

# Stepwise Regression Model of Development at Nonmetropolitan Interchanges

HENRY E. MOON, JR.

---

After a brief outburst of research on the Interstate highway system soon after its construction, the network and its varied and wide-ranging aptitude for reshaping the nonurban United States were bypassed as major points of contention by social scientists. Land use conversion brought about by the system is a logical measure of the network's microlevel impact on rural and nonmetropolitan areas and is the focal point of this paper. This study features a stepwise regression model that explains over 50 percent of the variation in interchange area development at 65 nonmetropolitan Kentucky interchanges. The procedure relies on four independent variables (preconstruction development, interchange type, whether the site permits the legal sale of alcoholic beverages, and whether the interchange is in Appalachia) to predict the dependent variable, which is current development. The results of this study indicate the widely diverse development process found around nonmetropolitan Interstate interchanges. Both the theoretical and applied aspects of transportation-related development are considered by this model.

---

After a brief outburst of research on the Interstate highway system soon after its construction, the network and its varied and wide-ranging aptitude for reshaping the nonurban United States were bypassed as major points of contention by social scientists. Land use conversion brought about by the system is a logical measure of the system's microlevel impact on rural and nonmetropolitan areas, but this factor is minimized in the literature. This undertreatment is true for nonmetropolitan areas in general. The current study seeks to complement the growing literature on urban land use and conversion with an examination of intensified change in nonurban land use. In addition, the paper presents a model of variable development in interchange areas.

Historically, the Interstate highway system has been an important element in the transportation engineering literature (1). Most efforts have concentrated on the design and traffic aspects of the network, but some have addressed the land use issue (2-4). Most studies have focused primarily on urban areas. As concentrations and centers of impact, nonmetropolitan interchanges captured some of the early attention of researchers who were interested in the Interstate network. More recently, however, these interchanges have drawn a minimum amount of attention relative to their innate ability to reshape landscapes on a large scale. Because of this limited amount of investigation, interchanges and their highly variable potential for changing land use are misunderstood throughout

the academic, planning, and professional communities. The impact has been viewed as a function of traffic volume and distance from urban places by those attempting to understand the process of transportation-related development. None of the investigations into the variation of changes in interchange-specific land use have resulted in a satisfactory model (5, 6). This study established the complexity of the land use modification process that occurs at nonmetropolitan interchanges by using site and situational characteristics to both explain and forecast interchange development.

## METHODOLOGY

Interchange construction and development are indicators of and significant factors in the land use conversion process and are the focus of this paper. To understand interchange development it is necessary to (a) define and measure the developmental complexity of a set of nonmetropolitan interchanges; (b) identify variables that contribute to, retard, or alter the measured complexity; and (c) test the worth of the independent factors in explaining variable interchange development by building a land use conversion model.

Kentucky is an ideal site for an analysis of nonmetropolitan interchanges. Since local Interstate highway construction began in Jefferson County in 1956, total state mileage has since grown to 738. In Kentucky the Interstate system constitutes 1.1 percent of the state's total highway mileage while it carries 23 percent of the traffic (a figure almost identical to that of the national system). Five highways, 40 counties, and all of Kentucky's major cities are incorporated into the national network. Although the system's function is to connect major metropolitan regions, its routes pass through rural areas lying between these nodal cities, providing the potential for direct, high-speed access to or through places that might previously have been remote and relatively inaccessible.

Because the focus of this study is nonmetropolitan counties with Interstate highway interchanges, only the counties that are categorized as nonmetropolitan by the U.S. Bureau of the Census and contain one or more interchanges are included. Every Interstate highway in Kentucky crosses through nonmetropolitan counties, providing a set of 65 interchanges for analysis (Figure 1). Interchanges are defined as points at which traffic can enter or exit the Interstate highway from or to another road. Because of the regional, temporal, directional, and design differences among Kentucky's Interstate highways, a study based on the state's nonmetropolitan interchanges facilitates a comparison with nonmetropolitan interchanges in other states, especially those east of the Mississippi River.

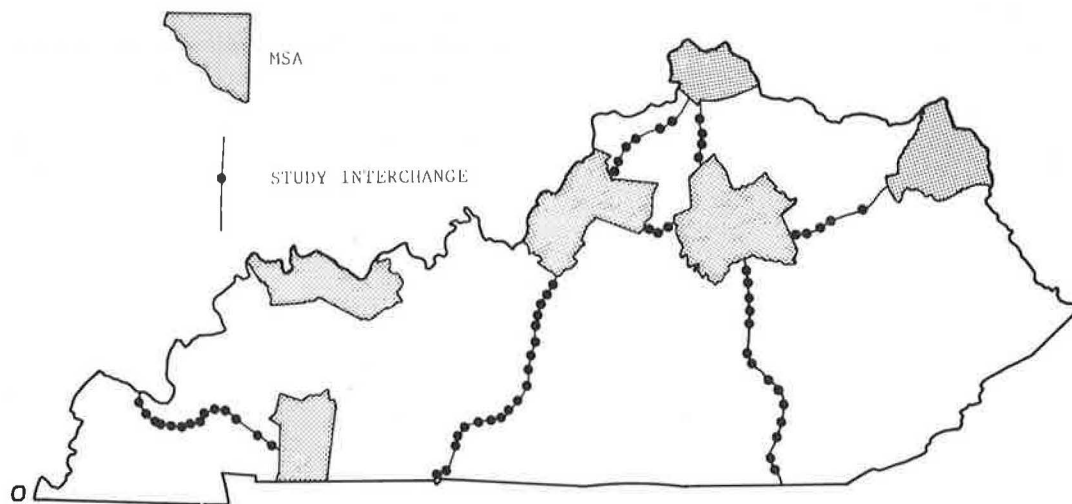


FIGURE 1 The Interstate highway system through Kentucky.

A 502-acre area around each of the 65 interchanges in Kentucky was studied (7). This size was selected for the study area because of its proven significance in an earlier analysis and its successful prestudy testing in Kentucky. R. D. Twark discovered that most new economic development at Interstate highway interchanges occurs within 0.5 mi of the intersection. For that reason, he defined the area within that distance of the crossroads as the "interchange community" and emphasized its usefulness as a study region. This particular acreage, then, is calculated from a 1-mi-diameter circle centered on the interchange. The reason for the choice of a circular study area is that linear strip analysis tends to deemphasize development on roads parallel to the Interstate highway. Furthermore, square or rectangular study areas overemphasize the corners, which are removed from the immediate influence of the interchange.

This analysis utilizes a simple, straightforward definition of development that is consistent with that found in the related literature. In this case, development is a cumulative term, indicating the presence of actual structures at an interchange site that possess both a structural and an activity component. Developmental complexity incorporates the size and scope of entities located at interchange sites and a measure of the activity that the particular facilities generate. Because structures may include a wide variety of patterns, sizes, and uses, a simple count of the individual buildings in each study area is inappropriate and misleading (i.e., a tobacco barn is not equal to a convenience store in size or in activity generated). Therefore buildings are divided into size categories on the basis of their size and ability to generate traffic: simple, nonresidential; single-family residential; multifamily residential; small commercial and small institutional; large commercial, large institutional, and small industrial; and large industrial (more than 20,000 ft<sup>2</sup>). Building numbers recorded by category in the field at each interchange are weighted on the basis of the buildings' size and ability to generate traffic. On the basis of extensive field checking and trial-and-error-type runs, the weightings selected for the six building categories are

- 1: Category I (simple, nonresidential);
- 2: Category II (single-family residential);
- 4: Category III (multifamily residential);

- 8: Category IV (small commercial and small institutional);
- 16: Category V (large commercial, large institutional, and small industrial); and
- 32: Category VI (large industrial).

The developmental complexity of each study area is calculated by summing the number of building present in each category and multiplying by the appropriate category weight, then totaling the category sums. For example, if a particular study area contains 20 single-family homes, 5 barns, and 4 service stations, its developmental complexity score is

$$[(20 \times 2) + (5 \times 1) + (4 \times 8)] = 77$$

This number, representing the level of land use conversion or construction at each interchange, is the dependent variable recorded at the 65 observation sites in this study.

The impact of Interstate highway interchange construction can be expected to vary from place to place, depending on a variety of local and site-specific characteristics. According to the literature, traffic volume, interchange type and age, topography, and distances to urban areas are factors that influence developmental complexity. Yet these factors have proven incapable of explaining variation in interchange development. A conclusive approach apparently involves the examination of a broader and more comprehensive set of variables. Variables that measure the historical, spatial, economic, population, and social attributes of a place must be included.

Independent variables of land use, engineering, traffic, social, regional, population, and geographic natures were introduced into the analysis. For each study interchange the following data were collected. Variable 1 serves as the dependent variable, whereas variables 2–31 serve as independent variables in this analysis:

- 1. Current developmental complexity,
- 2. Average daily traffic count,
- 3. Interchange age,
- 4. On a north-south highway,
- 5. In the Bluegrass Region,
- 6. In the Pennyroyal Region,

7. In the Jackson Purchase Region,
8. On a state highway,
9. On a federal highway,
10. Number of preconstruction land owners,
11. Cloverleaf type,
12. Diamond type,
13. Half type,
14. Leg type,
15. Double type,
16. Diamond/cloverleaf type,
17. Distance to the nearest neighboring interchange,
18. Distance to the farthest neighboring interchange,
19. Exits to a major tourist attraction,
20. Distance to the nearest city,
21. Distance to the nearest Metropolitan Statistical Area (MSA),
22. Distance to the nearest city with a population greater than 25,000,
23. Number of workers commuting into the county,
24. Number of workers commuting out of the county,
25. County population,
26. Percentage of the county's population that is urban,
27. Preconstruction developmental complexity,
28. Dominant soil capability classification at the interchange,
29. Areal size of the developable portion of the interchange study area,
30. Whether the area restricts the legal sale of alcoholic beverages, and
31. Whether the interchange operates within Appalachia (a measure of latent demand).

Variables 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, and 19 are entered as dummy variables. For example, if an interchange connects an Interstate highway and a county road, the values introduced at that particular observation for variables number 8 and 9 would be 0. The effect of that characteristic would be felt in the constant or *Y*-intercept, an especially critical figure in this type of analysis because it holds the combined effect of all null dummy variable measures. Had the intersecting road in question been a federal highway, variable number 9 for that observation would have a value of 1.

The SPSS-X stepwise regression procedure combines forward and backward entry to construct a model. The routine introduces an initial variable into the regression equation according to its explanatory power, then searches for a variable that will complement the previously selected variable and increase the multivariate correlation coefficient. Throughout the model-building process, this technique searches the matrix of correlation coefficients to find the minimum number of variables that provides the maximum amount of explained variation. A significant advantage that this more advanced stepwise procedure holds over prior techniques is its ability to emulate a backward entry procedure by removing once-entered variables from the equation to achieve the optimal equation. This ideal end product is a combination of variables that fit together in a way that maximizes explanatory power and minimizes collinearity within the model. By addressing the collinearity problem, this method greatly reduces the tendency that many regression techniques have toward explaining only a

small portion of the total variation in the dependent variable. This technique selects variables for entry into the equation on the basis of their ability to explain a different portion of overall variation. As the primary ingredient in this methodology, stepwise regression allows a common and easily reproducible *modus operandi* for scholars and planners alike.

## MEASURED COMPLEXITY

Among the 65 study interchanges, developmental complexity scores range from 13 at one Marshall County interchange to 1,616 at an interchange in McCracken County. The mean score among all study interchanges is 323.815. The typical interchange study area of 502 acres contains over 102 structures, including all types of both residential and nonresidential buildings. Most dominant among building types are single-family residences, whereas industrial facilities are most infrequent. As percentages of total interchange complexity, the following category values indicate the level of influence that each structure type carries in the overall analysis:

- 40.94 percent: Category IV (small commercial and small institutional);
- 29.29 percent: Category II (single-family residential);
- 14.82 percent: Category V (large commercial, large institutional, and small industrial);
- 10.08 percent: Category I (simple, nonresidential);
- 3.50 percent: Category III (multifamily residential); and
- 1.37 percent: Category VI (large industrial).

The relationships between the dependent variable and the set of independent variables as pictured in the matrix of correlation coefficients reveal a complex picture of the factors that influence interchange development (Table 1). For example, the strongest relationship between developmental complexity and an independent variable involves preconstruction complexity ( $r = 0.563$ ). Other variables that exhibit relatively strong positive relationships with developmental complexity are number of preconstruction land owners ( $r = 0.368$ ) and number of out commuters ( $r = 0.362$ ). Among the more significant detriments to development that can be identified from the matrix are location in the Pennyroyal region, increased distance to cities of any size, and a more limited topography. An interchange's relative proximity to a city of any size provides the strongest negative relationship in the data set,  $-0.358$ , because complexity decreases as distance increases. Interchange areas near cities are generally more developmentally complex than those further removed from the influence of urban places. These figures alone only hint at the overall development process, for they merely suggest simplified measures of influence while holding other factors constant. From the correlation coefficients, no single independent variable can be identified as a satisfactorily strong factor in development, and collinearity is highly visible in certain areas of the matrix. These two characteristics of the matrix are resolved by the stepwise regression technique while it searches for multivariate influence and reduces collinearity.

TABLE 1 CORRELATIONS BETWEEN INDEPENDENT VARIABLES AND THE DEPENDENT VARIABLE (Y)

Independent Variable	Correlation With Y
2. Average daily traffic count	0.32
3. Interchange age	0.07
4. On a north-south highway	-0.01
5. In the Bluegrass Region	0.01
6. In the Pennyroyal Region	-0.26
7. In the Jackson Purchase Region	0.23
8. On a state highway	0.03
9. On a federal highway	0.08
10. Number of preconstruction land owners	0.37
11. Cloverleaf type	-0.12
12. Diamond type	0.10
13. Half type	-0.07
14. Leg type	-0.14
15. Double type	0.30
16. Diamond/cloverleaf type	0.06
17. Distance to the nearest neighboring interchange	0.00
18. Distance to the farthest neighboring interchange	-0.06
19. Exits to a major tourist attraction	-0.11
20. Distance to the nearest city	-0.36
21. Distance to the nearest MSA	0.15
22. Distance to the nearest city > 25,000	-0.07
23. Number of workers commuting into the county	0.27
24. Number of workers commuting out of the county	0.36
25. County population	0.32
26. Percent of the county's population that is urban	0.35
27. Preconstruction developmental complexity	0.56
28. Dominant soil capability classification	-0.29
29. Size of the developable part of the study area	-0.02
30. Whether the area restricts the legal sale of alcoholic beverages	0.33
31. Whether the interchange is in Appalachia (a measure of latent demand)	0.20

## THE MODEL

The form of the stepwise regression model utilized here is as follows:

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_4$$

where

- $Y$  = expected developmental complexity;
- $a$  = constant or  $Y$ -intercept (particularly important in this analysis because it contains the joint effect of all null dummy variable measures);
- $b_{1...4}$  = regression coefficients accompanying each variable; and
- $X_{1...4}$  = measured value of each  $X$  at the proposed interchange site.

Because preconstruction developmental complexity has a stronger relationship with the dependent variable (current developmental complexity), it is selected by the procedure for initial entry into the regression equation. Preconstruction developmental complexity is measured in the same way as the current complexity. This historical measure identifies the level of development in a study area prior to highway placement and is drawn from preconstruction aerial photographs. The mean preconstruction developmental complexity score for the 65 nonmetropolitan sites is 49.94. Unlike current developmental complexity, preconstruction complexity is most influenced by

single-family residences and not commercial or institutional establishments. Of the total complexity measure at an average interchange, 57.55 percent can be attributed to single-family housing, as compared to less than 30 percent of current scores. A typical interchange study area contained 28.13 structures prior to construction of the Interstate highway through the area. In general, the 502-acre site was the location of 14.37 single-family residences; 12.77 simple, nonresidential structures; 0.71 small commercial or small institutional facilities; 0.14 multi-family residential buildings; 0.14 large commercial, large institutional, or small industrial operations; and no large industrial enterprises. Preconstruction developmental complexity alone explains 31.66 percent of the variation in current complexity.

Second, the procedure enters the double interchange variable as the next independent explanatory factor. The technique omits the ownership variable in spite of the fact that its  $r$  score is greater than that of the chosen variable. Its omission is based on the strength of its relationship (0.684) with the previously selected variable (preconstruction developmental complexity). This is an excellent example of the procedure's ability to minimize collinearity in the equation. Seven geometrically distinct interchange types are currently open to Interstate traffic in nonmetropolitan Kentucky. Of the 65 study interchanges, 55 are diamond-shaped, five are trumpet-shaped, three are leg or M-shaped, two are halves of double interchanges, two are diamond/cloverleaf combined, one is a cloverleaf, and one is a half-diamond. Interchanges that serve as halves or ends of a double interchange encourage large-scale development, whereas trumpet interchanges discourage land use change. Double interchanges generate and are surrounded by 170 percent more development than the mean figure. Variable levels of access and visibility to goods and services at double and trumpet interchanges result in these higher levels of investment. With the addition of this second variable, the model's explanatory power is increased to 39.32 percent.

Third, the procedure enters the "wet" variable (i.e., variable 30, which indicates restriction on alcohol sales) into the equation to produce an explanatory capability of 46.10 percent. Social preferences of surrounding populations retard or spur interchange development. Of the 65 study interchanges, only 19 are legally categorized as "wet" (i.e., allowing sale of alcohol) while the remaining 46 forbid the sale of alcoholic beverages, including beer and wine. Wet interchange areas are 50 percent more developed than their dry counterparts.

Finally, a fourth independent variable, an interchange's presence in Appalachia, is entered into the equation. Of the 65 study interchanges, 19 lie in the part of Kentucky so identified, whereas 46 fall outside the region. Of the 19, 5 are located along Interstate 64, east of Lexington, and the remaining 12 lie south of the same city on Interstate 75. One of these southern interchanges connects with the Appalachian Development Highway System at London, Kentucky, providing impetus for interaction between and within Appalachia. Appalachian interchange areas are almost 20 percent more developed than their non-Appalachian counterparts.

Given the results of the stepwise procedure above, the following values would be introduced into the equation:

$$a = -33.098874,$$



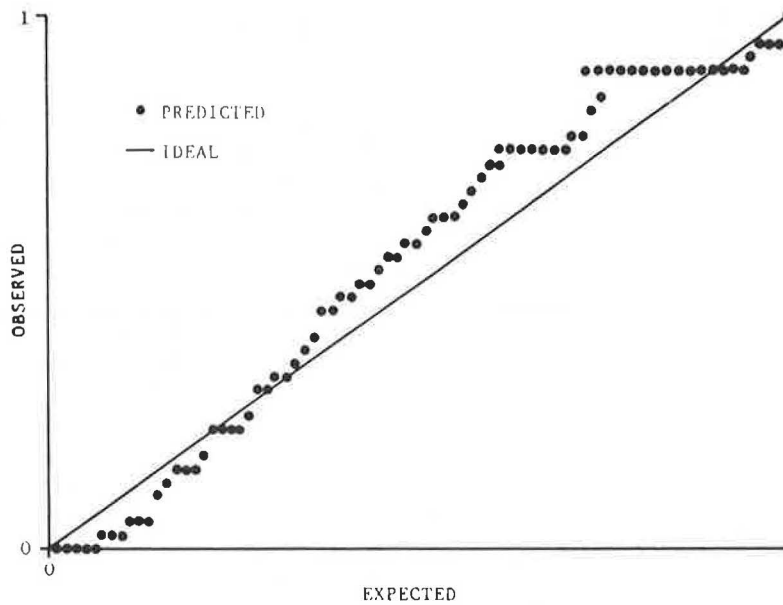


FIGURE 2 Plot of standardization residuals (65 observations).

$$\begin{aligned}
 b_1 &= 5.387014, \\
 b_2 &= -1,025.030050, \\
 b_3 &= 214.671303, \text{ and} \\
 b_4 &= 193.920165.
 \end{aligned}$$

The model has an overall explanatory capability of 53.07 percent (Figure 2). Both logarithmic and square root data transformations failed to significantly improve explanatory power.

#### SUMMARY AND CONCLUSION

The land use conversion process documented in this study is a complex one. Whereas prior studies attempted to explain and forecast interchange development with a limited number of independent variables, the model presented here utilizes 4 variables drawn from a list of 30 independent variables. These results should encourage planners to recognize the role of nonsystem factors in development and include them in the transportation planning process. Land use conversion is a complicated process, driven by a variety of variables at work in different spatial circumstances.

#### REFERENCES

1. W. L. Garrison. Supply and Demand for Land at Highway Interchanges. *Bulletin 288*, HRB, National Research Council, Washington, D.C., 1961, pp. 61-66.
2. P. D. Cribbins, W. T. Hill, and H. O. Seagraves. Economic Impact of Selected Sections of Interstate Routes on Land Value and Use. In *Highway Research Record 75*, HRB, National Research Council, Washington, D.C., 1965, pp. 1-30.
3. R. H. Ashley and W. F. Berard. Interchange Development Along 180 Miles of I-94. In *Highway Research Record 96*, HRB, National Research Council, Washington, D.C., 1965, pp. 46-58.
4. F. I. Thiel. Highway Interchange Area Development. In *Highway Research Record 96*, HRB, National Research Council, Washington, D.C., 1965, pp. 24-42.
5. R. D. Twark. *A Predictive Model of Economic Developments at Non-Urban Interchange Sites on Pennsylvania Interstate Highways*. Pennsylvania State University, College Station, June 1967.
6. T. H. Eighmy and J. J. Coyle, Jr. *Toward a Simulation of Land Use for Highway Interchange Communities*. Pennsylvania State University, University Park, 1967.
7. H. E. Moon, Jr. *Modeling Land Use Change Around Interstate Highway Interchanges in Nonmetropolitan Areas: A Multivariate Statistical Analysis*. University of Kentucky, Lexington, 1986.