanges.** HRB Bull. 288, p. 1. Washington, D.C.: NRC, National Academy of Sciences.

[3] Ferber, R. and P. J. Verfoon. 1962. **Research Methods in Economics and Business.** p. 407. New York: Macmillan Company.

[4] Hood, W. C. and T.C. Koopmans. 1961. **Studies in Econometric Methods.** p. 112. New York: Wiley.

[5] Johnston, J. 1972. **Econometric Methods.** p. 376. New York: McGraw Hill.

[6] Pendleton, W.C. and R.D. Vlasin. 1961. "The Impact of Interstate Highways in Rural Areas, "**Engineering Bulletin of Purdue University** (Proceedings of the 47th Annual Road School, March 27-30, 1961). p. 106. Lafayette, Indiana: Purdue University.

[7] Lehr, R.L. **et al.** 1965. **The Relationship of the Highway Interchange and the Use of Land in the State of Oklahoma.** Pt. I, p. I-6. Oklahoma City: University of Oklahoma Research Institute.

[8] United States Department of Transportation. 1976. **Social and Economic Effects of Highways.** Washington, D.C.: Federal Highway Administration, USDOT.

[9] Malinvaud, E. 1966. **Statistical Methods of Econometrics.** P. 70. Chicago: Rand McNally.

[10] Goldberger, A. 1964. **Econometric Theory.** p. 329. New York: Wiley.