

Multilevel Modeling of the Local and Regional Economic Impacts of Bypasses

Joshua B. Mills^a, Research Assistant, jbmills@purdue.edu
Jon D. Fricker^a, Professor of Civil Engineering, fricker@purdue.edu

^aSchool of Civil Engineering
Purdue University
550 Stadium Mall Drive
West Lafayette IN 47907-2051
USA

March 2010

Josh Mills will be the Corresponding Author.
Phone 765-494-2205
Fax 765-496-7996

Introduction

Bypasses, which redirect through traffic around a community's downtown area, could have substantial impacts. Identifying these impacts and attributing these impacts solely to the presence of a bypass has posed a challenge to decision-makers and researchers. The economic impacts have the potential to be long-term and far-reaching. Because impacts take place over a long period of time, and because these impacts cannot be easily isolated due to economic externalities (such as recessions), researchers have found it difficult to establish a standard estimation framework for determining bypass impacts. Recent studies (Mills and Fricker 2009, for example) have used econometric and other quantitative methods to determine the extent of these economic impacts. This study will use longitudinal mixed-effects models to study the economic impacts of eight bypasses located in north central Indiana.

Study Area and Sources of Data

The study area in north central Indiana (see Figure 1) includes several major US routes (the majority of which are four-lane divided highways) and state roads (most of which are two-lane roads). A total of 65 ZIP codes, including eight communities with bypasses, comprise the study area. The study area is roughly bordered by South Bend and Elkhart to the north, Fort Wayne to the east, Kokomo and Marion to the south, and the White County-Cass County border to the west.

Employment, payroll, and establishment data for various industry sectors were gathered from Zip Business Patterns, an annual set of ZIP Code-level data published by the Census Bureau since 1994. Because true ZIP Codes do not have "polygon"-type definitions, the actual unit of aggregation for ZBP data are Zip Code Tabulation Areas (ZCTA), a unit of aggregation defined by the Census Bureau that generally conform to

the actual boundaries of ZIP Codes. Because of their basic equivalence, the terms ZCTA and ZIP Code will be used interchangeably. These data are similar to County Business Patterns (CBP), which were used in a similar study (Mills 2007), with one major difference. County Business Patterns include employment and payroll data for all industry sectors, as well as number of establishments by number of employees (e.g. 1-4 employees, 5-10 employees, etc.). Due to disclosure laws, only industry-specific establishment by employment size are available for Zip Business Patterns, although total employment and payroll figures are still available for each ZIP Code.

ZIP Code-level economic data were collected for the years 1998-2006. Other data were gathered from long-form Census 2000 population, income, education, commute, and tax data from Summary File 1 and Summary File 3. These data were converted and aggregated by SAS-based online applications developed by the Missouri Census Data Center (2008).

The study area is shown in Figure 1. The map shown in Figure 1 illustrates the advantage of using ZIP Code-level data as opposed to county-level data. While the study area encompasses several counties, the affected area consists of 65 distinct ZIP Codes. Selection of ZIP Codes for the study was based on a subjective examination of the ZIP Codes most likely to be affected by the study area.

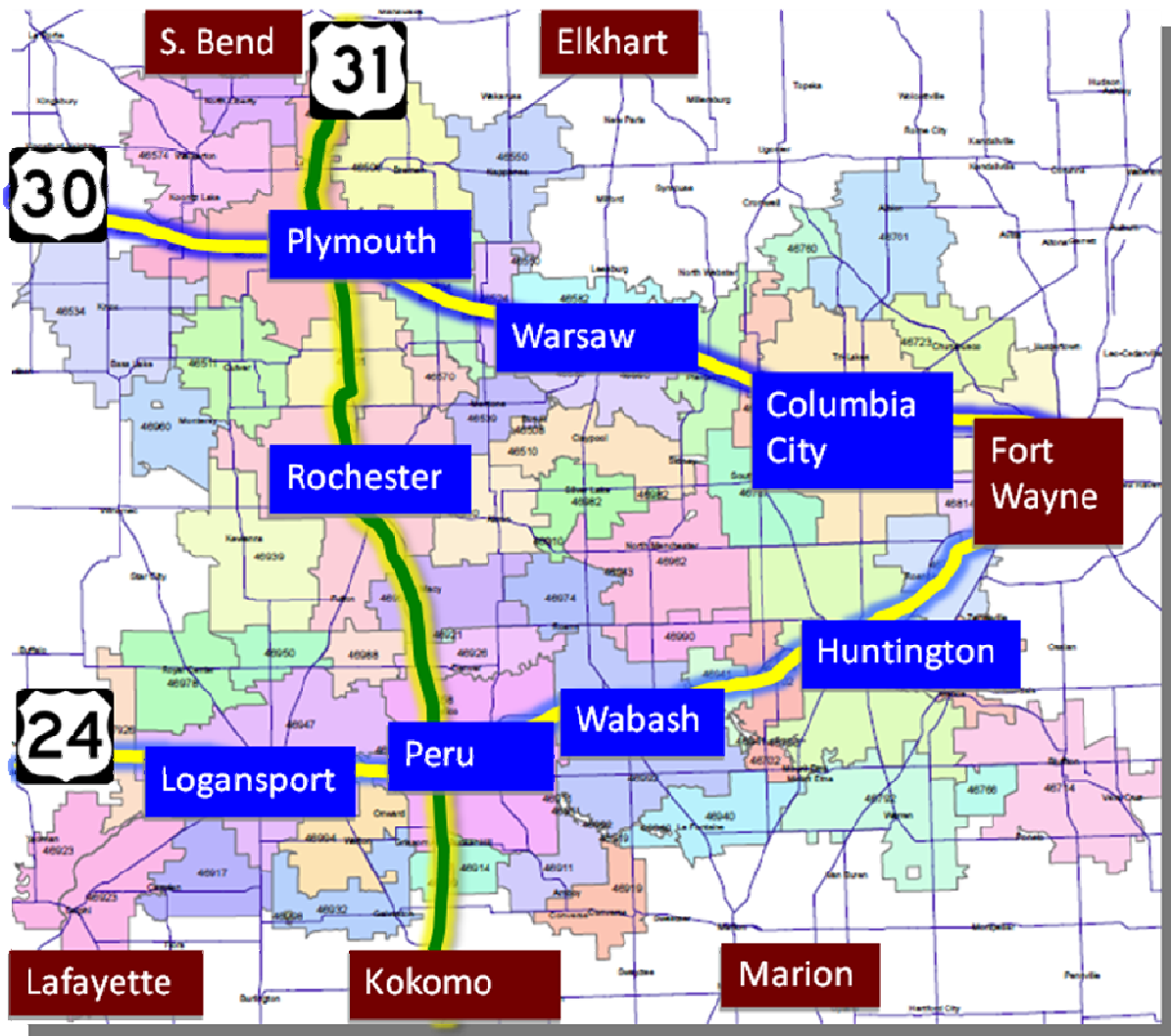


FIGURE 1 Study area selected for analysis, with principal routes. Bypassed communities are coded blue, and nearby large cities are coded red.

Population figures and percentage change over time are shown below in Table 1.

TABLE 1 Descriptive statistics for bypasses communities in study area.

| City | Population | | | % Change 1990-2000 | % Change 2000-2008 | County | Year Bypass Opened |
|------------|------------|--------|--------|-----------------------|-----------------------|------------|--------------------------|
| | 1990 | 2000 | 2008 | | | | |
| Logansport | 16,812 | 19,684 | 18,663 | 17.1% | -5.2% | Cass | 1999 |
| Peru | 12,843 | 12,994 | 12,301 | 1.2% | -5.3% | Miami | 1979 |
| Wabash | 12,127 | 11,743 | 10,815 | -3.2% | -7.9% | Wabash | 1964 |
| Huntington | 16,389 | 17,450 | 16,521 | 6.5% | -5.3% | Huntington | 1964 |
| Columbia | | | | | | | |
| City | 5,706 | 7,077 | 8,283 | 24.0% | 17.0% | Whitley | 1963 |
| Warsaw | 10,968 | 12,415 | 13,627 | 13.2% | 9.8% | Kosciusko | 1972 |
| Plymouth | 8,303 | 9,840 | 11,038 | 19.0% | 12.0% | Marshall | 1975 |
| Rochester | 5,969 | 6,414 | 6,457 | 7.0% | 1.0% | Fulton | 1975 |

Mixed-effects models can be used to capture within-group and between-group variation. These types of models are similar to panel data models in that both types of models allow for estimated coefficients to vary by individual (or cross-section) or over time period. Mixed-effects models differ from panel data models in that while panel data models focus on accounting for unobserved heterogeneity through the use of different intercepts for each cross-section or time-period, mixed-effects models allow for covariates to vary by group. The other key difference between panel data models and mixed-effects models lies in the method of estimation. Panel data models have closed-form, analytical solutions which can be solved via linear algebra. In contrast, mixed models do not have closed-form, analytical solutions and thus must be solved by nonlinear numerical optimization (Croissant and Millo 2008). While general serial correlation and heteroskedasticity can be controlled for using feasible GLS (FGLS) in the panel data framework, the mixed effects framework allows for a more general specification of heteroskedasticity (such as allowing the variance to have nonlinear covariates) and within-group serial correlation (through the use of ARMA correlation structures).

The mixed-effects model takes the form:

$$y_{it} = CITY_{it} \cdot \varphi + INDUSTRY_{it} \cdot \psi + BYPASS_{it} \cdot \xi + Z_i \gamma_i + \varepsilon_{it} \quad (1)$$

$$\begin{aligned} \gamma_i &\sim N(0, \Psi) & \varepsilon_i &\sim N(0, \Sigma) \\ i &= 1, \dots, N, t = 1, \dots, T \end{aligned} \quad (2)$$

where

y_{it} is the dependent variable,

$CITY_{it}$ represents characteristics relating to the bypassed community (such as population and the distance to the nearest large city) with associated vector of coefficients φ , $INDUSTRY_{it}$ represents characteristics of related industries, such as employment and payroll levels, with associated vector of coefficients ψ , $BYPASS_{it}$ represents characteristics of the bypass of each county seat, primarily in the form of indicators representing how many years the bypass had been open, with associated vector of coefficients ξ , Z_i represents the matrix of random effects with associated vector of coefficients γ_i , which vary by ZIP Code, and ε_{it} represents the vector of normally distributed idiosyncratic errors.

The φ , ψ , and ξ vectors represent the fixed effects, or coefficients that do not vary by cross-section or over time.

To illustrate, consider a ZIP Code that has had a bypass open for the last 43 years of data stored in the database. Assume this ZIP Code has three separate bypass age indicators, one indicator for years 31-35, another indicator for years 36-40, and another indicator for years 41-45. The indicator for years 31-35 would be set to 1 for the years in which the bypass had been open for 31-35 years, inclusive.. Thereafter, the indicator would be set equal to 0. For the next 5 years, only the age 36-40 indicator would be set to 1. This indicator would then be set to 0 after those 5 years. For the last 3 years in the database, the age 41-45 indicator would be set to 1. This brief example shows that at most one indicator is “switched on” for a given observation.

Estimation

Maximum likelihood estimation (MLE) has typically been used to estimate mixed-effects models. The MLE estimator for the variance matrix, however, has been found to be downward biased (Verbeke and Molenberghs 2001). An alternative estimator, the Restricted Maximum Likelihood (REML) estimator, corrects for this downward bias. Eq. 1 can be rewritten as

$$y_{it} = X_{it}\beta + Z\gamma_i + \varepsilon_{it} \quad (3)$$

where

$$X_{it} = \begin{bmatrix} CITY_{it} \\ INDUSTRY_{it} \\ BYPASS_{it} \end{bmatrix}$$

$$\beta = \begin{bmatrix} \varphi \\ \psi \\ \xi \end{bmatrix}$$

Let $V_i = Z_i\gamma Z_i' + \Sigma$ and α denote all coefficients to be estimated in γ and Σ . Using the specification of Laird and Ware (1982), the maximum likelihood estimator for β is:

$$\hat{\beta} = \left(\sum_{i=1}^{NT} X_{it}' V_i^{-1} X_{it} \right)^{-1} \sum_{i=1}^{NT} X_{it}' V_i^{-1} y_{it} \quad (4)$$

For p parameters, the likelihood function for restricted maximum likelihood estimation is (Verbeke and Molenberghs 2001):

$$L(\alpha) = (2\pi)^{-(NT-p)/2} \left[\sum_{i=1}^{NT} X_{it}' X_{it} \right]^{1/2} \left[\sum_{i=1}^{NT} X_{it}' V_i^{-1} X_{it} \right]^{-1/2} \left[\prod_{i=1}^{NT} |V_i|^{-1/2} \right] \cdot \exp \left\{ -\frac{1}{2} \sum_{i=1}^{NT} (y_i - X_{it}\hat{\beta})' V_i^{-1} (y_i - X_{it}\hat{\beta}) \right\} \quad (5)$$

Correcting for Autocorrelation and Heteroskedasticity

Analysis of the within-group autocorrelation functions (ACFs) for each of the models showed that significant autocorrelation was present in the models. Given the likely presence of heterogeneity within the residuals, three correlation structures for correcting for this autocorrelation with heterogeneity were considered: Heterogeneous AR(1), Toeplitz, and first-order antedependence. The form of each of these correlation structures are listed below in Table 2.

TABLE 2 Autocorrelation structures used in models

| Name | Model |
|---|--|
| Heterogeneous AR(1) (SAS Institute 2008) | $\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_2\rho^2 & \cdots & \sigma_1\sigma_T\rho^{T-1} \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \cdots & \sigma_2\sigma_T\rho^{T-2} \\ \sigma_3\sigma_1\rho^2 & \sigma_3\sigma_2\rho & \sigma_3^2 & \cdots & \sigma_3\sigma_T\rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_T\sigma_1\rho^{T-1} & \sigma_T\sigma_2\rho^{T-2} & \cdots & \sigma_T\sigma_{T-1}\rho & \sigma_T^2 \end{bmatrix}$ |
| Toeplitz (SAS Institute 2008) | $\begin{bmatrix} \sigma^2 & \sigma_1 & \sigma_2 & \cdots & \sigma_{T-1} \\ \sigma_1 & \sigma^2 & \sigma_1 & \cdots & \sigma_{T-2} \\ \sigma_2 & \sigma_1 & \sigma^2 & \cdots & \sigma_{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{T-1} & \sigma_{T-2} & \cdots & \sigma_1 & \sigma^2 \end{bmatrix}$ |
| First-order Antedependence (Zimmerman et al. 1995) | $\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_2\rho_1\rho_2 & \cdots & \sigma_1\sigma_{T-1}\prod_{i=1}^{T-1}\rho_i \\ & \sigma_2^2 & \sigma_2\sigma_3\rho_2 & \cdots & \sigma_1\sigma_{T-1}\prod_{i=2}^{T-1}\rho_i \\ & & \sigma_3^2 & \cdots & \sigma_1\sigma_{T-1}\prod_{i=3}^{T-1}\rho_i \\ & & & \ddots & \vdots \\ \text{Symmetric} & & & & \sigma_T^2 \end{bmatrix}$ |

Antedependence models are of particular interest. The typical AR(1) correlation structure assumes that the autocorrelation process is stationary and thus has uniform variance among observations with decaying correlation over time. Antedependence correlation structures are used when within-group autocorrelation is nonstationary, or when correlations between within-group observations do not uniformly decrease over

time (Zimmerman et al. 1995; Zimmerman and Nuñez-Antón 2009). The Toeplitz model typically results as the result of a moving average process and retains the stationarity assumptions of the typical ARMA(p,q) models (Zimmerman and Nunez-Anton 2009).

Corrections for heteroskedasticity were performed using the following variance function (Pinheiro and Bates 2000);

$$\text{Var}(\varepsilon_{ij}) = \sigma^2 \cdot \exp(2 \cdot \vartheta) \quad (6)$$

where ϑ represents the covariate for the variance function.

Models and Results

Three models were estimated: Total employment, to capture the overall economic impacts; number of manufacturing establishments, to capture the impacts on basic industry; and the number of full-service restaurants, to capture the impacts on the service industries. ZIP-level figures were normalized against the state in order to account for economic externalities, resulting in a “ZIP-to-State” ratio for each of these models. The ZIP-to-State ratio provides an at-a-glance evaluation of how much each ZIP Code contributes to the state’s overall economy. The “best” fitting correlation structures were chosen based on log-likelihood, AIC, and BIC values.

For the total employment model (Table 3), the age indicators for 1-5 years and 6-10 years are both negative and statistically significant; this trend was mirrored in the other two models. This could indicate that there are initial negative impacts for the first ten years a bypass has been open. However, only one community in the study area (Logansport) has a bypass that has been open for such a short amount of time, so these results should be taken with caution. The random effects estimates show that between-

ZIP variance significantly outweighs within-ZIP variance. This is to be expected, as this regional model includes ZIP Codes with low population, often clustered around ZIP Codes with high population (the ZIP Codes with cities). This high degree of variance is also captured through the use of the exponential variance function. Variance in the Toeplitz model appears to decrease over time (Table 4). The model with the Toeplitz autocorrelation structure yielded the highest log-likelihood values and thus is the best fitting model.

For the manufacturing model (Table 5), the coefficients for the Logansport, Huntington, and Marion indicators are negative (although the Marion coefficient is only significant in the antedependence model). The communities located along the US-24 corridor – Logansport, Peru, Wabash, and Huntington – experienced a decline in employment during the 1980s as a result of the closure of several manufacturing plants and the Grissom Air Force Base. Marion, located 25 miles south of these communities, has also experienced a steady decline in employment. This contrasts with the relatively good fortunes of communities located along the US-30 corridor – Warsaw, Plymouth, and Columbia City. Warsaw in particular continues to maintain a high level of employment due to its concentration of orthopedic manufacturing firms. As can be expected, establishment levels of other industrial and civic sectors has a positive relationship on the number of manufacturing establishments; the cause and effect, relationship, is unclear. As with the total employment model, the age indicators for 31-35 years and 36-40 years are both positive and significant. This could indicate that in the long run, bypasses will induce economic growth for the affected community. Both the ARH(1) and antedependence models (see Table 6) showed variance over time; curiously, autocorrelation for the first three years of analysis was not significant in the antedependence model.

Similar findings arise for the full-service restaurant establishments model (Table 7). Impacts are positive and statistically significant for communities with bypasses that

have been open for at least 30 years. The positive coefficient of the Columbia City indicator and the negative coefficient of the Peru indicator reflect the divergent employment trajectories along the US-30 and US-24 corridors, respectively. As observed from the log-likelihood values, the antedependence model has the best fit. While both the ARH(1) and antedependence models (Table 8) both show decreasing variance over time, the antedependence model reveals that the degree of autocorrelation (the rho parameters) significantly change in magnitude over time, a trend not captured by the ARH(1) model.

**TABLE 3 Fixed Effects Estimates for Total Employment
ZIP to State Ratio (Natural Log)**

| Variable | Coefficient | |
|---|---|--|
| | No Correction for Autocorrelation or Heteroskedasticity (1) | Toeplitz, with Correction for Heteroskedasticity (2) |
| Intercept | -6.678*** | -6.675*** |
| Median Real Estate Taxes / Aggregate Real Estate Taxes (2000) | -146.940*** | -146.300* |
| Distance from population center of ZIP Code to bypass | -0.056* | -0.05633*** |
| Rural Population (Farm) / Total Population (Squared) | -74.772* | -75.573*** |
| Bypass Age Indicator: 1 to 5 years | -0.038*** | -0.024*** |
| Bypass Age Indicator: 6 to 10 years | -0.1613*** | -0.054*** |
| Bypass Age Indicator: 31 to 35 years | 0.049* | 0.074*** |
| Bypass Age Indicator: 36 to 40 years | -0.022 | 0.057*** |
| AIC | 181.50 | -306.10 |
| BIC | 185.80 | -280.00 |
| Log-Likelihood | -88.75 | 165.05 |
| Log-Likelihood (Intercept Only) | -112.65 | -112.65 |

*** - Significant at 1 percent ** - Significant at 5 percent * - Significant at 10 percent

**TABLE 4 Random Effects Estimates for Total Employment
ZIP to State Ratio (Natural Log)**

| Variable | Coefficient | |
|------------------------------|---|--|
| | No Correction for Autocorrelation or Heteroskedasticity (1) | Toeplitz, with Correction for Heteroskedasticity (2) |
| Intercept | 1.7430*** | 1.5728*** |
| Toeplitz Sigma Squared | - | 0.2031*** |
| Toeplitz Sigma 2 | - | 0.1990*** |
| Toeplitz Sigma 3 | - | 0.1911*** |
| Toeplitz Sigma 4 | - | 0.1811*** |
| Toeplitz Sigma 5 | - | 0.1700*** |
| Toeplitz Sigma 6 | - | 0.1578*** |
| Toeplitz Sigma 7 | - | 0.1462*** |
| Toeplitz Sigma 8 | - | 0.1393*** |
| Toeplitz Sigma 9 | - | 0.1336*** |
| Variance Fn: Exp(Population) | - | -0.0004*** |
| Residual | 0.04146*** | 0.03287*** |

**TABLE 5 Fixed Effects Estimates for Number of Manufacturing Establishments
ZIP to State Ratio (Natural Log)**

| Variable | Coefficient | | |
|---|---|-------------------------|--------------------------------|
| | No Correction for Autocorrelation or Heteroskedasticity (1) | Heterogeneous AR(1) (2) | First-order Antedependence (3) |
| Intercept | -6.4731*** | -6.6027*** | -6.6704*** |
| Indicator: 1 if nearest large city is Marion, 0 otherwise | -0.2898 | -0.2686 | -0.3521* |
| Indicator: 1 if nearest bypassed city is Logansport, 0 otherwise | -0.7211*** | -0.7350*** | -0.7988*** |
| Indicator: 1 if nearest bypassed city is Huntington, 0 otherwise | -0.3751 | -0.366 | -0.4007 |
| Health services number of establishments: ZIP to State Ratio | 179.87** | 200.35*** | 205.07*** |
| Religious organizations, number of establishments: ZIP to State Ratio | 194.62*** | 186.31*** | 185.74*** |

| | | | |
|---|------------|------------|------------|
| Number of Construction Establishments, ZIP to State Ratio (Natural Log) | 0.1675** | 0.1483*** | 0.1301*** |
| Transportation and Public Utilities, Number of Establishments: ZIP to State Ratio | 82.1533** | 46.0499*** | 49.3324*** |
| Bypass Age Indicator: 1 to 5 years | -0.1271*** | 0.04437*** | 0.02648** |
| Bypass Age Indicator: 31 to 35 years | 0.1334* | 0.05504 | 0.06833* |
| Bypass Age Indicator: 36 to 40 years | 0.1305** | 0.09108* | 0.08627* |
| AIC | 251.34 | 9.66 | 11.76 |
| BIC | 255.69 | 33.58 | 50.90 |
| Log-Likelihood | -123.67 | 6.17 | 12.12 |
| Log-Likelihood (Intercept Only) | -160.17 | -160.17 | -160.17 |

*** - Significant at 1 percent ** - Significant at 5 percent * - Significant at 10 percent

**TABLE 6 Random Effects Estimates for Number of Manufacturing Establishments
ZIP to State Ratio (Natural Log)**

| Variable | Coefficient | | |
|-----------|---|-------------------------|--------------------------------|
| | No Correction for Autocorrelation or Heteroskedasticity (1) | Heterogeneous AR(1) (2) | First-order Antedependence (3) |
| Intercept | 0.5510*** | 0.6162*** | 0.6016*** |
| Sigma 1 | - | 0.1552*** | 0.06521*** |
| Sigma 2 | - | 0.1327*** | 0.01792 |
| Sigma 3 | - | 0.1197*** | 0.0334 |
| Sigma 4 | - | 0.1170*** | 0.06178*** |
| Sigma 5 | - | 0.07646*** | 0.1008*** |
| Sigma 6 | - | 0.08380*** | 0.1210*** |
| Sigma 7 | - | 0.04593*** | 0.1399*** |
| Sigma 8 | - | 0.04149*** | 0.1379*** |
| Sigma 9 | - | 0.09575*** | 0.1816*** |
| Rho 1 | - | 0.7498*** | 0.2109 |
| Rho 2 | - | - | -0.0955 |
| Rho 3 | - | - | 0.3783 |
| Rho 4 | - | - | 0.6912*** |
| Rho 5 | - | - | 0.7691*** |
| Rho 6 | - | - | 0.8861*** |
| Rho 7 | - | - | 0.9360*** |
| Rho 8 | - | - | 0.8652*** |
| Residual | 0.0564*** | - | - |

TABLE 7 Fixed Effects Estimates for Number of Full-Service Restaurant Establishments, ZIP to State Ratio (Natural Log)

| Variable | Coefficient | | |
|---|---|-------------------------|--------------------------------|
| | No Correction for Autocorrelation or Heteroskedasticity (1) | Heterogeneous AR(1) (2) | First-order Antedependence (3) |
| Intercept | -7.5433*** | -7.5388*** | -7.3702*** |
| Indicator: 1 if nearest bypassed city is Columbia City, 0 otherwise | 0.4105* | 0.362 | 0.4081* |
| Indicator: 1 if nearest bypassed city is Peru, 0 otherwise | -0.2661* | -0.3098** | -0.3189** |
| Indicator: 1 if nearest bypassed city is Rochester 0 otherwise | 0.1991 | - | - |
| Indicator: 1 if nearest large city is Marion, 0 otherwise | -0.2647** | -0.3161*** | -0.2864** |
| Rural Population (Farm and Nonfarm) / Total Population | -0.5097 | -0.4908 | -0.6681* |
| Indicator: 1 if route from population center of ZIP Code to bypass was along a two-lane county, State Route, or US highway, 0 otherwise | 0.1964 | 0.2049 | 0.2190* |
| Health services number of establishments: ZIP to State Ratio | 395.66*** | 353.90*** | 332.66*** |
| Religious organizations, number of establishments: ZIP to State Ratio | 66.8824 | 96.007 | 78.9224 |
| Liquor stores, number of establishments: ZIP to State Ratio | 55.7598* | 58.0764** | 52.2429* |
| Bypass Age Indicator: 1 to 5 years | -0.2937*** | -0.2136*** | -0.3513*** |
| Bypass Age Indicator: 6 to 10 years | -0.1303*** | -0.1962*** | -0.2709*** |
| Bypass Age Indicator: 31 to 35 years | 0.1443*** | 0.1134** | 0.1863*** |
| Bypass Age Indicator: 36 to 40 years | 0.1478*** | 0.1320*** | 0.1488** |
| AIC | 329.98 | 260.40 | 222.54 |
| BIC | 334.33 | 284.32 | 261.68 |
| Log-Likelihood | -162.99 | -119.2 | -93.27 |
| Log-Likelihood (Intercept Only) | -223.295 | -223.295 | -223.295 |

*** - Significant at 1 percent ** - Significant at 5 percent * - Significant at 10 percent

TABLE 8 Random Effects Estimates for Number of Full-Service Restaurant Establishments, ZIP to State Ratio (Natural Log)

| Variable | Coefficient | | |
|-----------|---|-------------------------|--------------------------------|
| | No Correction for Autocorrelation or Heteroskedasticity (1) | Heterogeneous AR(1) (2) | First-order Antedependence (3) |
| Intercept | 0.2170*** | 0.2069*** | 0.2148*** |
| Sigma 1 | - | 0.06817*** | 0.09744*** |
| Sigma 2 | - | 0.06782*** | 0.08404*** |
| Sigma 3 | - | 0.1198** | 0.06531*** |
| Sigma 4 | - | 0.1324*** | 0.07047*** |
| Sigma 5 | - | 0.0976*** | 0.06805*** |
| Sigma 6 | - | 0.0944*** | 0.09630*** |
| Sigma 7 | - | 0.0833*** | 0.07850*** |
| Sigma 8 | - | 0.0651*** | 0.06378*** |
| Sigma 9 | - | 0.1136*** | 0.1263*** |
| Rho 1 | - | 0.4931** | 0.8457*** |
| Rho 2 | - | - | 0.1131 |
| Rho 3 | - | - | 0.05123 |
| Rho 4 | - | - | -0.06767 |
| Rho 5 | - | - | 0.5444*** |
| Rho 6 | - | - | 0.5115*** |
| Rho 7 | - | - | 0.4231*** |
| Rho 8 | - | - | 0.5585*** |
| Residual | 0.07337*** | - | - |

Conclusions, Limitations, and Future Work

Each of these models have shown that bypasses have positive economic impacts on affected communities. All three models showed that while the initial economic impacts may be negative for the first decade a bypass has been open, these will be outweighed by positive impacts in later years. The use of the Toeplitz, antedependence, and heterogeneous AR(1) error structures were warranted, as shown by the significant degree of variance over time.

However, extrapolation of these model results should be taken with caution. The sample size is rather small, and the inclusion of only one community with a bypass that

had been open for less than 10 years does bias the sample to an extent. The models do not address the spatial relationship of the ZIP Codes in the study area. Spatial econometric models developed previously (Mills and Fricker 2010) showed that spatial autocorrelation was both present and significant in this study area. The total employment model in this study and the spatial lag model in the previous study do share one common characteristic, however; both indicate that while bypassed communities may benefit economically over time, surrounding, more rural areas may be adversely affected. This was demonstrated explicitly through the marginal effects in Mills and Fricker (2010) and implicitly here by the manner in which the bypass age indicators were defined. Because the age indicators are nonzero only for the ZIP Codes with the bypassed communities, the models imply that while the ZIP Codes with bypasses will experience growth in employment over time, the surrounding ZIP Codes may not be so fortunate.

The functional form of the multilevel models allowed for multiple types of variation, be it over space or over time, that cannot be captured in a model estimated by OLS. While spatial relationships are not explicitly captured, the models do indicate that over time, the communities in this study area did experience growth in total employment and in the basic (represented by the number of manufacturing establishments) and service (represented by full-service restaurant establishments) industries over time. Future studies should jointly consider the spatial relationships and the different means in which other unobserved variance can be captured.

Disclaimer

Funding for this research was provided by the NEXTRANS Center, Purdue University under Grant No. DTRT07-G-005 of the U.S. Department of Transportation, Research and Innovative Technology Administration (RITA), University Transportation Centers Program. The contents of this presentation reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This

document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

References

Croissant, Y., and G. Millo (2008). Panel Data Econometrics in R: The plm Package. *Journal of Statistical Software*, July 2008, Vol. 27, Issue 2.

Faraway, J.J. (2006). *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Chapman & Hall/CRC.

Laird, N. M., and J. H. Ware (1982). *Random-effects Models for Longitudinal Data*, *Biometrics* 38: 963–974.

Mills, J. (2007). *Effects of Bypasses*. Master's Thesis, Purdue University, West Lafayette, IN.

Mills, J., and J.D. Fricker (2010). *An Analysis of the Socioeconomic Impacts of Bypasses on Small- and Medium-Sized Communities Incorporating Spatial Econometric Methods*, Paper 10-0672, Transportation Research Board 89th Annual Meeting, January 2010, Washington DC.

Mills, J., and J.D. Fricker (2009). *The Effects on Bypasses on Communities: An Integrated Approach*, Proceedings of the 12th Conference on Transportation Planning Applications, May 2009, Houston TX.

Pinheiro, J.C., and D.M. Bates (2000). *Mixed-Effects Models in S and S-PLUS*. Springer.

SAS Institute Inc. (2008). *The MIXED Procedure, SAS/STAT 9.2® User's Guide*. Cary, NC.

Verbeke, G., and G. Molenberghs (2001). *Linear Mixed Models for Longitudinal Data*, Springer.

Zimmerman, D.L., Nuñez-Antón, V., and H. El-Barmi (1995). *Computational Aspects of Likelihood-Based Estimation of First-Order Antedependence Models*, *Journal of Statistical Computation and Simulation*, 60, 1, 67-84.

Zimmerman, D.L., and V. Nuñez-Antón (2009). *Antedependence Models for Longitudinal Data*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability.