A Count Data Model with Endogenous Covariates: Formulation and Application to Roadway Crash Frequency at Intersections

Chandra R. Bhat*
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712
Phone: 512-471-4535; Fax: 512-475-8744
Email: bhat@mail.utexas.edu
and
King Abdulaziz University, Jeddah 21589, Saudi Arabia

Kathryn Born
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712
Phone: 512-471-4535, Fax: 512-475-8744
E-mail: born2@utexas.edu

Raghuprasad Sidharthan
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712
Phone: 512-471-4535, Fax: 512-475-8744
E-mail: raghu@mail.utexas.edu

Prerna C. Bhat
Harvard University
1350 Massachusetts Avenue, Cambridge, MA 02138
Phone: 512-289-0221
E-mail: prernabhat@college.harvard.edu

*corresponding author

July 20, 2013

ABSTRACT

This paper proposes an estimation approach for count data models with endogenous covariates. The maximum approximate composite marginal likelihood inference approach is used to estimate model parameters. The modeling framework is applied to predict crash frequency at urban intersections in Irving, Texas. The sample is drawn from the Texas Department of Transportation (TxDOT) crash incident files for the year 2008. The results highlight the importance of accommodating endogeneity effects in count models. In addition, the results reveal the increased propensity for crashes at intersections with flashing lights, intersections with crest approaches, and intersections that are on frontage roads.

Keywords: Count data; treatment-outcome models; accident analysis; generalized ordered response; flashing light control.

1. INTRODUCTION

This paper develops an estimation approach for count data models with endogenous covariates, where the endogenous covariates are based on a multinomial probit model of discrete choice. The proposed formulation model constitutes a specific version of the generalized Roy model that is referred to as the treatment effects model (see Heckman and Vytlacil, 2005 and Bhat and Eluru, 2009). In the empirical context studied in the paper, the type of control at an intersection constitutes the treatment. The type of control is represented in five categories: (1) no traffic control (including intersections with no control and intersections with some minimal form of control such as turn marks and marked lanes), (2) yield sign control on one more approaches with no other form of control, (3) stop sign control on one or more approaches, and (4) flashing light control (one or more approaches having a flashing red or yellow light), and (5) regular signal light control. The count outcome in the empirical context is the number of crashes at urban intersections. In this case, the type of traffic control may itself be determined by the frequency of crashes, as, in fact, is explicitly noted in the Manual on Uniform Traffic Control Devices or MUTCD (FHWA, 2009). For instance, the total entering volume of traffic on the approach roadways to an intersection may directly impact both the type of control as well as the frequency of crashes, creating an "endogeneity" of the type of traffic control in crash frequency analysis. But if the entering volume were an observed variable, then this type of "endogeneity" is easily accommodated by including the entering volume as an explanatory variable, along with traffic control type, in the modeling of crash frequency. More generally, if the determination of the control types at intersections were random conditional on observed characteristics, a traditional count model for crash frequency would suffice. However, many unobserved factors may affect both control type and crash frequency, rendering the random conditional (on observed characteristics) assumption untenable. For instance, at intersections with an unobserved terrestrial/topographic feature that limits sight distance, flashing lights may be installed instead of a stop sign. That same unobserved feature may be responsible for a lower or a higher frequency of crashes (one can argue that motorists are more careful when they encounter some unobserved topographic feature, resulting in a lower frequency of crashes; alternatively, it could also be that the unobserved feature results in a higher frequency of crashes). If one of these two situations exists, but is ignored, it would generate an inconsistent, spurious, and biased effect of flashing lights on intersection crash frequency. Of course, there are many other application

contexts where our proposed model formulation should be useful, including the type of insurance plan ("treatment") and the number of doctor visits (outcome), residential location type ("treatment") and the number of cancer incidents (outcome), the intensity of lighting at a train station and the number of crimes at the station, and one of many other contexts where count models are used to model the outcome decision. However, in the development of the model formulation in this paper, as in the empirical analysis context of the paper, the focus will be on traffic control type and the frequency of crashes at urban intersections.

Methodologically speaking, our parametric multinomial discrete-count model uses a general multinomial probit (MNP) specification for the treatment and ties this MNP model with a count model. In particular, we use Castro, Paleti, and Bhat's or CPB's (2012) recasting of a univariate count model as a restricted version of a univariate generalized ordered-response probit (GORP) system. In addition to providing substantial flexibility to accommodate high or low probability masses for specific count outcomes, the latent variable-based count specification of the GORP system provides a convenient mechanism to tie the count outcome with the MNP treatment model. In this regard, our proposed model (which we will refer to as the Count model with endogenous multinomial probit selection, or the CEMPS model) has some similarity with Bhat (1998) and Munkin and Trivedi's (2008) ordered probit model with endogenous selection, but with four important differences. First, the outcome variables in the earlier models were ordinal variables, while the outcome variable in our CEMPS model is a true count variable that can take on any non-negative integer value. Second, the earlier models did not allow random response variations (or unobserved heterogeneity) in the sensitivity to exogenous factors in both the selection (or treatment) component as well as the outcome component. On the other hand, it is now well established that ignoring such response variations when present will lead to inconsistent and biased parameters estimates in both multinomial discrete choice models as well as count models (see Chamberlain, 1980, Bhat, 1998). For instance, variations in the effect of entering volume on the type of control installed at an intersection and on crash frequency may result from the complex interactions between unobserved intersection characteristics and motorist learning/adaptation behavior in response to different levels of traffic volume. Accommodating such unobserved heterogeneity effects is not simply an esoteric econometric effort, but can have very real implications for accurately assessing the overall effects of variables on the outcome of interest (for example, to design countermeasures to reduce crash frequency in

our empirical context; see Anastasopoulos and Mannering, 2009 and Castro *et al.*, 2013). Third, we allow unobserved heterogeneity in the treatment effects themselves rather than *a priori* positing fixed treatment effects. For example, even after controlling for endogeneity effects, the "true" effect of control type on crash frequency itself may vary across intersections due to such unobserved intersection geometric features as curb radii and approach configuration. Fourth, unlike Bhat (1998), we use an MNP-based treatment model rather than a very restrictive multinomial logit treatment model, and use a simple frequentist inference approach rather than Munkin and Trivedi's relatively cumbersome Bayesian estimation approach. The frequentist approach is based on an analytic (as opposed to a simulation) approximation of the multivariate normal cumulative distribution (MVNCD) function that appears in the full likelihood function of the proposed model. Bhat (2011) discusses this analytic approach, which is based on earlier works by Solow (1960) and Joe (1996). The approach involves only univariate and bivariate cumulative normal distribution function evaluations in the likelihood function.

In summary, and to our knowledge, this is the first formulation and application of a flexible count outcome model with a multinomial probit selection model, which also accommodates unobserved heterogeneity effects.

2. MODEL FORMULATION

2.1. The Selection (Treatment) MNP Model

In the usual random discrete choice formulation, write the unobserved continuous random latent variable influencing the probability that intersection q is controlled by traffic control type i as follows:

$$U_{qi} = \boldsymbol{\beta}_q' \mathbf{x}_{qi} + \xi_{qi} \tag{1}$$

where $\mathbf{x}_{qi}$ is a $(D \times 1)$-column vector of exogenous attributes (including a dummy variable for each control type alternative except a base control type), $\boldsymbol{\beta}_q$ is an individual-specific $(D \times 1)$-column vector of corresponding coefficients that varies across intersections based on unobserved intersection attributes, and ξ_{qi} captures the idiosyncratic (unobserved) intersection characteristics that impact the latent propensity of control type i being installed at intersection q (in the rest of this paper, we will refer to U_{qi} as the propensity of control type i being installed at

intersection q). We assume that the error terms ξ_{qi} are multivariate normally distributed across control types i for a given intersection q: $\boldsymbol{\xi}_q=(\xi_{q1},\xi_{q2},...,\xi_{qI})' \sim MVN_I(\mathbf{0}_I,\boldsymbol{\Lambda})$, where $MVN_I(\mathbf{0}_I,\boldsymbol{\Lambda})$ indicates an I-variate normal distribution with a mean vector of zeros denoted by $\mathbf{0}_I$ and a covariance matrix $\boldsymbol{\Lambda}$. Such a specification captures the possibility that, for instance, topographical and unobserved features that hinder the approach line of sight to an intersection (and therefore increase or decrease the occurrence of crashes) may also impact the propensity of flashing lights or full signal lights being installed at the intersection. It also allows for unobserved features to impact different control type propensities differently. Of course, the precise reasons for covariances and heteroscedasticity in the underlying latent propensities across control types are, by definition, not observed, but it is not difficult to conceive of reasons why such effects may exist. At the least, it behooves the analyst to consider a general covariance matrix (but see identification issues discussed later) rather than *a priori* assuming an independent and identically distributed covariance matrix for $\boldsymbol{\xi}_q$. Further, to allow variation in the effect of observed intersection characteristics due to unobserved intersection attributes (as discussed in the previous section), we consider $\boldsymbol{\beta}_q$ as a realization from a multivariate normal distribution: $\boldsymbol{\beta}_q \sim MVN_D(\mathbf{b},\widetilde{\boldsymbol{\Omega}})$. The vectors $\boldsymbol{\beta}_q$ and $\boldsymbol{\xi}_q$ are assumed to be independent of each other. For future reference, we also write $\boldsymbol{\beta}_q = \mathbf{b}+\widetilde{\boldsymbol{\beta}}_q$, where $\widetilde{\boldsymbol{\beta}}_q \sim MVN_D\left(\mathbf{0}_D,\widetilde{\boldsymbol{\Omega}}\right)$.

The model above may be written in a more compact form by defining the following vectors and matrices: $\mathbf{U}_q=(U_{q1},U_{q2},...,U_{qI})'$ ($I\times 1$ vector), $\mathbf{x_q}=(\mathbf{x_{q1}},\mathbf{x_{q2}},\mathbf{x_{q3}},...,\mathbf{x_{qI}})'$ ($I\times D$ matrix), $\mathbf{V}_q=\mathbf{x_q b}$ ($I\times 1$ vector), $\boldsymbol{\xi_q}=(\xi_{q1},\xi_{q2},...,\xi_{qI})'$ ($I\times 1$ vector), $\breve{\boldsymbol{\Omega}}_q=\mathbf{x_q}\widetilde{\boldsymbol{\Omega}}\mathbf{x}_q'$ ($I\times I$ matrix), and $\boldsymbol{\Omega}_q=\breve{\boldsymbol{\Omega}}_q+\boldsymbol{\Lambda}$ ($I\times I$ matrix). Also, for later use, partition $\breve{\boldsymbol{\Omega}}_q$ so that the first alternative's utility covariance component is separated from those of the remaining utility covariances attributable to random coefficients: $\breve{\boldsymbol{\Omega}}_q=\begin{bmatrix}\breve{\boldsymbol{\Omega}}_{q1} & \breve{\boldsymbol{\Omega}}'_{q1,>1}\\ \breve{\boldsymbol{\Omega}}_{q1,>1} & \breve{\boldsymbol{\Omega}}_{q,>1}\end{bmatrix}$. Then, we may write, in matrix notation, $\mathbf{U_q}=\mathbf{V_q}+\boldsymbol{\xi_q}$ and $\mathbf{U_q}\sim MVN_I(\mathbf{V_q},\boldsymbol{\Omega}_q)$. Also, let $\mathbf{u}_q=(u_{q1},u_{q2},\ldots u_{qI})'$ ($i\neq m_q$) be an $(I\text{-}1)\times 1$ vector, where m_q is the actual observed control type at intersection q, and

$u_{qi} = U_{qi} - U_{qm_q}$ $(i \neq m_q)$. Then, $\mathbf{u}_q < \boldsymbol{0}_{I-1}$, because control type m_q is the one installed at intersection q.

In the context of the formulation above, several important identification issues need to be addressed. First, a constant cannot be identified in the utilities for one of the I alternatives. Similarly, intersection-specific variables that do not vary across alternatives can be introduced for I–1 control type alternatives, with the remaining alternative being the base (but see also the fifth identification consideration discussed below; in the rest of this paper, we will use the first alternative, corresponding to the "no control" type alternative, as the base alternative). Second, the coefficients associated with the constants specific to each alternative have to be fixed parameters, because their randomness is already captured in the covariance matrix $\boldsymbol{\Lambda}$. Third, only the covariance matrix of the error differences is estimable. Taking the difference with respect to the first alternative, only the elements of the covariance matrix $\boldsymbol{\Lambda}_1$ of $\varepsilon_{qi1} = \xi_{qi} - \xi_{q1}$, $i \neq 1$ are estimable. However, the condition that $\mathbf{u}_q^* < \boldsymbol{0}_{I-1}$ takes the difference against the observed control type m_q at intersection q. Thus, during estimation, the covariance matrix $\boldsymbol{\Lambda}_{m_q}$ $\varepsilon_{qim_q} = \xi_{qi} - \xi_{qm_q}$, $i \neq m_q$ is desired. Since m_q will vary across intersections q, $\boldsymbol{\Lambda}_{m_q}$ will also vary across intersections. But all the $\boldsymbol{\Lambda}_{m_q}$ matrices must originate in the same covariance matrix $\boldsymbol{\Lambda}$ for the original error term vector $\boldsymbol{\xi}_q$. To achieve this consistency, $\boldsymbol{\Lambda}$ is constructed from $\boldsymbol{\Lambda}_1$ by adding an additional row on top and an additional column to the left. All elements of this additional row and column are filled with values of zeros. $\boldsymbol{\Lambda}_{m_q}$ may then be obtained appropriately for each intersection q based on the same $\boldsymbol{\Lambda}$ matrix. Fourth, an additional scale normalization needs to be imposed on $\boldsymbol{\Lambda}_1$. For this, we normalize the first diagonal element of $\boldsymbol{\Lambda}_1$ to the value of one. Fifth, in MNP models where the variables are all specific to the observational units (intersections in the current paper) and whose values do not vary across alternatives, empirical identification issues need to be considered. In particular, as discussed by Keane (1992) and Munkin and Trivedi (2008), identification is tenuous unless exclusion restrictions are placed in the form of at least one intersection characteristic being excluded from each control type propensity in addition to being excluded from a base alternative (but appearing in some other control type propensities). In our application, such exclusion restrictions were

identified based on the estimation of a simpler independent MNP model ($\boldsymbol{\Lambda}$ fixed to the identity matrix) and removing intersection variables that turned out to be statistically insignificant in impacting specific control type propensities.

With the normalizations above on the $\boldsymbol{\Lambda}$ matrix, the covariance matrix $\boldsymbol{\Omega}_q$ takes the form below:

$$\boldsymbol{\Omega}_q = \breve{\boldsymbol{\Omega}}_q + \boldsymbol{\Lambda} = \begin{bmatrix} \breve{\Omega}_{q1} & \breve{\boldsymbol{\Omega}}'_{q1,>1} \\ \breve{\boldsymbol{\Omega}}_{q1,>1} & \breve{\boldsymbol{\Omega}}_{q,>1} \end{bmatrix} + \begin{bmatrix} 0 & \mathbf{0}'_{I-1} \\ \mathbf{0}_{I-1} & \boldsymbol{\Lambda}_1 \end{bmatrix}. \tag{2}$$

2.2. The Count Outcome Model

Consider the recasting of the count model using a specific functional form for the generalized ordered-response probit (GORP) structure as follows:

$$y_q^* = \boldsymbol{\delta}_q' \mathbf{w}_q + \boldsymbol{\rho}_q \mathbf{A}_q + \eta_q,\ y_q = l_q \ \ if\ \ \psi_{q,l_q-1} < y_q^* < \psi_{q,l_q},\ \ l_q \in \{0,1,2,.....\}, \tag{3}$$

$$\psi_{q,l_q} = \Phi^{-1}\left[\frac{(1-c_q)^{\theta}}{\Gamma(\theta)} \sum_{r=0}^{l}\left(\frac{\Gamma(\theta+r)}{r!} c_q^r\right)\right] + \varphi_{l_q},\ c_q = \frac{\lambda_q}{\lambda_q+\theta},\ \text{and}\ \lambda_q = e^{\boldsymbol{\gamma}'\mathbf{z}_q}.$$

In the above equation, y_q^* is an underlying latent continuous crash propensity variable corresponding to intersection q that maps into the observed count l_q through the $\boldsymbol{\psi}_q$ vector (which is a vertically stacked column vector of thresholds $(\psi_{q,-1}, \psi_{q0}, \psi_{q1}, \psi_{q2}, ..., \infty)'$. $\boldsymbol{\delta}_q$ is an intersection-specific $(C \times 1)$-column vector of coefficients on a conformable $(C \times 1)$-column vector of observable covariate vector $\mathbf{w}_q$ (not including a constant), $\mathbf{A}_q$ is an *(I*-1)×1-column vector of binary (0/1) indicator variables for the absence/presence of each control type (except the base "no control" type) at intersection q $[\mathbf{A}_q = (a_{q2}, a_{q3}, ...a_{qI})']$. Thus, for intersection q with observed control type m_q $(m_q \neq 1)$, $a_{m_q} = 1$ and all other elements are zero; if $m_q = 1$, all elements of $\mathbf{A}_q$ take a value of zero. $\boldsymbol{\rho}_q$ is a corresponding intersection-specific vector of structural control type "treatment" effects (relative to the base category of "no control") on the frequency of crashes at intersection q. η_q is a random error term representing unobserved factors influencing the latent variable y_q^*, and therefore the observed counts. It is assumed to be

standard normal distributed.[1] $\boldsymbol{\gamma}$ is another ($\widetilde{C} \times 1$)-column vector of parameters corresponding to another vector of observable covariates $\mathbf{z}_q$ (including a constant). Φ^{-1} in the threshold function of Equation (3) is the inverse function of the univariate cumulative standard normal. θ is a parameter that provides flexibility to the count formulation, and, as we will see later, serves the same purpose as the dispersion parameter in a traditional negative binomial model (θ >0). $\Gamma(\theta)$ is the traditional gamma function; $\Gamma(\theta) = \int_0^\infty t^{\theta-1} e^{-t}\, dt$. The threshold terms in the $\boldsymbol{\psi}_q$ vector satisfy the ordering condition (*i.e.*, $\psi_{q,-1} < \psi_{q,0} < \psi_{q,1} < \psi_{q,2} .... < \infty \; \forall \; q$) as long as $\varphi_{-1} < \varphi_0 < \varphi_1 < \varphi_2 .... < \infty$. The presence of these φ terms provides substantial flexibility to accommodate high or low probability masses for specific count outcomes, beyond what can be offered by traditional treatments using zero-inflated or related mechanisms. For identification, we set $\varphi_{-1} = -\infty, \psi_{q,-1} = -\infty \; \forall q,$ and $\varphi_0 = 0$. In addition, we identify a count value e^* ($e^* \in \{0,1,2,.....\}$) above which φ_e ($e \in \{0,1,2,.....\}$) is held fixed at φ_{e^*}; that is, $\varphi_e = \varphi_{e^*}$ if $e > e^*$, where the value of e^* can be based on empirical testing. For later use, let $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \ldots \varphi_{e^*})'$ ($e^* \times 1$ vector).

To proceed, define $\mathbf{s}_q = (\mathbf{w}_q', \mathbf{A}_q')'$, and $\boldsymbol{\mu}_q = (\boldsymbol{\delta}_q', \boldsymbol{\rho}_q')'$. Assume that $\boldsymbol{\mu}_q$ is a realization from a multivariate normal distribution with mean vector **d** and covariance $\widetilde{\boldsymbol{\Gamma}}$. It is not necessary that all elements of $\boldsymbol{\mu}_q$ be random; that is, the analyst may specify fixed coefficients on some exogenous variables in the model, though it will be convenient in presentation to assume that all elements of $\boldsymbol{\mu}_q$ are random. Also, note that the submatrix of $\widetilde{\boldsymbol{\Gamma}}$ corresponding to the coefficient vector of $\boldsymbol{\rho}_q$ should be diagonal, because each intersection is controlled by a single control type. With the definitions above, Equation (3) may be equivalently written as:

$$y_q^* = \boldsymbol{\mu}'_q \mathbf{s}_q + \eta_q, \; y_q = l_q \; \textit{if} \; \psi_{q,l_q-1} < y_q^* < \psi_{q,l_q}, \; l_q \in \{0,1,2,.....\}, \tag{4}$$

[1] The exclusion of a constant in the vector $\mathbf{w}_q$ of Equation (3) and the use of the standard normal distribution (as opposed to a non-standard normal distribution) for the error term η_q are innocuous normalizations (see Zavoina and McKelvey, 1975; Greene and Hensher, 2010).

$$\psi_{q,l_q} = \Phi^{-1}\left[\frac{(1-c_q)^{\theta}}{\Gamma(\theta)}\sum_{r=0}^{l}\left(\frac{\Gamma(\theta+r)}{r!}c_q^r\right)\right] + \varphi_{l_q},\ c_q = \frac{\lambda_q}{\lambda_q+\theta},\ \text{and}\ \lambda_q = e^{\boldsymbol{\gamma}'\mathbf{z}_q}.$$

The specification of the GORP model in the equation above provides a very flexible mechanism to model count data. It subsumes the traditional count models as very specific and restrictive cases. In particular, if the vector $\boldsymbol{\mu}_q$ is degenerate with all its elements taking the fixed value of zero, and all elements of the $\boldsymbol{\varphi}$ vector are zero, the model in Equation (4) collapses to a traditional negative binomial model with dispersion parameter θ. If, in addition, $\theta \to \infty$, the result can be shown to be the Poisson count model. Also, note that the non-linear functional form for the effects of the variables in the $\mathbf{z}_q$ vector on the thresholds allows identification for any variables that are common in the $\mathbf{z}_q$ and $\mathbf{s}_q$ vectors. Further, we can write $y_q^* \sim N(D_q, \ddot{\Gamma})$, where $D_q = \mathbf{d}'\mathbf{s}_q$ and $\ddot{\Gamma} = 1 + \mathbf{s}_q\tilde{\boldsymbol{\Gamma}}\mathbf{s}_q' = 1 + \breve{\Gamma}$, where $\breve{\Gamma} = \mathbf{s}_q\tilde{\boldsymbol{\Gamma}}\mathbf{s}_q'$.

In the empirical context of crash counts at intersections, CPB interpret the GORP recasting of the count model as follows. There is a latent "long-term" crash propensity y_q^* associated with intersection q that is a linear function of a set of intersection-related attributes $\mathbf{w}_q$ and the endogenous intersection control type variables $\mathbf{A}_q$. On the other hand, there may be some specific intersection characteristics (embedded in $\mathbf{z}_q$ within the threshold terms) that may dictate the likelihood of a crash occurring at any given *instant of time* for a given long-term crash propensity y_q^* (there may be common elements in $\mathbf{w}_q$ and $\mathbf{z}_q$). Thus, two intersections may have the same latent long-term crash propensity y_q^*, but may show quite different observed number of crashes over a certain time period because of different y_q^*- to - y_q mappings through the cut points (y_q is the observed count variable).

2.3. The Joint Model System

The count model of the previous section can be estimated independently of the MNP-based control type model of Section 2.1. However, doing so would assume that the observed control type at intersections has nothing to do with the frequency of crashes, thus resulting in the inference that the estimates related to the elements of the $\boldsymbol{\rho}_q$ vector provide the "treatment"

effect if specific control types are put in place at an intersection. However, as discussed earlier in Section 1, there is reason to believe that the unobserved factors contained in the elements of the error vector $\boldsymbol{\xi}_q$ may also be manifested in the error term η_q. If this is the case, a non-random assignment of control types to intersections is being used to make inferences about the potential engineering effect of using different control strategies. But, assessing this engineering effect is effectively equivalent to a thought experiment in which differences in crash frequency are to be evaluated following the random assignment of control types to intersections. Econometrically speaking, the challenge then is to attempt to compute this engineering effect from a non-randomly assigned observed sample by accommodating the covariance effects between the vector $\boldsymbol{\xi}_q$ and the error term η_q.

Of course, as already discussed in Section 2.1, only differences in the control type propensities matter in the MNP model, and so, without loss of generality, we accommodate the covariance between $\boldsymbol{\xi}_q$ and η_q by specifying their joint covariance matrix as follows

$$\widetilde{\boldsymbol{\Sigma}} = \begin{bmatrix} 0 & \mathbf{0}'_{I-1} & 0 \\ \mathbf{0}_{I-1} & \boldsymbol{\Lambda}_{\mathbf{1}} & \boldsymbol{\Xi} \\ 0 & \boldsymbol{\Xi}' & 1 \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & \widetilde{\boldsymbol{\Sigma}}_{\mathbf{1}} \end{bmatrix}, \text{ where } \widetilde{\boldsymbol{\Sigma}}_{\mathbf{1}} = \begin{bmatrix} \boldsymbol{\Lambda}_{\mathbf{1}} & \boldsymbol{\Xi} \\ \boldsymbol{\Xi}' & 1 \end{bmatrix}, \tag{5}$$

where $\boldsymbol{\Xi}$ is the $(I-1)\times 1)$ column vector of the covariance elements of the error differences $\varepsilon_{qi1} = \xi_{qi} - \xi_{q1}$, $i \neq 1$ in the discrete choice model with the η_q error term in the count model. Now, consider the $[(I+1)\times 1)]$ vector $\mathbf{G} = \left[\mathbf{U}'_q, y_q^*\right]'$. Let $\mathbf{H}_q = (\mathbf{V}'_q, D_q)'$. Then, $\mathbf{G} \sim MVN_{(I+1)}(\mathbf{H}_q, \boldsymbol{\Sigma}_{\mathbf{q}})$, where

$$\begin{aligned} \boldsymbol{\Sigma}_{\mathbf{q}} &= \begin{bmatrix} \breve{\Omega}_{q1} & \breve{\boldsymbol{\Omega}}'_{q1,>1} & 0 \\ \breve{\boldsymbol{\Omega}}_{q1,>1} & \breve{\boldsymbol{\Omega}}_{q,>1} & \mathbf{0}_{I-1} \\ 0 & \mathbf{0}'_{I-1} & 0 \end{bmatrix} + \begin{bmatrix} 0 & \mathbf{0}'_{I-1} & 0 \\ \mathbf{0}_{I-1} & \boldsymbol{\Lambda}_{\mathbf{1}} & \boldsymbol{\Xi} \\ 0 & \boldsymbol{\Xi}' & 1 \end{bmatrix} + \begin{bmatrix} 0 & \mathbf{0}'_{I-1} & 0 \\ \mathbf{0}_{I-1} & \mathbf{0}_{I-1,I-1} & \mathbf{0}_{I-1} \\ 0 & \mathbf{0}'_{I-1} & \breve{\Gamma} \end{bmatrix} \\ &= \begin{bmatrix} \breve{\Omega}_{q1} & \breve{\boldsymbol{\Omega}}'_{q1,>1} & 0 \\ \breve{\boldsymbol{\Omega}}_{q1,>1} & \breve{\boldsymbol{\Omega}}_{q,>1} + \boldsymbol{\Lambda}_{\mathbf{1}} & \boldsymbol{\Xi} \\ 0 & \boldsymbol{\Xi}' & 1+\breve{\Gamma} \end{bmatrix}. \end{aligned} \tag{6}$$

The positive definiteness of the covariance matrix $\boldsymbol{\Sigma}_q$ can be ensured by making certain that the covariance matrices $\widetilde{\boldsymbol{\Omega}}$ (covariance matrix of $\widetilde{\boldsymbol{\beta}}_q$), the diagonal covariance matrix $\widetilde{\boldsymbol{\Gamma}}$

(covariance matrix of $\boldsymbol{\mu}_q$), and the covariance matrix $\widetilde{\boldsymbol{\Sigma}}_1$ (covariance matrix of the error differences $\varepsilon_{qi1} = \xi_{qi} - \xi_{q1}$, $i \neq 1$ in the discrete choice model and the η_q error term in the count model) are each positive definite. This is considered by using a Cholesky-decomposition of these three matrices, and undertaking the estimation with respect to these Cholesky-decomposed parameters. Note also that the first diagonal element of $\widetilde{\boldsymbol{\Sigma}}_1$ is normalized to one for identification, which can be ensured by fixing the first diagonal element of the Cholesky matrix to one. Further, the last diagonal element of $\widetilde{\boldsymbol{\Sigma}}_1$ also is normalized to one (this is the variance of the error term η_q). To maintain this restriction, the corresponding diagonal element of the Cholesky matrix is written as $\sqrt{1-\sum_{j=1}^{I-1} h_{Ij}^2}$, where the h_{Ij} elements are the Cholesky factors in the last row *I*.

To develop the likelihood function, define $\mathbf{M}_q$ as an identity matrix of size *I* with an extra column added at the m_q^{th} column (thus, $\mathbf{M}_q$ is a matrix of dimension $(I)\times(I+1)$). This m_q^{th} column of $\mathbf{M}_q$ has the value of '-1' in the first $(I\text{-}1)$ rows and the value of zero in the final row. Then, the vector $\widetilde{\mathbf{G}}_q = (\mathbf{u}_q' \; y_q^*)'$ (of size $I\times 1$) is distributed as follows: $\widetilde{\mathbf{G}}_q \sim MVN_I(\widetilde{\mathbf{H}}_q, \widetilde{\boldsymbol{\Sigma}}_q)$, where $\widetilde{\mathbf{H}}_q = \mathbf{M}_q\mathbf{H}_q$ and $\widetilde{\boldsymbol{\Sigma}}_q = \mathbf{M}_q\boldsymbol{\Sigma}_q\mathbf{M}_q'$. The likelihood that intersection *q* has control type m_q installed and has a crash frequency of l_q is equivalent to the joint probability that $\mathbf{u}_q < \boldsymbol{\theta}_{I-1}$ and $\psi_{q,l_q-1} < y_q^* < \psi_{q,l_q}$. Defining $\varpi = (\mathbf{b}', \overline{\widetilde{\boldsymbol{\Omega}}}', \overline{\boldsymbol{\Lambda}}', \mathbf{d}', \overline{\widetilde{\boldsymbol{\Gamma}}}', \boldsymbol{\gamma}', \boldsymbol{\varphi}', \theta)'$, where $\overline{\boldsymbol{\Lambda}}$ represents the vector of upper triangle elements of $\boldsymbol{\Lambda}$ (and similarly for other covariance matrices), $\boldsymbol{\omega}_\Delta$ for the diagonal matrix of standard deviations of matrix Δ, $F_E(.;\boldsymbol{\alpha},\boldsymbol{\Delta})$ for the multivariate normal cumulative distribution function of dimension *E* with mean vector $\boldsymbol{\alpha}$ and covariance matrix Δ, and $\Phi_E(.;\boldsymbol{\Delta}^*)$ for the multivariate standard normal cumulative distribution function of dimension *E* and correlation matrix $\boldsymbol{\Delta}^*$ such that $\boldsymbol{\Delta}^* = \boldsymbol{\omega}_\Delta^{-1}\boldsymbol{\Delta}\boldsymbol{\omega}_\Delta^{-1}$, the likelihood function contribution from intersection *q* may be written as:

$$
\begin{aligned}
L_q(\varpi) &= F_I\left[\left(\mathbf{0}'_{I-1}, \psi_{q,l_q}\right)'; \widetilde{\mathbf{H}}_q, \widetilde{\boldsymbol{\Sigma}}_q\right] - F_I\left[\left(\mathbf{0}'_{I-1}, \psi_{q,l_q-1}\right)'; \widetilde{\mathbf{H}}_q, \widetilde{\boldsymbol{\Sigma}}_q\right] \\
&= \Phi_I\left[\boldsymbol{\omega}_{\widetilde{\boldsymbol{\Sigma}}_q}^{-1}\left\{\left(\mathbf{0}'_{I-1}, \psi_{q,l_q}\right)' - \widetilde{\mathbf{H}}_q\right\}, \widetilde{\boldsymbol{\Sigma}}_q^*\right] - \Phi_I\left[\boldsymbol{\omega}_{\widetilde{\boldsymbol{\Sigma}}_q}^{-1}\left\{\left(\mathbf{0}'_{I-1}, \psi_{q,l_q}\right)' - \widetilde{\mathbf{H}}_q\right\}, \widetilde{\boldsymbol{\Sigma}}_q^*\right].
\end{aligned} \tag{7}
$$

The likelihood function of the observed sample is then developed as $L = \prod_q L_q(\varpi)$. The expression in Equation (7) may be computed using simulation-based methods or an analytic approximation approach to approximate the multivariate normal cumulative distribution (MVNCD) functions. Typical simulation-based methods can get inaccurate and time-consuming as the dimensionality increases. On the other hand, the analytic approximation approach of Joe (1995) and Bhat (2011) is based solely on univariate and bivariate cumulative normal distribution evaluations, regardless of the dimensionality of integration, which considerably reduces computation time compared to other simulation techniques to evaluate multidimensional integrals. This is the approach used in the current paper.

3. SIMULATION STUDY

In this section, we evaluate the ability of the analytic approximation to recover the parameters for the joint MNP-count model proposed in this paper, as well as assess the ability of the asymptotic standard errors from the analytic procedure to provide an estimate of the finite sample error for the typical size of samples employed in estimation.

3.1. Experimental Design

We consider three signal control type alternatives in the simulation experiments. Assume two independent variables for each of the three alternatives in the MNP model. The coefficient vector $\boldsymbol{\beta}_q$ for intersection q is assumed to be a realization from a multivariate normal distribution with a mean vector $\mathbf{b} = (1.5, -1)$ and covariance matrix $\widetilde{\boldsymbol{\Omega}}$ as follows:

$$
\widetilde{\boldsymbol{\Omega}} = \begin{bmatrix} 1.00 & 0.60 \\ 0.60 & 1.57 \end{bmatrix} = \mathbf{L}_{\widetilde{\boldsymbol{\Omega}}}\mathbf{L}'_{\widetilde{\boldsymbol{\Omega}}} = \begin{bmatrix} 1.00 & 0.00 \\ 0.60 & 1.10 \end{bmatrix}\begin{bmatrix} 1.00 & 0.60 \\ 0.00 & 1.10 \end{bmatrix} \tag{8}
$$

Then, there are three Cholesky matrix elements to be estimated in $\mathbf{L}_{\widetilde{\Omega}}$ ($l_{\widetilde{\Omega}1} = 1.00$, $l_{\widetilde{\Omega}2} = 0.60$, $l_{\widetilde{\Omega}3} = 1.10$). Collectively, these elements, vertically stacked into a column vector, will be referred to as $\boldsymbol{l}_{\widetilde{\Omega}}$.

For the count variable, we consider an exogenous variable in the $\mathbf{w}_q$ vector generated again from a standard univariate distribution. In addition, dummy variables corresponding to the choice of the second and third alternatives from the three signal type alternatives are included as structural effects in the count specification through the $\mathbf{A}_q$ vector. The coefficient vector $\boldsymbol{\mu}_q$ for intersection q is assumed to be a realization from a multivariate normal distribution with a mean vector $\mathbf{d} = (1, -1, -1.5)$ and a diagonal covariance matrix as follows:

$$\widetilde{\boldsymbol{\Gamma}} = \begin{bmatrix} 0.25 & 0.00 & 0.00 \\ 0.00 & 0.50 & 0.00 \\ 0.00 & 0.00 & 1.00 \end{bmatrix} = \mathbf{L}_{\widetilde{\Gamma}} \mathbf{L}'_{\widetilde{\Gamma}} = \begin{bmatrix} 0.500 & 0.000 & 0.000 \\ 0.000 & 0.707 & 0.000 \\ 0.000 & 0.000 & 1.000 \end{bmatrix} \begin{bmatrix} 0.500 & 0.000 & 0.000 \\ 0.000 & 0.707 & 0.000 \\ 0.000 & 0.000 & 1.000 \end{bmatrix} \tag{9}$$

There are three Cholesky matrix elements to be estimated in $\mathbf{L}_{\widetilde{\Gamma}}$ ($l_{\widetilde{\Gamma}1} = 0.5$, $l_{\widetilde{\Gamma}2} = 0.707$, $l_{\widetilde{\Gamma}3} = 1.0$). Collectively, these elements, vertically stacked into a column vector, will be referred to as $\boldsymbol{l}_{\widetilde{\Gamma}}$. The dispersion parameter θ is fixed at 2, and the $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \ldots \varphi_{e^*})'$ vector (labeled $\boldsymbol{\varphi}$ here) is set so that $\boldsymbol{\varphi} = (\varphi_1) = (0.75)$. We consider an exogenous variable in the $\mathbf{z}_q$ vector (embedded in the threshold function) generated again from a standard univariate distribution. The corresponding coefficient (labeled as $\boldsymbol{\gamma}_1$) is set to 0.5.

The covariance matrix that generates the jointness among the dependent variables (that is, the covariance matrix of the error differences $\varepsilon_{qi1} = \xi_{qi} - \xi_{q1}$, $i \neq 1$ in the discrete choice model and the η_q error term in the count model) is specified as follows:

$$\widetilde{\boldsymbol{\Sigma}}_1 = \begin{bmatrix} \boldsymbol{\Lambda}_1 & \boldsymbol{\Xi} \\ \boldsymbol{\Xi}' & 1 \end{bmatrix} = \begin{bmatrix} 1.00 & 0.60 & 0.00 \\ 0.60 & 1.00 & 0.48 \\ 0.00 & 0.48 & 1.00 \end{bmatrix} = \mathbf{L}_{\widetilde{\Sigma}_1} \mathbf{L}'_{\widetilde{\Sigma}_1} = \begin{bmatrix} 1.0 & 0.0 & 0.0 \\ 0.6 & 0.8 & 0.0 \\ 0.0 & 0.6 & 0.8 \end{bmatrix} \begin{bmatrix} 1.0 & 0.6 & 0.0 \\ 0.0 & 0.8 & 0.6 \\ 0.0 & 0.0 & 0.8 \end{bmatrix}, \tag{10}$$

In the above $\widetilde{\boldsymbol{\Sigma}}_1$ matrix, the first element is normalized (and fixed) to the value of 1, as is the third diagonal element. The sub-matrix of the first two columns and the first two rows of $\widetilde{\boldsymbol{\Sigma}}_1$

corresponds to the matrix $\mathbf{\Lambda_1}$, which is the covariance matrix of the utility differentials of the second and third alternatives (with respect to the first alternative) in the traffic control type variable. In the simulation exercise, for convenience, we fix the covariance of ε_{q21} and η_q to zero (as reflected by the zero entry in the first row and third column of $\widetilde{\mathbf{\Sigma}}_\mathbf{1}$. There are three Cholesky matrix elements to be estimated in $\mathbf{L}_{\widetilde{\Sigma}_1}$ ($l_{\widetilde{\Sigma}_1 1} = 0.6$, $l_{\widetilde{\Sigma}_1 2} = 0.8$, $l_{\widetilde{\Sigma}_1 3} = 0.6$). Collectively, these elements, vertically stacked into a column vector, will be referred to as $\boldsymbol{l}_{\widetilde{\Sigma}_1}$.

The set-up above is used to develop the covariance matrix $\widetilde{\mathbf{\Sigma}}$ for the error vector $\boldsymbol{\tau}_q = (\xi_{q1}, \xi_{q2}, \xi_{q3}, \eta_q)'$. The mean vector $\mathbf{V}_q = \left(V_{q1}, V_{q2}, V_{q3}\right)'$ for the latent variable vector $\mathbf{U}_q = \left(U_{q1}, U_{q2}, U_{q3}\right)'$ is also computed. Then, for each of the 2000 observations, a specific realization of the $\boldsymbol{\tau}_q$ vector is drawn from the multivariate normal distribution with mean $\mathbf{0}_4$ and covariance structure $\widetilde{\mathbf{\Sigma}}$. The realization corresponding to $\boldsymbol{\xi}_q = (\xi_{q1}, \xi_{q2}, \xi_{q3})'$ is added to the mean vector $\mathbf{V}_q$ to obtain the realization of the vector $\mathbf{U}_q$ for each observation. The alternative with the highest value from the vector $\mathbf{U}_q$ is then picked, and designated as the chosen alternative for each observation. Next, the value for $y_q^* = \boldsymbol{\mu}_q' \mathbf{s}_q + \eta_q$ is generated, and is translated into an observed count based on the computed threshold values

The above data generation process is undertaken 50 times with different realizations of the $\boldsymbol{\tau}_q$ vector to generate 50 different data sets, each with 2000 observations. The estimator is applied to each data set to estimate data specific values of $\left(\mathbf{b}', \boldsymbol{l}_{\widetilde{\Omega}}', \mathbf{d}', \boldsymbol{l}_{\widetilde{\Gamma}}', \theta, \varphi_1, \gamma_1, \boldsymbol{l}_{\widetilde{\Sigma}_1}'\right)'$. A single random permutation is generated for each individual (the random permutation varies across individuals, but is the same across iterations for a given individual) to decompose the multivariate normal cumulative distribution (MVNCD) function into a product sequence of marginal and conditional probabilities (see Section 2.1 of Bhat, 2011).[2] The estimator is applied to each dataset 10 times with different permutations to obtain the approximation error.

[2] Technically, the MVNCD approximation should improve with a higher number of permutations in the MACML approach. However, when we investigated the effect of different numbers of random permutations per individual, we noticed little difference in the estimation results between using a single permutation and higher numbers of permutations, and hence we settled with a single permutation per individual.

3.2. Performance Evaluation

The performance of the proposed inference approach in estimating the parameters of the proposed model and the corresponding standard errors is assessed as follows:

(1) Estimate the parameters for each data set and for each of 10 independent sets of permutations. Estimate the standard errors (s.e.) using the Godambe (sandwich) estimator.

(2) For each data set *s*, compute the mean estimate for each model parameter across the 10 random permutations used. Label this as MED, and then take the mean of the MED values across the data sets to obtain **a mean estimate**. Compute the **absolute percentage (finite sample) bias** (APB) of the estimator as:

$$APB = \left| \frac{\text{mean estimate - true value}}{\text{true value}} \right| \times 100$$

(3) Compute the standard deviation of the MED values across datasets, and label this as the **finite sample standard error or FSEE** (essentially, this is the empirical standard error).

(4) For each data set, compute the mean standard error for each model parameter across the 10 draws. Call this MSED, and then take the mean of the MSED values across the 30 data sets and label this as **the asymptotic standard error or ASE** (essentially this is the standard error of the distribution of the estimator as the sample size gets large).

(5) Next, to evaluate the accuracy of the asymptotic standard error formula as computed using the MACML inference approach for the finite sample size used, compute a **relative efficiency** (RE) value as:

$$RE = \frac{\text{ASE}}{\text{FSEE}} \times 100$$

(6) Compute the standard deviation of the parameter values around the MED parameter value for each data set, and take the mean of this standard deviation value across the data sets; label this as the **approximation error** (**APERR**).

3.3. Simulation Results

The results of the simulation experiments are presented in Table 1. These results indicate that the proposed method does very well in recovering the parameters, as can be observed by comparing

the mean estimates of the parameters with the true values. The absolute percentage bias (APB) is no more than 10% for any parameter (see column titled "Absolute Percentage Bias". The overall APB across all the parameters is only 3.1%. Among all the parameters, the dispersion parameter of the underlying negative binomial distribution, θ, is recovered least accurately with an APB value of 9.3%. The reason for this is that this parameter appears very non-linearly in the model system of Equation (4), and through the ψ_{q,l_q} threshold parameters. Besides, the threshold parameters do not change very substantially as the θ parameter increases and asymptotically approaches a Poisson distribution. As a result, it is difficult to pin down the value of the θ parameter through estimation, leading to higher APB values. Similarly, the parameter $l_{\tilde{\Sigma}_1 3}$ also has a relatively large APB of 8.3%. This parameter determines the Ξ part of the $\tilde{\boldsymbol{\Sigma}}_1$ covariance matrix that is responsible for generating the covariance between the MNP control type model with the count model of the number of crashes. Again, small changes in this element do not have too much effect on the likelihood function in Equation (7), and thus it is somewhat more difficult to pin down the value of this parameter. In any case, from a magnitude standpoint, it is clear that there is very little difference between the mean estimate of the parameters from the MACML method and the true parameter values.

The finite sample standard errors (FSSE) are small and are on an average about 13% of the true value of the parameters, indicating good empirical efficiency of the proposed estimator for the model. Another observation from the FSSE estimates is that these estimates (as a percentage of the mean estimates) are generally lower for the mean parameters of the vectors $\boldsymbol{\beta}_q$ and $\boldsymbol{\mu}_q$ (*i.e.*, for the for **b** vector and **d** vector elements) than for the corresponding Cholesky elements of the covariance matrices of these vectors (*i.e.*, for the $\boldsymbol{l}_{\tilde{\Gamma}}$ and $\boldsymbol{l}_{\tilde{\Omega}}$ vector elements). In particular, the ASE estimates (as a percentage of the mean parameter values) are, on average, about 13% for the mean parameters of the vectors $\boldsymbol{\beta}_q$ and $\boldsymbol{\mu}_q$ relative to 17.5% for the Cholesky elements of these vectors. This is to be expected since the mean parameters enter into the likelihood function of Equation (7) rather linearly (through the mean vector $\tilde{\mathbf{H}}_q$), while the Cholesky elements enter much more non-linearly through the $\boldsymbol{\Sigma}_q$ covariance matrix. Among the other parameters, the ASE values for the θ and $l_{\tilde{\Sigma}_1 3}$ parameters are relatively high in both

absolute magnitude as well as a percentage of the mean estimate (the FSEE of the θ parameter is 20% of the mean θ estimate, while the FSEE of the $l_{\tilde{\Sigma}_1 3}$ parameter is 21% of the mean $l_{\tilde{\Sigma}_1 3}$ estimate), reinforcing the finding earlier that these parameters are more difficult to recover than other parameters.

The finite sample standard errors and the asymptotic standard errors obtained are also close, with the relative efficiency (RE) value between 0.8-1.2 for all but three parameters. The efficiency values are outside this range for the Cholesky elements $l_{\tilde{\Gamma}1}$ and $l_{\tilde{\Sigma}_1 1}$, and the dispersion parameter θ. However, for the Cholesky parameters, the relatively high RE value is an artifact of the very low FSEE values in the first place. Indeed, in terms of testing whether these parameters are different from zero, there is no change in inference whether one uses the FSEE or the ASE. Compared to all other parameters, the difference between the FSEE and ASE is highest (in magnitude) for the θ parameter. This is consistent with the findings from earlier that the θ parameter appears to be the most difficult to pin down, which implies that accurate and precise estimation of this parameter, as well as the accuracy of the ASE estimate for this parameter from finite samples, would be particularly improved with large sample sizes. But, again, even for this parameter, there is effectively no difference in inference whether the FSEE or ASE is used to test if the model collapses to a Poisson model or not (as indicated earlier, the Poisson model results if $\theta \to \infty$; even when one uses a very conservative bound of 4 to replace ∞ in this test, both the FSEE and ASE return the same result that the Poisson model is rejected). Overall, across all parameters, the average relative efficiency is 1.11, indicating that the asymptotic formula is performing well in estimating the finite sample standard error. Further, as for the FSEE values, the ASE estimate, on average across all parameters, is only 16.5% of the mean estimate, indicating very good efficiency even using the ASE estimate for the FSEE.

Finally, the last column of Table 1 presents the approximation error (APERR) for each of the parameters, because of the use of different permutations. These entries indicate that the APERR is on an average only 0.010 and the maximum is only 0.028. More importantly, the approximation error (as a percentage of the FSEE or the ASE), averaged across all the parameters, is of the order of 7.5% of the sampling error. This is clear evidence that even a single permutation (per observation) of the approximation approach used to evaluate the MVNCD function provides adequate precision, in the sense that the convergent values are about the same

for a given data set regardless of the permutation used for the decomposition of the multivariate probability expression.

3.3.1 Effects of Ignoring the Joint Distribution of the Error Structures

This section presents the results of the estimation when the endogeneity of the treatment variable (in our case, the endogeneity of the intersection control type) on the count outcome (in our case, the frequency of crashes at an intersection) is ignored. That is, we examine the effect of constraining $l_{\tilde{\Sigma}_1 3}$ to zero when the data actually reflects a correlation. The simulation results for this restricted model (which we label as the "independent model") is presented in Table 2. For comparison purposes, we also present the results of the joint model proposed in the current paper. For the purpose of Table 2, we run only 50 estimations for each of the independent and joint models, corresponding to each of the 50 data sets generated as per the experimental design of Section 3.1. That is, we use only one set of permutations per data set to evaluate the MVNCD functions and do not run ten estimations per data set with different sets of permutations. We do so rather than run 10 replications because, as we presented in the earlier section, the approximation error in the parameters is negligible for any given data set. However, for each data set, we use the same set of permutations for the joint model and the independent model, so that we are able to appropriately compare the ability to recover parameters from the two models. In addition, we also compute a likelihood ratio test (LRT) statistic between the joint and independent models for each data set. This statistic needs to be compared against the table chi-squared value with one degree of freedom, which is equal to 3.84 at the 5% level of significance. In this paper, we identify the number of times (corresponding to the 50 data sets) that the LRT value rejects the independent model in favor of the joint model.

As can be observed from Table 2, the APB values are higher for almost all parameters in the independent model. The overall APB across all parameters is 7.5% in the independent model relative to only 2.6% in the joint model (as discussed earlier, the joint model results in Table 2 are slightly different from those in Table 1 because we use only one set of permutations for the estimates in Table 2). Moreover, when confining attention to the count model estimates, the APB in the independent model rises to 10.4% relative to only 2.1% for the joint model. This is a more appropriate comparison than comparing the APB across all parameters, since ignoring the covariance (when present) renders the count model estimates inconsistent while leaving the MNP

estimates still consistent. Further, the largest bias is in the d3 parameter, as can be observed in the sharp decrease in the d3 parameter estimate to a value of -0.755 in the independent model (leading to a 32.5% APB for this parameter in the independent model). This is to be expected because the $l_{\tilde{\Sigma}_1 3}$ element is associated with the correlation between (1) the utility difference of the third alternative from the first and (2) the count propensity effect of the third alternative relative to the first. In particular, if unobserved factors that increase the likelihood of the third treatment (or alternative) relative to the first also increase the propensity of the count variable, ignoring this correlation will lead to an understatement of a "true" reduction in the count propensity due to the third treatment. Intuitively, if, for example, full phased traffic signals are placed at intersections with intrinsically high propensities of incidents (due to unobserved factors of these intersections), then ignoring this positive correlation will dilute the "true" reduction in crashes due to converting an intersection with no control to that with a full-phased traffic signal.

The LRT test toward the bottom of Table 2 clearly indicates that the joint model rejects the independent model in all the 50 data sets, further reinforcing the need to consider jointness in the MNP and count components when present.

4. APPLICATION TO INTERSECTION ACCIDENT COUNTS

4.1. Background

Motorized vehicle use is a routine part of life for most American households. Motorized vehicles enable face-to-face connectivity with far-flung friends or relatives, as well as facilitate participation in daily activities. However, motorized vehicle use is not without its risks. Motorized vehicle crashes can cause property damage, injuries, and fatalities. Indeed, the leading cause of death for people between the ages of 11 and 27 is motorized vehicle crashes, according to the National Highway Traffic Safety Administration (NHTSA) (NHTSA 2013). In 2011, an average of 89 people died daily from motor vehicle crashes, which is about 1 person every 16 minutes (NHTSA 2013). The Center for Disease Control estimates the 2005 lifetime costs for fatal and non-fatal motor vehicle occupant injuries at around $70 billion, an estimate which includes treatment costs, rehabilitation, and lost productivity (CDC 2011). The US National Safety Council estimates the average cost of a death from a motor vehicle crash at $1.42 million (2011). A crash with only property damage and non-disabling injuries has an average economic cost of $9,100 (National Safety Council, 2011). In 2011, 2.2 million people were non-fatally

injured in motor vehicle crashes, and another 3.7 million crashes led to property damages (NHTSA 2013).

Motor vehicle crashes often occur at intersections, as turning maneuvers and even proceeding straight through an intersection can bring a vehicle into conflict with other vehicles whose drivers are pursuing their own path of travel. Among all roadway-related crashes, NHTSA estimates that 40% occur at intersections (NHTSA, 2010). In the pool of serious intersection crashes (those involving one or more fatalities), 60% occur at urban intersections. It is no surprise, therefore, that many earlier transportation and roadway safety studies have examined the frequency and severity of intersection crashes as a function of intersection control characteristics, roadway design features, and traffic volumes, with the ultimate objective of suggesting possible countermeasures to reduce such crashes (see, for example, CPB, 2012, Haque *et al.*, 2010, Huang and Chin, 2010, Mitra, 2009, and Haleem and Abdel-Aty, 2010). However, none of these studies consider the potential endogeneity of "independent" variables such as intersection control type when modeling the frequency of crashes, as we do in the current paper. Alternatively, some studies have attempted to examine the effects of flashing lights through before-after studies at intersections that changed night-time flashing conditions to regular signal phasing (see, for example, Gaberty and Barbaresso, 1987, Barbaresso, 1987, Polanis, 2002, and Srinivasan *et al.,* 2008). But these studies once again do not consider the possibility that the flashing lights may have been placed in the before case at a sample of intersections in a selective manner, and thus they ignore potential endogeneity considerations. Besides, these studies are typically based on observations on a very small number (12 to 15) of intersections.

4.2. The Data

Our analysis uses crash data drawn from the Texas Department of Transportation (TxDOT) Crash Record Information System (CRIS). The CRIS compiles police and driver reports of crashes into multiple text files, including complete crash, vehicle, person, and weather-related

details for each crash.[3] In this study, crashes designated as intersection or intersection related were extracted from the CRIS data base.[4] We further confined attention to intersections from Irving, Texas. This is because the CRIS does not include traffic flow information on intersection approach movements, a variable that has been well established as a key determinant of intersection crash risk propensity (see Quddus, 2008 and CPB). However, the North Central Texas Council of Governments offers a listing of 24-hour weekday traffic counts in its planning area, as well as an online map interface to visualize the count records. The visualization showed very good coverage across the many (>1000) intersections in the city of Irving. From this traffic count data, we extracted out the two-way flows on the approach streets for each intersection in Irving (of course, at some intersections, such as T-intersections or one-way approaches to an intersection, total one-directional flows were used as the approach volume on the appropriate approach streets). The sum of the flows on all approach streets to an intersection was computed to obtain an estimate of the total daily entering traffic at the intersection.

For this study, we further confined attention to 1032 intersections from Irving that had the same intersection control type installed (as obtained from the crash records) throughout the CRIS data collection period of 2003-2009.[5]

[3] The Texas law enforcement agency officially maintains the records of those crashes reported by police and drivers that involve property damage of more than $1,000 and/or the injury of one or more individuals (of course, records of crashes that involve fatalities on the spot are also maintained). Thus, the CRIS does not include minor crashes that involve only property damage of less than $1,000. However, in the rest of this paper, we will not belabor over this distinction, and will use the CRIS crashes as the measure for all crashes.

[4] TxDOT defines a crash as being intersection-related if it occurs within the curb-line limits of the intersection or on one of the approaches/exits to the intersection within 200 feet from the intersection center point.

[5] A handful of residential and local collector road intersections among the 1032 intersections did not have (in the NCTCOG data base) traffic counts on one or more approach streets. In such instances, rather than discarding such intersections, we used imputation procedures to estimate approach volumes. Also, for 52 of the 1032 intersections, the CRIS data base did not have a control type. For these, we identified the control type through visual identification based on Google Street View. Of the 52 intersections, two intersections were signalized, 44 were stop controlled, three were yield-controlled, and the remaining three intersections fell into the 'no control' category. These control types had not changed over time (based on visual identification over time). A final note regarding control types. Intersections that operated under flashing lights on one or more approaches for part of the nights (and as normal signals during the day) were designated as being under flashing light control. This is because we are using a 24-hour period for analysis, and also because intersection locations operating under flashing lights during any time may lead to driver confusion even during regular signal phasing operation (see Hunter *et al.,* 2011).

4.2.1. The Dependent Variable Statistics

The dependent variables in the model are the intersection control type and the count of crashes. In our analysis, the latter variable corresponds to crashes in 2008 at each of the 1032 intersections in the sample.

The intersection control type statistics in Table 3a indicate that about 35% of the intersections do not have any traffic control (such as residential street intersections, or intersections with minimal form of traffic control devices such as center stripes/dividers, turn marks, and marked lanes on one or more approaches). A substantial fraction of intersections are also controlled by stop signs (36.5%), while flashing light intersections are also represented in the sample (3.0%).

The distribution of crashes across the intersections is provided in Table 3b, and ranges from 0 to 21. The total number of crashes across the 1032 intersections is 959, yielding a mean number of crashes of 0.929 per intersection. A large fraction of the intersections (651 of the 1032 intersections, or 63% of intersections) did not have any reported crashes in 2008. As discussed in Section 1, our recasting of the count model as a GORP model incorporates the flexibility of handling such "excess" zeros if the explanatory variables are unable to accommodate this spike in zero count.

4.2.2. Independent Variable Characteristics

The characteristics of the 1032 intersections, in terms of the independent variables used in the model, are summarized in Table 4. Of the intersections in the sample, a third of the intersections have three entering roads in a T configuration, while 7% of the intersections have three entering roads arranged in a Y configuration, and the remaining 60% have 4 entering roads. In terms of approach roadway type combination, 84.6% of the intersections have approach streets all of which are city streets. Of the remaining 15.4% of the intersections, at least one approach roadway is not a city street (primarily state highways, farm to market roadways, and county roads).

The statistics for the approach roadway alignment indicate that a majority of intersections (81.4%) are composed of straight and level vertical grade approaches on all approach roadways, but 7.9% of intersections have at least one approach with a vertical grade, 6.3% include at least

one approach roadway with a horizontal curve, and another 4.4% have at least one roadway with a crest[6]. Also, a vast majority (92.8%) of intersections are not on frontage roads.

The traffic daily entering volume at an intersection is an important determining factor in explaining both the intersection control type as well as the count of crashes at the intersection. This volume varies between 200 and 78,288 vehicles.

The final variable used in the analysis is a Flow Split Imbalance factor (FSIMB), as introduced by CPB, and formulated as follows:

$$FSIMB = \frac{V_1 - V_2}{V_1 + V_2}, \tag{11}$$

where V_1 and V_2 correspond to the daily traffic volumes on the major and minor roadways, respectively ($V_1 \geq V_2$). The FSIMB factor can take a value between zero (when there is no imbalance in flows on the approach roads) and one (when there is complete imbalance in the flows, theoretically obtained when there is zero flow on the minor road). In the sample, the mean FSIMB statistic is 0.626, indicating that, on average, the major road volume is 4.35 times the minor road volume at the sampled intersections.

4.3. Variable Specification and Model Formulation

The variables in Table 4 are of two distinct types – the number and configuration of entering roads, the approach roadway type combination, the approach roadway alignment, and whether the intersection location is on a frontage road are categorical variables. On the other hand, the daily entering volume is a continuous variable, while the FSIMB factor is continuous but bounded between 0 and 1. The base category for the categorical variables is as follows: (a) four entering roads (for the number and configuration of entering roads), (b) At least one approach road is a non-city street (for approach roadway type combination) (b) straight and level approach streets (for approach roadway alignment), and (d) none of the approaches is a frontage road (for whether an intersection location is on a frontage road). For the continuous daily entering volume

[6] The CRIS data base does not define what exactly is a vertical grade and how if differs from a crest. This designation was made by the peace officer responsible for recording details of the crash. In general, an approach roadway is considered to have a vertical grade if there is a reasonable vertical slope that has not yet crested before the intersection point. On the other hand, an approach roadway is considered to have a crest if it has a vertical slope that crests very close to the intersection point. The specifications of what constitutes a reasonable vertical grade and how close should the crest be to the intersection for designation as a crest are not objective, but rather based on the subjective perspective of the peace officer.

variable, we attempted alternative functional forms, including the linear form, the natural logarithm form, a piecewise linear form, and dummy variables for different threshold values. Further, various interactions of the continuous and the categorical variables were also considered whenever adequate observations were available to test such interaction effects, such as between traffic volume and roadway alignment, and number of entering roads and traffic volume. But none of these interaction terms came out to be statistically significant. The final model was obtained based on goodness of fit, intuitiveness, and parsimony considerations. In some cases, we retained variables even if they were not statistically significant at the 5% level of significance because of intuitive considerations and also because the results should be useful for further exploration of crash determinants in future studies.

4.4. Model Estimation Results

We first discuss the effects of variables in the MNP model of intersection control type as well as the elements of the covariance matrix Λ_1 for the error differences of the control type propensities (Section 4.4.1), next the impacts of variables on the long term crash propensity in the count model (Section 4.4.2), subsequently the determinants of the thresholds in the count model (Section 4.4.3), and finally the elements of the covariance matrix Ξ between the control type propensity differentials and the count long-term propensity (Section 4.4.4).

4.4.1 Intersection Control Type MNP model

Table 5 presents the MNP model results for the joint model proposed in the current paper. The MNP model is estimated with the "no traffic control" alternative as the base alternative. If a '-' appears for a row variable in Table 5 corresponding to a column alternative, it implies that the column alternative forms a base alternative along with the "no traffic control" alternative (for the effect of the row variable).

The constants do not have any intuitive interpretation; they combine any unobserved biases with adjustment factors that accommodate the continuous sample values of the entering volume and FSIMB variables. Among the other variables, a T-intersection has, on average, a lower propensity of being under stop sign control relative to being under yield sign control or no control at all; however, there is heterogeneity in this coefficient, as observed by the (marginally significant) standard deviation on this coefficient. According to the results, 90% of T-

intersections have a higher propensity of being in "no control" or "yield control" states than in a "stop sign control" state. A T-intersection is also less likely, on average, to be under flashing light control relative to the no control or yield control states, and the least likely to be in a signal controlled state. Finally, under the category of "number and configuration of entering roads", the results indicate that Y-intersections have the lowest propensity to be under a signal control state, and the most likely to be in a yield control state. All of the above results are not surprising, given that T and Y intersections tend to be at low-volume, low-speed residential street locations, where "no control" or yield control are likely to be the most prevalent (see the Manual on Uniform Traffic Control Devices or MUTCD; FHWA, 2009).

When all the approach streets of an intersection are city streets (as opposed to one or more of the approach streets being a non-city street), the intersection is more likely to be controlled by a stop sign. In the category of approach roadway alignment variables, the presence of a vertical grade increases the likelihood of having flashing lights at the intersection, while the presence of a crest elevates the probability that flashing lights or signal lights will be placed at the intersection. These results are reasonable, because the presence of vertical grades and crests reduces sight distance. In such conditions, forms of control such as stop signs that can be seen only when close enough to the intersection may not be adequate. On the other hand, flashing lights and signals can be mounted high and be seen from far away, surmounting vertical grade or crest-related visual limitations. Indeed, the MUTCD advises the use of several warning signs to warn the motorists of the dangers ahead.

The total daily entering traffic volume variable, as defined in Section 4.2, was introduced in several ways, but the best data fit was obtained using a simple logarithmic transformation of the daily entering volume. High entering volume locations are most likely to be controlled by signal control. This is consistent with the warrants for signal placement, as documented in the MUTCD (FHWA, 2009). High entering volume locations are least likely to be under a no control or a stop-sign control state. Finally, after controlling for total entering traffic volume, Table 5 indicates that the split of traffic between the major and minor roads also plays an important role in control type placement. In particular, a higher value of the FSIMB factor (minor road approach volume being very less relative to major road approach volume) implies a higher propensity of the intersection being in no control than in other states, yield control relative to other forms of control, stop control relative to flashing lights and signals, and flashing lights

relative to complete signal phasing. This is intuitive, since, as documented in the MUTCD, the hierarchy of control placement is inversely related to the level of flow imbalance (a higher imbalance leads to fewer conflict points at the intersection).

A general specification was considered for the covariance matrix $\mathbf{\Lambda_1}$ of the error differences of the control type propensities (taken with respect to the "no control" alternative's propensity). But, in our empirical context, we could not reject the null hypothesis that this matrix has ones in its diagonals and 0.5 entries in its off-diagonals. This, of course, is equivalent to an independent and identical distribution specification for the original error terms (that is, the $\mathbf{\Lambda}$ covariance matrix of the original error terms turns out to be an identity matrix multiplied by 0.5). However, this result is specific to the current empirical context. In general, one needs to specify the more general model proposed in this paper before testing for more restrictive variants. Also, it should be noted that the random parameter on the T-intersection variable does generate heteroscedasticity across the overall random components of the control type propensities.

4.4.2. Long Term Crash Risk Propensity

The constant term in the long term crash risk propensity (or simply crash risk propensity from hereon) is normalized to zero, as discussed in Section 2.2 (see the column entitled "long term propensity" under "count parameters" in Table 5). The other results indicate that intersections with T configurations have a higher crash propensity relative to Y intersections and regular four-legged intersections. This may seem inconsistent with the studies of Abdel-Aty and Wang (2006) and Srinivasan *et al.* (2008), both of which found that three-legged intersections are less prone to crashes than four-legged intersections (presumably because the former type of intersections presents "fewer vehicle conflict points" than the latter). But these earlier studies included only signalized intersections in their samples, while our study includes all kinds of control types at intersections. Further, the net effect at T intersections will be a combination of the crash risk propensity and the threshold effect, which we discuss in the next section. As we will note when we compute the overall elasticity effects, we do find that three-legged intersections are less prone to crashes relative to four-legged intersections. But our study disentangles the "risk" effect from the "translation of the risk to actual crash outcome" effect.

Intersections with approach streets that are all city streets have a lower crash risk propensity, perhaps because of their location and the generally lower speeds of travel on these

city streets relative to non-city streets. As expected, intersections with approaches involving vertical grades, horizontal curves, or a crest are more prone to crashes, because of limiting sight distance considerations (see Savolainen and Tarko, 2005 for a similar result). The problem seems most acute for the case of crest approaches, consistent with such approaches creating the most problems associated with sight distance (but see also discussion in the next section). Table 5 also suggests that a higher FSIMB factor increases the crash risk propensity; that is, intersections where the volumes on the minor and major roadways are relatively unbalanced are more crash-prone than intersections where the minor and major roadways have about the same traffic volume (this is after controlling for total entering traffic volume). This perhaps reflects the fact that such intersections tend to require more gap-related judgment on the part of those approaching the intersection on the minor roadways. However, this effect is not statistically significant.

The traffic control type variables are considered to be endogenous in the proposed model. Thus, the effects of the control type variables in Table 5 are "cleansed" of unobserved factors that generate a correlation between the propensity of a specific control type being installed and the crash risk propensity at an intersection. For completeness, we report the results for the effects of all control type dummy variables on crash risk propensity in Table 5, though the ones for yield and stop control types are (statistically speaking) no different from having no control. Also, while we tested for unobserved heterogeneity in the effects of these control type variables (treatment effects) rather than *a priori* positing fixed treatment effects, our results did not find any statistically significant unobserved heterogeneity effects. The positive and highly statistically significant parameters in Table 5 for flashing lights and signal control indicate a higher crash propensity for these types of controls relative to when there is no control (note that this is after controlling for a whole suite of other factors, as already discussed). This effect is particularly large for flashing lights (red or yellow) on one or more approaches, an observation also made by Poškienė and Sokolovskij (2008). This may be a reflection of confusion on the part of drivers regarding how to respond on seeing a flashing light

4.4.3 Threshold Parameters

The thresholds are responsible for the "instantaneous" translation of the long-term crash risk propensity to whether or not a crash occurs at any given time (that is, they determine the

mapping of the latent propensity to the observed count outcome). The thresholds in Equation (4) are functions of the $\boldsymbol{\varphi}$ and $\boldsymbol{\gamma}$ vector, and the θ scalar. Among these, the elements of the vector $\boldsymbol{\varphi}$ provide flexibility to accommodate spikes in specific values of counts. However, in our estimations, none of these elements were statistically significant, indicating that there was no need to adjust the thresholds beyond accounting for the effects of exogenous variables on these thresholds through the $\boldsymbol{\gamma}$ vector. The column labeled "threshold estimates" in Table 5 represents these effects of exogenous variables.

The constant does not have any particular interpretation. For the other variables, a positive coefficient shifts all the thresholds toward the left of the crash propensity scale, which has the effect of reducing the probability of zero crashes (see CPB). On the other hand, a negative coefficient shifts all the thresholds toward the right of the crash propensity scale, which has the effect of increasing the probability of zero crashes. The effect of the variable corresponding to T intersection configuration indicates an increase in the probability of zero crash outcomes (decrease in probabilities of non-zero crash outcomes) at intersections with T configurations relative to other intersection configurations (for a given long-term crash propensity). That is, the translation of risk into the occurrence of a crash is depressed for T intersections, perhaps because the fewer conflict points (compared to four-legged intersections) and more conventional flow (compared to Y intersections) at T intersections allows drivers to more easily take evasive maneuvers even as they see a crash in the making. In combination with the approach roadway configuration effects on the long term crash propensity, the net implication is that, while the risk of crashes is higher at T intersections, these intersections also offer more of "out" to prevent a crash.

The estimate on the "all approach roadways are city streets" variable suggests that, given two intersections with the same crash risk propensity, an intersection with all city street approaches has a lower probability of zero crashes (higher probability of non-zero crashes) than an intersection with at least one non-city street approach. This indicates that motorists have less of an "out" as a crash starts to develop on city streets relative to on non-city streets. Thus, motorists may not be able to get into a different lane or maneuver in a different direction because of other simultaneous movements taking place and because of the clearly delineated and channeled traffic movements at intersections with city street approaches. The effect of the decreased risk propensity at intersections with city street approaches, combined with the elevated

probability of a crash outcome given a certain risk propensity at such intersections, is another indication (as with the T intersection case discussed earlier) of the complex interplay that is at work in terms of crash frequency at intersections. A similar situation is also at work, but in the reverse, for intersections with a crest approach and intersections with a high flow split imbalance (the last parameter under the column "threshold estimates"). While crest intersections and intersections where the volumes on the minor and major roadways are relatively unbalanced are more crash-prone than non-crest intersections, and intersections where the minor and major roadways have about the same traffic volume, respectively, it also appears that crest intersections and intersections with unbalanced volumes offer more of an option to "wiggle out" from a crash waiting to happen. The other results indicate a lower probability of zero crashes (higher probability of non-zero crashes) for intersections with a frontage road approach (compared to intersections without a frontage road approach) and intersections with high entering volumes. These results are not surprising, and suggest more difficulty in preventing a crash at such intersections due to more conflict points (see Haleem *et al*., 2010 and Oh *et al*., 2009) and less room to maneuver out of crash situations.

Finally, the last row presents the estimate of θ and its standard error. To test against a Poisson distribution assumption within the cumulative inverse function in the thresholds, one can obtain the inverse of θ and its standard error (using the delta method). This estimate can then be tested against a value of zero. The corresponding statistic is 4.83, clearly rejecting the simple Poisson distribution assumption in favor of the negative binomial distribution used here.

4.4.4 Error Covariance

Many different specifications were considered for the covariance vector $\boldsymbol{\Xi}$ between the control type propensity differentials (in the MNP model) and the crash risk propensity error (in the count model). Of these, only the elements corresponding to the covariance between the stop sign and flashing light propensity errors (relative to the "no control" propensity error) with the crash risk propensity turned out to be of some statistical significance. The covariance between the stop sign propensity (relative to the "no control" propensity) and the crash risk propensity was 0.208 (t-statistic of 1.23) and the covariance between the flashing light propensity (relative to the "no control" propensity) and the crash risk propensity was -0.439 (-2.82). A convenient, and not unreasonable, way to interpret these covariance terms would be to assume that the error

covariance between the "no control" propensity and the crash risk propensity is zero. Then, the implication is that unobserved factors that increase the propensity of stop control being installed at an intersection also increase crash risk propensity at the intersection. But unobserved factors that increase the propensity of flashing lights being installed decrease crash risk propensity. The latter result suggests that flashing lights are generally located at relatively low crash risk propensity locations (after controlling for a suite of observed factors that impact crash propensity). This is an interesting result, and suggests that drivers internalize risky situations and drive cautiously at such intersection locations. This reduces crash risk propensity (not unlike the case of severe-injury crashes decreasing at times of rainy/snowy conditions because motorists drive more slowly during such conditions; see, for example, Khattak and Knapp, 2001). At the same time, traffic engineers may also put up flashing lights at what they deem to be high crash risk intersections. The net result is that flashing lights may get placed, in a twist of what is intended, at locations of low frequency of crashes. This observation is also of practical value. It suggests that traffic engineers may be placing flashing lights at what they (and motorists) perceive as locations that may be prone to crashes, but because motorists also perceive this and drive more cautiously, the net result is that such intersections actually have a lower crash risk. The specific implication is that rather than deciding on the type of control at an intersection based primarily on limited sight distance or other geometry/volume considerations, traffic engineers may want to consider placing more emphasis on observed crash frequency in deciding the control type. Also, a broader implication is that countermeasures to reduce crashes should perhaps not be targeted as much on what are readily perceived by motorists to be risky situations, but focus on situations that are inappropriately perceived by motorists as safe situations when there are hidden dangers lurking. That is, the focus on countermeasures should be on those intangibles that do not get registered as being risky situations, rather than on geometry and other obvious (to motorists) "risky" factors.

Another observation from the results is that ignoring the unobserved covariance between the control type propensities and crash risk propensity (that is, ignoring the endogeneity of control type) leads to a substantially under-estimated positive effect of flashing lights on crash propensity. Specifically, in an independent model that ignores the above covariance, the coefficient on the flashing light control variable was only 0.985, instead of 2.136 in the joint model of Table 5. That is, the negative covariance between the flashing light propensity and

crash risk propensity dilutes the "true" positive causal impact of flashing light signals on crash propensity. The results reinforce the results from earlier studies (for example, Srinivasan *et al*., 2008 and CPB, 2012) that flashing lights increase the number of crashes. However, the results are indicative that this is not because flashing lights are placed at high crash risk propensity locations in the first place (indeed, our results suggest the opposite is true), and points to driver confusion when an intersection operates in flashing light mode. If anything, the results imply that the driver confusion due to flashing lights may have been understated in earlier studies that ignore the endogeneity of control type.

4.5 Assessing Elasticity Effects and Treatment Effects

4.5.1. Procedure for Elasticity Effects of Non-Signal Control Type Variables

The parameters on the exogenous variables in Table 5 do not directly provide the magnitude of the effects of variables on crash frequency. To do so, we compute the aggregate-level "elasticity effects" of variables to discern the magnitude and direction of variable impacts. Specifically, we examine the effects of variables on the expected number of crashes at each intersection, given the intersection characteristics. This effect itself can be computed unconditionally of the traffic control type at an intersection, or after controlling for the moderating effect of traffic control type. The unconditional effect may be computed by writing the expected number of crashes at intersection q unconditional of control type as:

$$E(y_q) = \sum_{k=0}^{\infty} \sum_{i=1}^{I} \left(P(a_{qi} = 1, y_q = k) \right) \cdot k, \tag{12}$$

where $P(a_{qi} = 1, y_q = k)$ is the probability that intersection q has control type i and k crashes. Although the summation in the equation above extends until infinity, we consider counts only up to k = 21, which is the maximum crash frequency observed in the dataset (in all subsequent formulas, we will retain the summation to ∞, though it will be understood that the summation is taken only until k = 21). This should not affect the elasticity computations because the probabilities associated with higher crash outcomes are very close to zero. Note also that, in Equation (12), the impact of a variable on the expected number of crashes may be through the $\mathbf{x}_q$ vector, through the $\mathbf{w}_q$ vector, through the $\mathbf{z}_q$ vector, or through more than one of these vectors. Alternatively, one can compute the effect of a variable after controlling for the moderating effect of control type by writing the expected number of crashes at intersection q as:

$$E(y_q) = \sum_{k=0}^{\infty} \left(P(y_q = k)\right) \cdot k. \tag{13}$$

In either case, the expected aggregate numbers of crashes is then computed by summing the above intersection-level number of crashes across all intersections Q. With the preliminaries above, one can compute the aggregate-level "elasticity" as a measure of the aggregate percentage change in crash frequency due to a change in an exogenous variable. For dummy variables, the procedure is as follows: (1) set the value of the dummy variable to zero for all intersections in the sample and compute the expected aggregate number of crashes, (2) set the value of the dummy variable to one for all intersections in the sample and compute the expected aggregate number of crashes, and (3) compute the effective percentage change in the expected total number of crashes across all intersections in the sample by taking the difference between the expected number of crashes obtained in step (2) and step (1) and dividing by the result from step (1). To compute the aggregate level "elasticity" effect of a multinomial exogenous variable (such as the number of entering roads), the procedure is as follows: (1) set the value of the multinomial variable to the base category for all intersections in the sample and compute the expected aggregate number of crashes, (2) set the value of the multinomial variable to each other non-base category for all intersections in the sample and compute the expected aggregate number of crashes for each of the non-base categories, and (3) compute the effective percentage change for each non-base category relative the base category. For continuous variables, we increase the value of the variable by 10% for each intersection and compute the percentage change in the expected total number of crashes per year across all intersections. For the FSIMB factor that is contained between 0 and 1, we increase the factor by 0.1 at each intersection.

The "elasticity" effects and their standard errors are presented in Table 6 for the non-control variables. The first entry in the table indicates that the number of crashes at T intersections is, on average, 34.6% less than the number of crashes at four-legged intersections unconditional of intersection control type, and 24.5% less than the number of crashes at four-legged intersections conditional on intersection control type. The higher magnitude of the unconditional elasticity effect is because T intersections are less likely to be controlled by flashing light and signal systems relative to four-legged intersections, and the presence of flashing light and signal systems increase crash risk propensity. As a result, when unconditioned out, the results recognize the lower likelihood of flashing light and signal control systems at T

intersections, while the conditional elasticities do not take this into consideration. Other entries may be similarly interpreted. The zero entry for the conditional elasticities for Y intersections is because the results indicate no statistically significant differences between Y intersections and four-legged intersections in the count model (after controlling for the intersection control type). However, signal control is unlikely to be in place at Y intersections (see the MNP model results), which, because of the higher crash risk propensity at signal controlled intersections, implies a lower (unconditional) count of crashes at Y intersections relative to four-legged intersections (as reflected by the entry "-6.8", though this is not statistically significant).

The results also summarize the effects of other variables that have opposite effects on the crash risk propensity and on the thresholds. For instance, if all approach roadways are city streets, this decreases crash propensity, but also reduces the probability of zero crashes through the threshold effects. When these effects are taken together, and over the entire population of intersections, the results suggest that, in general, an intersection with all city street approaches is 13.4% more likely to have crashes compared to an intersection with at least one non-city street. Other results may be interpreted similarly.

Overall, the results reveal that intersections with a crest on one approach (or on more than one approach) and intersections with a frontage road approach are the most crash-prone among the set of categorical independent variables, suggesting particular attention to such intersections. Specifically, the number of crashes is projected to be about 150% (or about 2.5 times) higher, on average, at an intersection with a crest approach relative to an intersection with straight and level approaches. Similarly, the number of crashes is projected to be again about 150% (or about 2.5 times) higher, on average, at an intersection on a frontage road than at an intersection that is not on a frontage road. While the crash-related effect of a crest approach may be attributed to sight distance limitations, the increased crash rate at frontage road intersections may be perhaps because motorists are not reducing speed enough after exiting off a highway (and as they approach an intersection on a frontage road). Further investigation of this effect will be helpful in improving intersection designs as well as for appropriate outreach and dissemination campaigns to inform the driving public.

4.5.2 Procedure for Treatment Effects for Traffic Control Type Variables

The observed data for each intersection includes the installed traffic control type and the frequency of crashes. Using the proposed model, we would like to assess the impact of traffic control type (the "treatment") on crash frequency ("the outcome"), after controlling for other observed and unobserved variable effects. An important measure to do so is the Average Treatment Effect (ATE) (see Heckman and Vytlacil, 2000 and Heckman *et al*., 2001). The ATE measure provides the expected crash frequency change for a random intersection if it were controlled by a specific control type *i* as opposed to another control type $j \neq i$. The measure is estimated as follows:

$$\hat{\text{ATE}}_{ij} = \frac{1}{Q}\sum_{q=1}^{Q}\left(\sum_{k=0}^{\infty} k \cdot \left[P(y_q = k \mid a_{qi} = 1) - P(y_q = k \mid a_{qj} = 1)\right]\right). \quad (14)$$

The analyst can compute the ATE for all the combinations of control types. Here, we focus on the common case of deciding between a flashing light operation and a full signal operation. According to the joint model (that jointly models intersection control type and crash outcomes), if a randomly selected (after controlling for other factors) intersection is controlled by a flashing light control rather than a full signal control, then the intersection is likely to have, on average, 2.9 more annual crashes. The t-statistic for the test against zero is 2.8, indicating that flashing lights pose significantly more risk of traffic accidents than signal control. In terms of percentage, a flashing light controlled intersection, on average, will have 208% more crashes (standard error of 89%) relative to a fully signal controlled intersection; that is, the count of crashes is projected to be 3 times higher at a flashing light-controlled intersection relative to a fully signal controlled intersection (after controlling for other factors). Hunter *et al*. (2011) also report a similar finding, and attribute this to confused driver behavior as drivers approach flashing lights. Interestingly, and consistent with the discussion in Section 4.4.4., the independent model (that ignores the endogeneity of intersection control type when modeling crash outcomes) shows a relatively small and statistically insignificant annual crash increase of 0.41 crashes (33%) if an intersection is controlled by flashing lights rather than signals. This demonstrates the importance of accommodating endogeneity considerations, and suggests that traffic engineers should think more than twice before using flashing light control.

5. CONCLUSIONS

In this paper, we propose a formulation and estimation approach for count data models with endogenous covariates. Our parametric multinomial discrete-count model uses a general multinomial probit (MNP) specification for the treatment and ties this MNP model with a count outcome model. The approach uses a recasting of a univariate count model as a generalized ordered-response probit (GORP) system, and allows random response variations (or unobserved heterogeneity) in the sensitivity to exogenous factors in both the treatment component as well as the outcome component, as well as accommodates potential heterogeneity in the treatment effects themselves on the count outcome. To our knowledge, this is the first formulation and application of a flexible count outcome model with a multinomial probit selection model, which also accommodates unobserved heterogeneity effects. An analytic approximation for the multivariate cumulative normal distribution is used to estimate the proposed model.

A simulation experiment is undertaken to evaluate the ability of the analytic approximation to recover the model, as well as to assess the ability of the asymptotic standard errors from the analytic procedure to provide an estimate of the finite sample error for the typical size of samples employed in estimation. These experiments show that our estimation approach recovers the underlying parameters very well and is efficient from an econometric perspective. The experiments also reveal the biases that accrue in the parameter estimates if endogeneity in covariates is ignored.

The empirical analysis uses crash data from intersections in the City of Irving, Texas. The data is drawn from the Texas Department of Transportation (TxDOT) Crash Record Information System (CRIS). Several intersection characteristics are used as explanatory variables, and the control type indicators are treated as endogenous variables. The results show the important effects (on both the control type and count of crashes) of (a) the number and configuration of entering roadways, (b) whether the approach roadways are all city streets or not, (c) approach roadway alignment, (d) whether or not at least one of the approach roadways is a frontage road, and (e) total intersection traffic volume as well as the distribution of the volume between the major and minor roads. Specifically, the number of crashes is projected to be about 150% (or about 2.5 times) higher, on average, at an intersection with a crest approach relative to an intersection with straight and level approaches. Similarly, the number of crashes is again projected to be about 150% (or about 2.5 times) higher, on average, at an intersection on a

frontage road than at an intersection that is not on a frontage road. Also, a flashing light controlled intersection, on average, is predicted to have 208% more crashes relative to an observationally equivalent fully signal controlled intersection; that is, the count of crashes is projected to be 3 times higher at a flashing light-controlled intersection relative to a signal-controlled intersection.

A practical implication of our results is that countermeasures to reduce crashes should perhaps not target what are readily perceived by motorists to be risky situations, but rather focus on situations that are inappropriately perceived by motorists as safe situations when there are hidden dangers lurking. That is, the focus of countermeasures should be on those intangibles that do not get registered as being risky situations, rather than on geometry and other obvious (as perceived by motorists) factors.

To summarize, this paper proposes and demonstrates the use of a count model with endogenous covariates. While the empirical context in the current paper pertains to intersection crashes, the proposed model should be applicable in a wide variety of contexts and in many different fields.

ACKNOWLEDGEMENTS

The authors are grateful to Lisa Macias for her help in formatting this document.

REFERENCES

Abdel-Aty, M., Wang, X. (2006). Crash estimation at signalized intersections along corridors: Analyzing spatial effect and identifying significant factors. *Transportation Research Record*, 1953, 98-111.

Anastasopoulos, P.C., Mannering F.L. (2009). A note on modeling vehicle accident frequencies with random-parameters count models. *Accident Analysis and Prevention*, 41(1), 153-159.

Barbaresso, J.C. (1987). Relative accident impacts of traffic control strategies during low-volume nighttime periods. *ITE Journal*, 57(8), 41-46.

Bhat, C.R. (1998). Accommodating flexible substitution patterns in multi-dimensional choice modeling: formulation and application to travel mode and departure time choice. *Transportation Research Part B*, 32(7), 455-466.

Bhat, C.R. (2011). The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. *Transportation Research Part B*, 45(7), 923-939

Bhat, C.R., Eluru, N. (2009). A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. *Transportation Research Part B*, 43(7), 749-765.

Castro, M., Paleti, R., Bhat, C.R. (2012). A latent variable representation of count data models to accommodate spatial and temporal dependence: Application to predicting crash frequency at intersections. *Transportation Research Part B*, 46(1), 253-272.

Castro, M., Paleti, R., Bhat, C.R. (2013). A spatial generalized ordered response model to examine highway crash injury severity. *Accident Analysis and Prevention*, 52, 188-203

CDC (2011). Nonfatal motor vehicle occupant injuries and seat belt use among adults. *Centers for Disease Control and Prevention.*

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *The Review of Economic Studies*, 47(1), Econometrics Issue, 225-238.

FHWA, (2009). *Manual on Uniform Traffic Control Devices (MUTCD), 2009 Edition*. U.S. Department of Transportation, Washington D.C.

Gaberty, I.I., Barbaresso, J.C. (1987). A case study of the accident impacts of flashing signal operations along roadways. *ITE Journal*, 57(7), 27-28.

Greene, W.H., Hensher, D.A. (2010). *Modeling Ordered Choices: A Primer*. Cambridge University Press, Cambridge.

Haleem, K., Abdel-Aty, M. (2010). Examining traffic crash injury severity at unsignalized intersections. *Journal of Safety Research*, 41(4), 347-357.

Haleem, K., Abdel-Aty, M., Mackie, K. (2010). Using a reliability process to reduce uncertainty in predicting crashes at unsignalized intersections. *Accident Analysis & Prevention*, 42(2), 654-666.

Haque, M.M., Chin, H.C., Huang, H. (2010). Applying Bayesian hierarchical models to examine motorcycle crashes at signalized intersections. *Accident Analysis & Prevention*, 42(1), 203-212.

Heckman, J.J., Vytlacil, E. (2000). The relationship between treatment parameters within a latent variable framework. *Economics Letters*, 66(1), 33-39.

Heckman, J.J., Vytlacil, E. (2005). Structural equations, treatment effects, and econometric policy evaluation1. *Econometrica*, 73(3), 669-738.

Heckman, J., Tobias, J.L., Vytlacil, E. (2001). Four parameters of interest in the evaluation of social programs. *Southern Economic Journal*, 68(2), 210-223.

Huang, H., Chin, H. (2010). Modeling road traffic crashes with zero-inflation and site-specific random effects. *Statistical Methods and Applications*, 19(3), 445-462.

Hunter, M., Jenior, P., Bansen, J., Rodgers, M. (2011). Mode of flashing for malfunctioning traffic signals. *Journal of Transportation Engineering*, 137(7), 438-444.

Joe, H. (1995). Approximations to multivariate normal rectangle probabilities based on conditional expectations. *Journal of the American Statistical Association*, 90(431), 957-964.

Joe, H. (1996). Time series models with univariate margins in the convolution-closed infinitely divisible class. *Journal of Applied Probability*, 33(3), 664-677.

Keane, M.P. (1992). A note on identification in the multinomial probit model. *Journal of Business & Economic Statistics*, 10(2), 193-200.

Khattak, A.J., Knapp, K.K. (2001). Interstate highway crash injuries during winter snow and nonsnow events. *Transportation Research Record: Journal of the Transportation Research Board*, 1746, 30-36.

Mitra, S., (2009). Spatial autocorrelation and Bayesian spatial statistical method for analyzing intersections prone to injury crashes. *Transportation Research Record: Journal of the Transportation Research Board*, 2136, 92-100.

Munkin, M.K., Trivedi, P.K. (2008). Bayesian analysis of the ordered probit model with endogenous selection. *Journal of Econometrics*, 143(2), 334-348.

National Highway Traffic Safety Administration (NHTSA) (2010). Crash Factors in Intersection-Related Crashes: An On-Scene Perspective. Publication DOT HS 811366, National Center for Statistics and Analysis, National Highway Traffic Safety Administration, Washington D.C.

National Highway Traffic Safety Administration (NHTSA) (2013). Traffic Safety Facts: 2011 Data. Publication DOT HS 811753, National Center for Statistics and Analysis, National Highway Traffic Safety Administration, Washington D.C.

National Safety Council. (2011) Estimating the costs of unintentional injuries. Statistics Department, National Safety Council, Itasca, IL.

Oh, J., Washington, S., Lee, D. (2009). Expected safety performance of rural signalized intersections in South Korea. *Transportation Research Record: Journal of the Transportation Research Board*, 2114, 72-82.

Polanis, S.F. (2002). Right-angle crashes and late-night/early-morning flashing operation: 19 case studies. *ITE Journal*, 72(4), 26-28.

Poškiene, J., Sokolovskij, E. (2008). Traffic control elements influence on accidents, mobility and the environment. *Transport*, 23(1), 55-58.

Quddus, M. (2008). Time series count data models: an empirical application to traffic accidents. *Accident Analysis and Prevention*, 40(5), 1732-1741.

Savolainen, P.T., Tarko, A.P. (2005). Safety impacts at intersections on curved segments. *Transportation Research Record: Journal of the Transportation Research Board*, 1908, 130-140.

Solow, R.M. (1960). On a family of lag distributions. *Econometrica*, 28(2), 393-406.

Srinivasan, R., Council, F., Lyon, C., Gross, F., Lefler, N., Persaud, B. (2008). Safety effectiveness of selected treatments at urban signalized intersections. *Transportation Research Record: Journal of the Transportation Research Board*, 2056, 70-76.

Zavoina, W., McKelvey R.D. (1975). A statistical model for the analysis of ordinal-level dependent variables. *Journal of Mathematical Sociology*, 4(1), 103-120.

LIST OF TABLES

Table 1. Simulation Results for 50 Datasets of 2000 Observations Each

Parameter	Component of	Parameter Estimates			Standard Error Estimates			
		True	Mean Estimate	APB	FSSE	ASE	RE	APERR
b1	MNP	1.500	1.464	2.4%	0.175	0.177	1.02	0.013
b2	MNP	-1.000	-0.956	4.4%	0.155	0.125	0.80	0.008
$l_{\tilde{\Omega}1}$	MNP	1.000	0.983	1.7%	0.163	0.149	0.91	0.012
$l_{\tilde{\Omega}2}$	MNP	0.600	0.604	0.6%	0.141	0.138	0.97	0.010
$l_{\tilde{\Omega}3}$	MNP	1.100	1.028	6.5%	0.189	0.164	0.87	0.012
d1	Count	0.500	0.505	0.9%	0.046	0.045	0.99	0.001
d2	Count	-0.500	-0.502	0.4%	0.087	0.096	1.10	0.003
d3	Count	-1.000	-1.047	4.7%	0.118	0.137	1.16	0.011
$l_{\tilde{\Gamma}1}$	Count	0.500	0.492	1.6%	0.060	0.085	1.40	0.003
$l_{\tilde{\Gamma}2}$	Count	0.707	0.697	1.5%	0.158	0.164	1.04	0.006
$l_{\tilde{\Gamma}3}$	Count	1.000	1.040	4.0%	0.129	0.151	1.17	0.009
θ	Count	2.000	2.186	9.3%	0.396	0.561	1.41	0.028
□1	Count	0.750	0.755	0.6%	0.069	0.074	1.08	0.003
γ1	Count	0.500	0.494	1.3%	0.039	0.042	1.09	0.001
$l_{\tilde{\Sigma}_1 1}$	MNP	0.600	0.587	2.2%	0.081	0.127	1.56	0.015
$l_{\tilde{\Sigma}_1 2}$	MNP	0.800	0.776	3.0%	0.115	0.133	1.16	0.012
$l_{\tilde{\Sigma}_1 3}$	Joint	0.600	0.650	8.3%	0.128	0.142	1.11	0.026
Across all Parameters				3.1%	0.132	0.148	1.11	0.010

Table 2. Effects of Ignoring Endogenous Effects

Parameter	Component of	True	CEMPS		Independent	
			Mean Estimate	APB	Mean Estimate	APB
b1	MNP	1.500	1.463	2.5%	1.459	2.8%
b2	MNP	-1.000	-0.957	4.5%	-0.952	5.0%
$l_{\tilde{\Omega}1}$	MNP	1.000	0.984	1.6%	0.985	1.5%
$l_{\tilde{\Omega}2}$	MNP	0.600	0.601	0.2%	0.601	0.2%
$l_{\tilde{\Omega}3}$	MNP	1.100	1.034	6.4%	1.026	7.1%
d1	Count	0.500	0.504	0.8%	0.490	1.8%
d2	Count	-0.500	-0.504	0.9%	-0.548	8.8%
d3	Count	-1.000	-1.028	2.7%	-0.755	32.4%
$l_{\tilde{\Gamma}1}$	Count	0.500	0.481	3.9%	0.433	15.4%
$l_{\tilde{\Gamma}2}$	Count	0.707	0.700	1.0%	0.672	5.2%
$l_{\tilde{\Gamma}3}$	Count	1.000	1.033	3.2%	0.987	1.2%
θ	Count	2.000	2.117	5.5%	1.633	22.4%
□1	Count	0.750	0.752	0.3%	0.789	5.0%
γ1	Count	0.500	0.496	0.9%	0.504	1.0%
$l_{\tilde{\Sigma}_1 1}$	MNP	0.600	0.581	3.3%	0.573	4.6%
$l_{\tilde{\Sigma}_1 2}$	MNP	0.800	0.772	3.7%	0.762	4.9%
Overall mean value across parameters				2.6%		7.5%
Mean log-likelihood at convergence			-3265.9		-3276.9	
Number of times the likelihood ratio test (LRT) statistic favors the CEMPS model			All fifty times when compared with $\chi^2_{1,0.95} = 3.84$ value (mean LRT statistic is 21.3)			

Table 3a. Intersection Control Type

Type of Traffic Control	Percentage
No traffic control or minimal traffic control	34.6%
Yield sign	6.1%
Stop sign	36.5%
Flashing light	3.0%
Regular signal light	19.8%

Table 3b. Crash Frequency Distribution across Intersections

Number of Crashes	Number of Intersections	Percentage of total	Cumulative percentage
0	651	63.1%	63.1%
1	195	18.9%	82.0%
2	77	7.5%	89.4%
3	30	2.9%	92.3%
4	24	2.3%	94.7%
5	19	1.8%	96.5%
6	10	1.0%	97.5%
7	6	0.6%	98.1%
8	5	0.5%	98.5%
9	2	0.2%	98.7%
10	5	0.5%	99.2%
11	2	0.2%	99.4%
12	1	0.1%	99.5%
13	1	0.1%	99.6%
15	1	0.1%	99.7%
16	1	0.1%	99.8%
20	1	0.1%	99.9%
21	1	0.1%	100.0%

Table 4. Sample Characteristics (1032 Observations)

Variable	Sample share
Number and Configuration of Entering Roads	
Three	
T-intersection	32.9%
Y-intersection	7.0%
Four	60.1%
Approach Roadway Type Combination	
All approach roadways are city streets	84.6%
At least one approach roadway is a non-city street	15.4%
Approach Roadway Alignment	
Straight and level approach streets	81.4%
At least one approach has a vertical grade	7.8%
At least one approach roadway has a horizontal curve	6.3%
At least one approach roadway has a hillcrest	4.4%
Is Intersection Location on a Frontage Road	
None of the approaches is a frontage road	92.8%
At least one approach roadway is a frontage road	7.2%

Descriptive statistics

	Minimum	Maximum	Mean	Std. Dev.
Total daily entering volume (vehicles/day)	200.000	78288.000	14517.210	13235.758
Flow split imbalance (FSIMB) factor	0.000	0.997	0.626	0.306

Table 5. Estimation Results for the MNP and Count Model – Joint Model

Variables	Joint MNP and Count Model					
	MNL Parameters				Count Parameters	
	Yield	Stop	Flashing	Signal	Long-term propensity	Threshold Estimates
Constant	-2.887 (-4.42)	0.423 (2.38)	-4.619 (-4.39)	-7.798 (-13.97)	0.000 (fixed)	-5.070 (-5.15)
Number and Configuration of Entering Roads (base is "four entering roads")						
Intersection is a T-intersection	-	-0.629 (-1.67)	-0.609 (-2.43)	-0.993 (-6.32)	0.671 (2.06)	-1.252 (-2.15)
Standard Deviation	-	1.891 (1.45)	-	-	-	-
Intersection is a Y-intersection	0.863 (5.30)	-	-	-0.835 (-3.19)	-	-
Approach Roadway Type Combination (base is "at least one approach road is a non-city street")						
All approach roadways are city streets	-	0.279 (2.06)	-	-	-0.324 (-1.14)	0.570 (1.29)
Approach Roadway Alignment (base is "straight and level approach streets")						
At least one approach has vertical grade	-	-	0.469 (1.92)	-	0.460 (3.22)	-
At least one approach has horizontal curvature	-	-	-	-	0.454 (2.99)	-
At least one approach has a crest	-	-	0.829 (2.79)	0.416 (1.63)	1.996 (4.77)	-1.483 (-2.91)
Is Intersection Location on a Frontage Road (base is "none of the approaches is a frontage road")						
At least one approach roadway is a frontage road	-	-	-	-	-	1.025 (3.90)
Traffic Volume-Related Variables						
Logarithm of daily entering volume (veh/day)	0.257 (3.33)	-	0.494 (4.42)	0.953 (15.81)	-	0.487 (4.78)
Flow split imbalance (FSIMB) factor	-0.599 (-2.25)	-0.793 (-4.58)	-1.544 (-4.52)	-1.989 (-8.77)	0.424 (1.03)	-1.234 (-1.77)
Intersection Control Type (base is "no control")						
Yield control type	-	-	-	-	0.200 (1.09)	-
Stop control type	-	-	-	-	-0.164 (-0.66)	-
Flashing lights control type	-	-	-	-	2.136 (5.00)	-
Signal control type	-	-	-	-	0.675 (4.86)	-
θ					0.589 (Standard error: 0.201)	

Table 6. Elasticity Estimates of Non-Control Variables (estimate with std. error in parenthesis)

Variable	Unconditional	Conditional
Number and Configuration of Entering Roads (base is "four entering roads")		
Intersection is a T-intersection	-34.59 (9.68)	-24.52 (12.84)
Intersection is a Y-intersection	-6.81 (5.82)	0.00 (0.00)
Approach Roadway Type Combination (base is "at least one approach road is a non-city street")		
All approach roadways are city streets	13.41 (23.79)	20.95 (24.86)
Approach Roadway Alignment (base is "straight and level approach streets")		
At least one approach has vertical grade	81.42 (22.90)	58.38 (21.24)
At least one approach has curvature	57.80 (25.26)	53.11 (23.20)
At least one approach has a crest	150.36 (32.85)	97.10 (27.90)
Is Intersection Location on a Frontage Road (base is "none of the approaches is a frontage road")		
At least one approach roadway is a frontage road	157.94 (65.30)	154.71 (63.78)
Traffic Volume-Related Variables		
Logarithm of daily entering volume	5.81 (0.64)	4.07 (0.83)
Flow split imbalance (FSIMB) factor	-8.68 (2.02)	-6.00 (2.12)