ABSTRACT

Title: PEDESTRIAN - VEHICULAR CRASHES: THE INFLUENCE OF PERSONAL AND ENVIRONMENTAL FACTORS

Carolina V. Burnier, M.Sc., 2005

Directed By: Dr. Kelly J. Clifton, Department of Civil and Environmental Engineering

This thesis examines the relationship between land use and pedestrian-vehicle crashes. Analysis focuses on how the pedestrian crashes vary by personal characteristics (age, sex, condition, and severity of injury) and physical characteristics of the crash area (location type, population density, land use, pedestrian activity, and demand). The data for this study are pedestrian-vehicle crashes in Baltimore City, MD occurring between 2000 and 2002, supplied by the State of Maryland Motor Vehicle Accident Report. The results from the analysis suggest that in general, there are significant effects of land use on pedestrian crashes and, more importantly, pedestrian exposure. When controlling for demand, urban downtown areas with high population and roadway densities and good commercial accessibility are found to have negative relationships to pedestrian exposure. The results may justify the promoting of denser urban neighborhood designs and advocate the need to guide safety policy investments to these urban areas with high pedestrian activity.

PEDESTRIAN - VEHICULAR CRASHES: THE INFLUENCE OF PERSONAL AND ENVIRONMENTAL FACTORS

By

Carolina Valdemarin Burnier

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2005

Advisory Committee:
Assistant Professor Kelly J. Clifton, Chair
Professor Paul Schonfeld
Associate Professor David J. Lovell

Dedication

To my husband, John, for his constant support throughout this process. I am eternally grateful for his kindness, patience and unconditional love. Thank you for always believing in me.

To my parents who have always been great examples of hard work and dedication. They have been the inspiration for my academic career. Thank you for teaching me to not only pursue knowledge but to love its pursuit.

To my brothers and sister, Eugenio, Miguel, and Julia for encouraging me in all my undertakings. Thank you for being my best friends all along.

Acknowledgements

This thesis would not be complete without the mention of my sincere appreciation to Dr. Kelly Clifton for her guidance and support.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Concept

The National Highway Traffic Safety Administration (2001) reports that although pedestrians are involved in a small proportion of vehicle crashes (2-3%), they represent a much higher proportion of crash fatalities (11-13%). In a majority of these deaths (55%), improper pedestrian behavior was a contributing factor. Several factors may contribute to these crashes including pedestrian characteristics, characteristics of the location of the crash and urban form.

Age, gender, socioeconomic status and pedestrian behavior are among the pedestrian characteristics that seem to have an effect on pedestrian-vehicular crashes. Children tend to be less educated about pedestrian safety and commit critical errors such as running across the street to reach a park or other destination. Gender is also an important factor in analyzing pedestrian-vehicular crashes. Studies show that men tend to engage in more risk-taking behavior than women and therefore may be more susceptible to being involved in a crash. Other socioeconomic characteristics (race, income, education) must also be included in the crash study. A greater percentage of the lower income population tends to rely on walking as a primary mode of transportation when compared to populations of higher income levels.

It is also important to consider pedestrian behavior when studying pedestrian crashes. Pedestrian obedience to traffic signals and substance use may contribute to the incidence of pedestrians being involved in crashes.

It appears that the built environment could also have significant influence on pedestrian crash rates. Pedestrian generators, such as high-density development and commercial land uses, would expect to see more crashes because they have more pedestrians.

This thesis examines the factors that contribute to pedestrian-vehicular crashes. Figure 1.1 shows a concept diagram of the variables being studied to understand the involvement of personal, location and urban form characteristics on crashes.

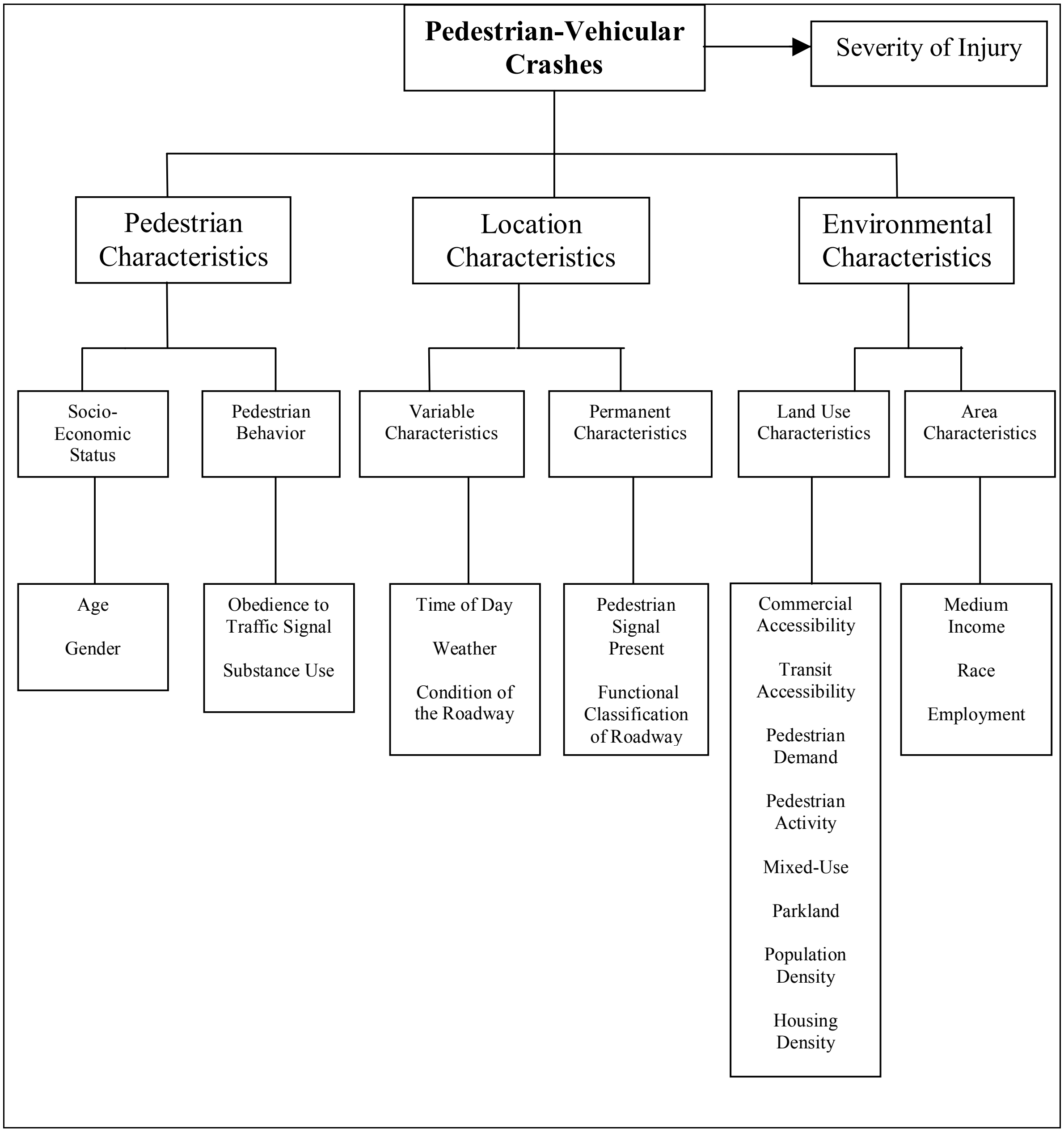


Figure 1.1 Concept Diagram for Crashes

1.2 Background on Baltimore City

The City of Baltimore is the largest city in the State of Maryland and has a population of over 650 thousand residents with a population density of over 8,000 persons per square mile. The median household income for Baltimore City, $30,078, is lower than the national median, at $41,994 and significantly lower than the Maryland state median, $52, 868. Because Baltimore is an urban city, it is not surprising that twenty-three percent of Baltimore City residents are below the poverty level, which is significantly higher than the national average of 12 percent. Over six percent of Baltimore City's employees walk to work, more than double the national average. Additionally, Baltimore City provides numerous public transit options/facilities: bus routes, light rail, a subway system and a commuter rail connecting to Washington, DC.

1.3 Purpose

This paper examines effects on pedestrian-vehicle crashes for the City of Baltimore, MD. The objective of the analysis is to determine the relationships, if any, between characteristics of pedestrians and the physical characteristics of crash locations. Specifically, this study aims to examine age and gender differences in pedestrian crashes with respect to: a) involvement; b) severity of injury; c) personal factors and behaviors; d) environmental conditions and e) demand. To do this, we examine pedestrian crash data for the City of Baltimore, using descriptive statistics and multivariate analysis. The next chapter details the data source and methodology in more detail.

The analysis hopes to obtain clear relationships between pedestrian-vehicle crashes and land use characteristics. These findings may aid in the planning of environments that are conducive to safe pedestrian activity. The results of this study should help guide safety policy decisions to ensure the well being of pedestrians.

Chapter 2: Literature Review

2.1 Pedestrian Behavior and Personal Characteristics

Research shows that personal characteristics have an important impact on mode choice. Hanson (1982) determines that socio-economic variables such as age, income and gender have the largest impact on mode choice. Dieleman et al. (2002) examines the effect of personal attributes on travel behavior with the Netherlands National Travel Survey. Their results show that personal characteristics such as higher income and presence of children in the household tend to increase trips by car and therefore decrease walking. Jang (2003) discusses models to forecast trips as a result of trip chaining. Jang finds that personal characteristics such as gender, education, marital status and presence of children have an effect on the allocation of travel modes.

Studies of personal characteristics and mode choice have focused on the travel differences between men and women; however, the majority of this work has focused on motorized travel. Rosenbloom (1998) found that women are more dependent upon transit, tend to make more household maintenance trips, and engage in more trip chaining. Pedestrian trips are often linked to transit because transit users frequently walk to access their transit stops. Therefore it is likely that more women walk then men.

2.2 Mode Choice and the Built Environment

There is extensive research work on the relationship between land use characteristics and mode choice. Rajamani et al. (2003) models land use measures and their effect on mode share. The land use measures include park area per housing unit, land use mix index, percent households within walking distance to a bus stop, population density, and percentage cul-de-sac. The study found that mixed land use and high residential densities increase walking and transit trips for non-work trips and an increased cul-de-sac percentage will decrease walking trips. Handy (1996) also explored the relationship between mode choice and land use characteristics. Handy found that the quality of the pedestrian environment at a destination played an important role in the choice to walk.

Supporting these findings, Cervero (2002) studied the effects of the built environment (density, diversity and design) on mode choice in Montgomery County, MD. Cervero found that higher densities and land use mixtures had a positive relationship with riding transit versus driving. However, Cervero warns that studies on urban form and travel mode choice often use models that are flawed due to mis-specification or omitted variables.

2.3 Crashes and the Personal Characteristics

Another thread of literature important to the topic of pedestrian activity is the interaction between personal characteristics and pedestrian crashes. Research work on the relationship between pedestrian-vehicular crashes and pedestrian behavior suggests that a large number of all crashes may be attributed to improper pedestrian behavior. Preusser et al (2002) found that in the Washington, D.C. and Baltimore area 50% of pedestrians involved in pedestrian vehicular crashes are judged culpable in the crash. Some pedestrian crashes can be associated with risk-taking behavior such as alcohol consumption, mid-block crossing, and crossing under unsafe conditions such as low visibility or high vehicular volume. Alcohol and substance abuse by the pedestrian are involved in fifteen percent of pedestrian crashes (Stutts et al. 1996).

In 2002, 4,808 pedestrians were killed in the United States in traffic crashes. However, data from the National Highway Traffic Safety Administration (2004) reveal that although traffic fatalities have increased over the last decade, pedestrian fatalities have steadily decreased (13% since 1992). Alcohol involvement was reported in 46% of the crashes resulting in pedestrian fatalities for either the motorist or pedestrian (National Highway Traffic Safety Administration, 2004). Age and gender seem to have an effect on the severity of injury in pedestrian-vehicular crashes. In 2000, 31% of pedestrians involved in crashes that were injured were children (less than 16 years of age) and 11% of pedestrian fatalities involved children (National Highway Safety Administration, 2001). In 2002, men were involved in greater numbers of fatal crashers per population than women in every age category (National Highway Safety Administration, 2004).

Research on the behavior of children shows that children tend to be the most likely group to make a critical error in pedestrian behavior such as mid-block dashes (Preusser et al., 2002). Much research has also been dedicated to the differences between men and women in risk-taking behaviors. This may help explain the differences in crash involvement by gender. Men may be placing themselves at greater risks by crossing an intersection improperly, disobeying a pedestrian signal, or crossing at midblock. Alcohol and substance abuse by the pedestrian are involved in fifteen percent of pedestrian crashes (Stutts et al., 1996) and men are more likely than women to be drinking heavily and using illicit drugs (Thom, 2003).

2.4 Crashes and the Built Environment

Less research has focused on environmental factors, specifically land use factors, associated with pedestrian crashes. Because over 70% of all pedestrian crashes occur in urban areas (Herbert-Martinez et al., 2004 and National Highway Safety Administration, 2004), it is important to understand the interaction between pedestrian crashes and the built environment.

Retting et al (2003) provides an assessment of the changes to the built environment that can impact the occurrence of pedestrian crash injury. These measures include speed control, separation of pedestrians from roadway and increased visibility of pedestrians. Their results show that physical changes to the built environment can significantly decrease pedestrian-vehicle crashes.

Intuitively, it appears that the built environment could have significant influence on pedestrian crash rates. Pedestrian generators, such as high-density development and commercial land uses, would expect to see more crashes because they have more pedestrian demand.

2.5 Conclusions

As evidenced by previous research, differences in travel behavior, including mode choice, trip purpose, and time of day, may lead to distinct patterns in walking behavior. In turn, these travel patterns, along with other factors may result in possible crash involvement. This, in turn, raises questions about the relationship of pedestrian crashes with different urban forms.

This research will focus on the relationship of pedestrian crashes and the physical environment characteristics with an emphasis on land use attributes of the crash location for the City of Baltimore. The study will also include population characteristics such as population density, sociodemographics, age and gender of pedestrians involved in crashes and land use characteristics such as proximity to transit, commercial district or parkland and housing density. This study will also emphasize differences in pedestrian-vehicular crashes by age group and gender.

Chapter 3: Methodology

3.1 Data

The Maryland State Highway Administration provided the data source for this study. The data consist of pedestrian-vehicle crashes in Baltimore City, Maryland obtained from the State of Maryland Motor Vehicle Accident Report. These crashes occurred over a three year period (2000-2002). The data consist of individual crash records with personal characteristics of the victim and location characteristics of the crash associated with each crash record. The personal characteristics include information on the victim's age, sex, clothing, obedience to pedestrian signals and substance use. The location characteristics for the crash include the type of intersection, whether pedestrian signalization is present, surface condition and location of victim at the time of crash. Information regarding other crash characteristics such as time of day and weather are also incorporated on the database.

Using a Geographic Information System (GIS), the individual crash records were given specific geographic coordinates of the nearest intersection in order to locate the crashes spatially on a map of Baltimore City. Through GIS, each crash was identified within a Census blockgroup using shapefiles provided by the Census website. A Census blockgroup is composed of Census blocks, which are the smallest geographic area for which the Census Bureau collects and tabulates decennial census data. Census blockgroups have similar populations (approximately 1000 people) and have a minimum area of 0.69 acres (Bureau of the Census, 1995). Boundary files for the block groups were obtained from the TIGER files from U.S. Census 2000.

Using the 2000 U.S. Census demographic information tables, additional information about the blockgroup area where each crash occurred was added to the crash database. This information includes median income, total population, population density, number of households, employment, and percent population that walks to work.

Six land use variables were aggregated to the blockgroup level: single family residential dwelling units (SFRDU) density, percent SFRDU within bus stop buffer, percent SFRDU within commercial uses buffer, percent land dedicated to parkland, population density, roadway density and mixed use. Point data, including SFRDU and commercial parcel data were obtained from Maryland Property View provided by the Maryland Department of Planning. The Baltimore City government, through a cooperative agreement, supplied the park polygon file where percent park land for each block group was obtained. In addition, bus stop access information for each block group was gathered from Transit View 2000 supplied by the Maryland Department of Transportation. From the Enhanced Tiger Files, the functional classification of the roadways where also obtained to verify whether different types of roadways have a higher occurrence of pedestrian-vehicular crashes.

The data were then analyzed using STATA and SPSS to obtain descriptive and multivariate statistical analysis to discern the differences in crashes and their interaction with land use variables.

3.2 Methodology

GIS was used to show the spatial distribution of the pedestrian-vehicular crashes as shown on Figure 3.1. The crash data contained 3408 crashes, however only 3009 crashes had sufficient information to be geocoded. Therefore the analysis is performed on 3009 pedestrian-vehicular crashes. Land use variables and census information about the blockgroup area where the crash occurred were added spatially to the each crash point using GIS. Two databases resulted from the spatial analysis. The first database lists every crash with its respective attributes including personal and environmental characteristics, land use data and crash location information. The second database is structured by blockgroups and includes aggregate information of the crashes and the land use characteristics to each particular blockgroup of Baltimore City. Baltimore City is divided into 710 blockgroup areas that can be seen in Figure 3.2.

Descriptive statistics and multivariate analysis were performed to determine relationships between pedestrian-vehicle crashes and land use characteristics. The variables in the study included continuous variables such as age, time of day, population density and transit accessibility. Others were categorical, such as severity of injury, pedestrian signal obedience and presence of substance. To better evaluate these pedestrian and environmental characteristics, new categories have been designated for each variable. For most cases three categories are chosen for each variable.

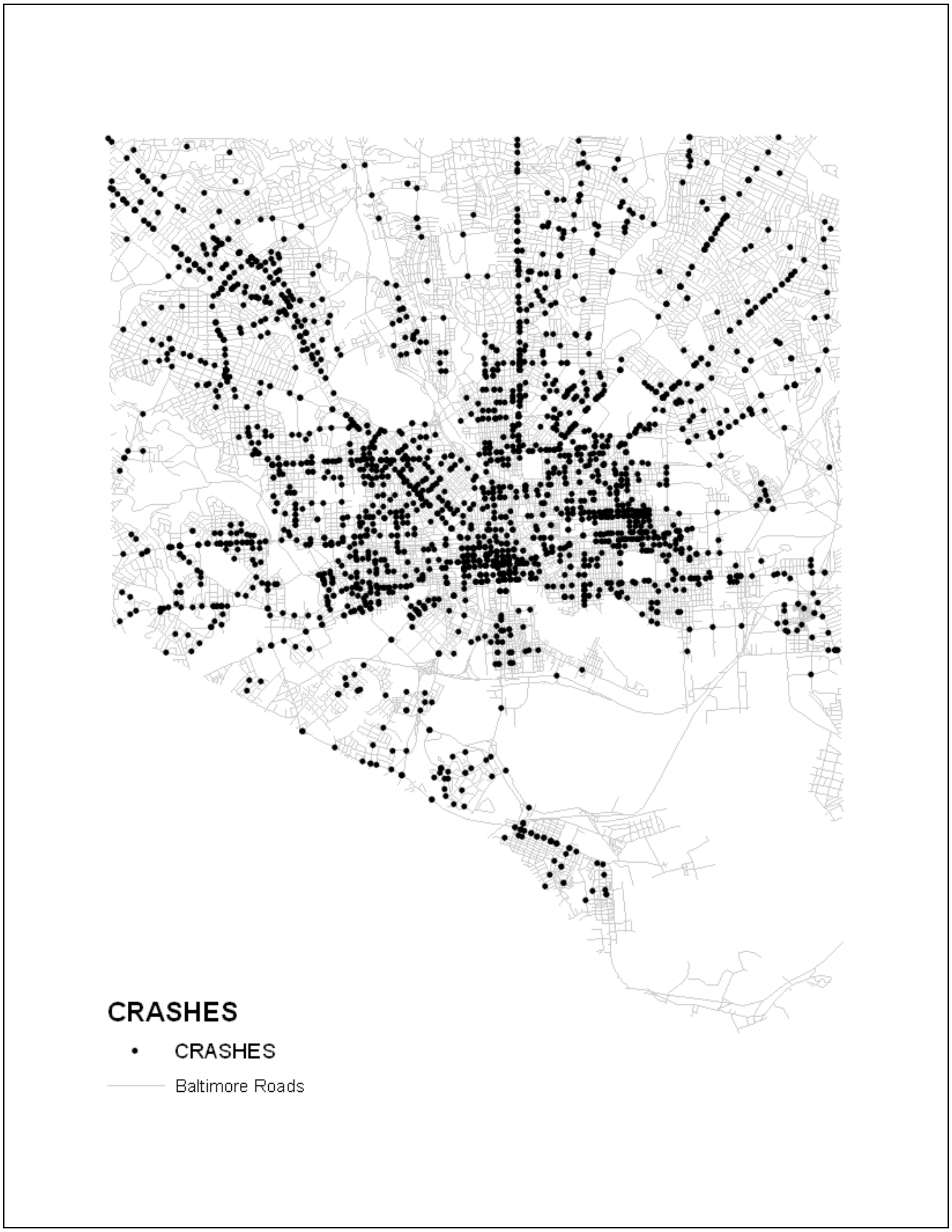


Figure 3.1: Distribution of Pedestrian-Vehicular Crashes

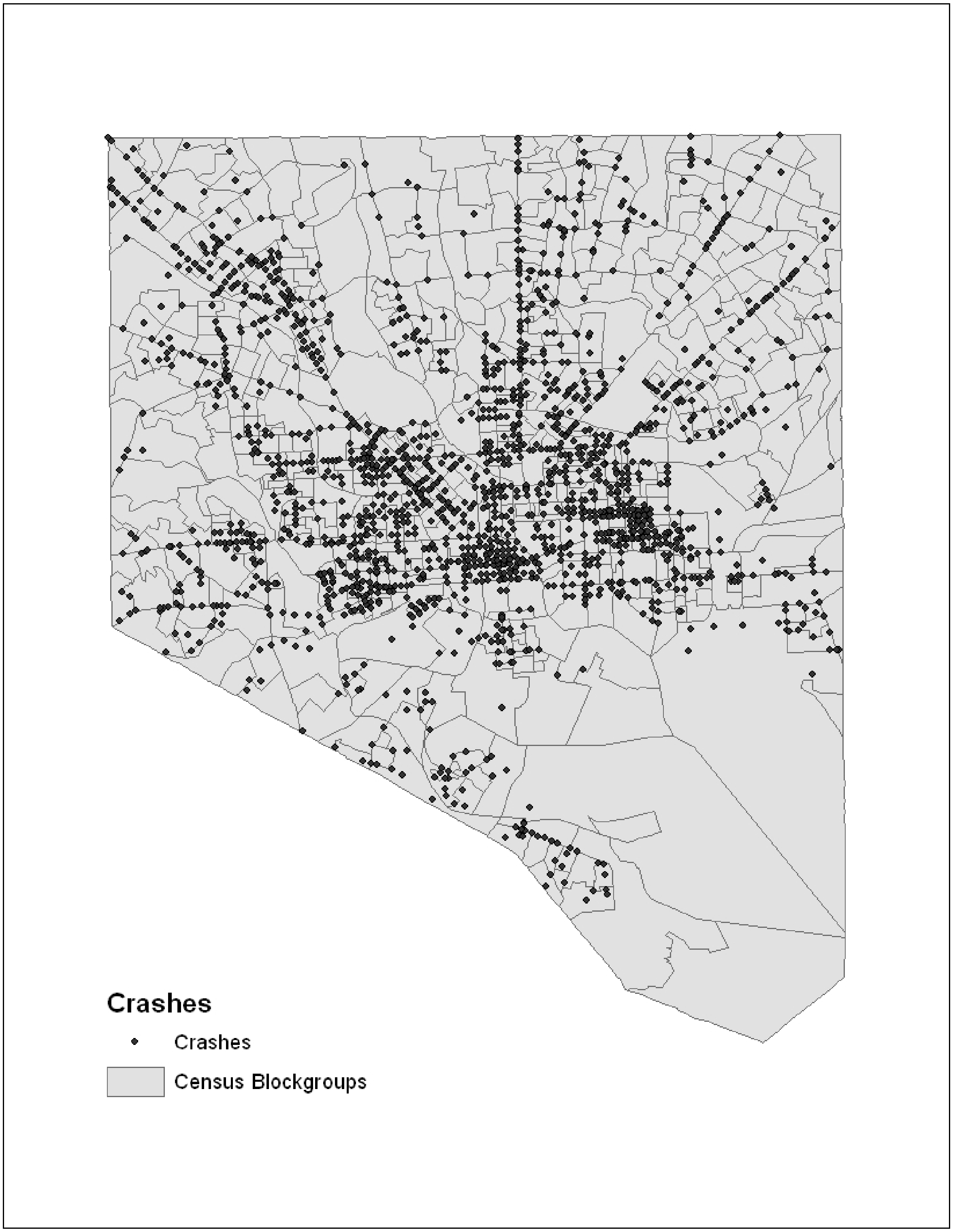


Figure 3.2: Pedestrian-Vehicular Crashes by Blockgroup

3.2.1 Pedestrian Characteristics

The analysis for the study investigates the relationship between crashes and land use characteristics by victims of different age groups and gender. The age groups in this study are defined as child (0 to 15 years of age), adult (16 to 64 years of age), and elderly (65 years of age or over).

The severity of injury for each crash was reported as one of five categories: not injured or unknown, possible injury, not incapacitated injury, disabled injury and fatal injury. For the purpose of this study, three severity of injury categories have been designated. The category for no injury comprises of crashes with no or unknown injury and possible injury. Not incapacitated and disabled injury categories have been combined into the injury category and fatal injuries remains its own category.

The crash data also included information on whether a substance such as alcohol and illegal or medicinal drug was detected on the pedestrian involved in the crash. Three categories are used in this study to evaluate the effect of substance use on the crashes: no substance detected, substance present, and substance contributed.

Data on pedestrian traffic law obedience are also provided in the crash reports. There were four categories provided: pedestrian obeyed signal, pedestrian disobeyed signal, no pedestrian signal or malfunctioning pedestrian signal. The no pedestrian signal category incorporates crashes that occurred at intersections with no pedestrian signal or where pedestrians were crossing mid-block. The data set contained pedestrian traffic signal obedience for over 2500 crashes but only 6 crashes occurred in a location with a malfunctioning pedestrian signal, therefore malfunctioning and no signal categories are combined into one for the purpose of this paper.

3.2.2 Characteristics of the Crash Location

The location of the pedestrian at the time of the crash is also included in the analysis. The police reports identify where the pedestrian was located such as on the roadway, crosswalk, shoulder, or sidewalk. The three categories used in the study are on roadway at crosswalk, on roadway not at crosswalk and other locations. The other locations category includes shoulder, curb, sidewalk, bikeway, school bus zone and outside right of way.

The crash data set include the time of the crash. However, because of the disparity in daylight hours among different days of the year, the variable for light is used instead to account for time of day. The variable for light is categorized as daylight, dawn or dusk, and dark.

Roadway condition is also recorded at the time of crash in various categories, such as shoulder defect, hole or ruts, loose surface material or obstruction; however this analysis only considers whether there was a defect on the roadway at the time of crash or not.

Information on roadway classification is included in the street network files provided by the US Census Bureau. There are many roadway classes used but for the purposes of this study, classification of the roadways is grouped into four categories. These categories are primary separated, primary unseparated, local and neighborhood separated, and local and neighborhood unseparated. The Primary roadway categories include all primary interstate highways, primary state highways, and secondary state or county highways. The local streets categories include local, neighborhood and city streets. To assign the functional classification of the roadways to the crash location, the two databases were combined by the primary street where the crash occurred. For the great majority of roadways, the functional classification used in this paper is unchanged for the entire length of the road. However for those few roadways which have segments classified as primary and local, the primary classification was chosen.

3.2.3 Land Use Characteristics

To observe the effect of land use characteristics on pedestrian-vehicular crashes, land use variables are constructed using blockgroup level data from the U.S. 2000 Census. The built environment characteristics that are calculated in this study are as follows: single family residential dwelling units (SFRDU) density, percent SFRDU within bus stop buffer, percent SFRDU within commercial uses buffer, percent mixed-use, percent population that walk to work, percent land dedicated to parkland, population density, and roadway density.

Journey to work data at the blockgroup level from the U.S. Census 2000 are used as one indicator of pedestrian activity. The GIS quantile feature is used to find the range of percentage of the population that walk to work for each category so that each class contains the same number of members. The areas are categorized into low pedestrian activity (less than 8% of workers walk), medium (8-23% walk) and high (23% and greater).

For the percent park area measure, GIS is used to superimpose Baltimore City's park layer onto the blockgroup layer. The area designated to parkland of each blockgroup is then obtained by performing an intersect operation in GIS of the parks and blockgroups. The area used for park purposes is then divided by the total area of each blockgroup to obtain the percentage park area measure. Similarly to the pedestrian activity measure, the quantile feature is used so that blockgroups are categorized by the amount of park land with low (less than 10% of tract area), medium (10 – 40% of tract area), and high (greater than 40%).

The population density of each blockgroup area is calculated by dividing the population (number of residents in the area) by area of each blockgroup obtained from the U.S. Census 2000. Again, the quantile feature in GIS is used to categorize blockgroups by the levels of population density with low (less than 10,000 person per square mile), medium (10,000 to 20,000 persons per square mile) and high density areas (greater than 20,000 persons per square mile).

Information on the Median Household Income of each area is provided by the US Census 2000. The three categories are obtained so that each class contains the same number of features. The categories are low (less than $25,000), medium ($25,000 to $40,000) and high (more than $40,000).

Using GIS, a measure of pedestrian accessibility to commercial uses (retail and services) is calculated. This measure is computed as the percentage of single-family residential units (SFRDU) in the area within ¼ mile of commercial uses. To perform this measure, Property View data are used to obtain the GIS layers for commercial parcels and for single family residential units. The superimposed layers are then used to create a ¼ mile buffer zone for commercial parcels. The number of single family residential units in each buffer zone is calculated. The number of SFRDUs for each blockgroup that fall within commercial buffer zones are summed and then divided by the total number of SFRDUs in that particular blockgroup. By dividing the blockgroups in three quantiles, this measure is then categorized as high (over 85% of households that live within ¼ of commercial), medium (55-85%), and low (less than 55%).

The measure used in this study for transit accessibility is percent single family residential units within ¼ mile from bus stops. To perform this measure, Property View data are used to obtain the GIS layers for single family residential units and Transit View is used to obtain the bus stops layer. The superimposed layers are then used to create a ¼ mile buffer zone for bus stops. Similarly, the number of single family residential units in each buffered zone of commercial uses is calculated. The number of SFRDUs for each blockgroup that fall within transit buffer zones are summed and then divided by the total number of SFRDUs in that particular blockgroup. This measure is also categorized by divining the blockgroups into quantiles. The lowest percentage of transit accessibility in Baltimore City was 40%. Therefore the categories are as follows: high (over 95% to 100% of households that live within ¼ of transit), medium (80% to 95%), and low (40% to 80%).

Mixed use variables are also important when analyzing land use characteristics. For this study, the mixed use variable is defined as the ratio of commercial area to total area for each blockgroup. The total commercial square mileage for each blockgroup is computed by summing the first-floor area for each business in that blockgroup. The caveat with this measure however is that if the business is located on multiple floors, then only the first floor is considered. However, if there are two businesses located on the same building but on different floors, then the area of each commercial entity is included in the sum. Thus, it is possible that very few of the blockgroups show ratios that exceed 1. The categories for this variable were also obtain using quantiles and are defined as high (ratio greater than 0.45 of commercial area divided by total area), medium (0.017 to 0.450), and low (ratio less than 0.017).

The density of roads is also included as a land use measure in the analyses. This measure is created in GIS by summing the road segments in each blockgroup to obtain a total linear distance of roads per blockgroup. This distance is then divided by the total area of each blockgroup to generate a density of roads. Again this measure is divided into three categories, each with the same number of members. The classes produced for this measure are high (greater than 45 linear miles per square mile), medium (23 to 45), and low (less than 23).

3.2.4 Descriptive Statistics

The pedestrian and crash characteristics and land use measures described above are used to generate descriptive statistics to identify their relationships with the incidence of pedestrian-vehicular crashes. Each variable is analyzed by age and gender and where applicable by severity of injury. The majority of the variables are classified into three categories (low, medium and high). The analysis is done in SPSS and results are summarized in tables in the discussion section. The χ^2 value of each relationship is computed to show statistical significance. The tables show the degree of statistical significance of each relationship. However, unless otherwise stated, statistical significance discussed in the text is at the 95% level of confidence or better.

3.2.5 Multivariate Statistics

In order to better understand the effects of the pedestrian, location and environmental characteristics, multivariate analysis of the crashes is performed. A series of linear regressions are used to determine a model that best explains the crashes variable. The analysis is performed at the blockgroup level. For this analysis, a number of additional variables for each blockgroup were computed in order to better control the crashes variable. These control variables are as follows: total number of households, employment (percent residents over 16 years of age who are employed), race (percent non-white population), children (percent people under 16 years of age), education (percent population that attended at least some college) and vehicle ownership (percent households that have at least one vehicle).

The linear regressions are performed for the number of crashes in each area (blockgroup). However, because blockgroups have varied areas, the dependent variable used is density of crashes per blockgroup. Density of crashes is defined by equation 1:

$$crash\ density = \frac{\#crashes_{blockgroup}}{area_{blockgroup}} \qquad (1)$$

To account for the differing populations in each area, models are also estimated with the dependent variable as the density of crashes per population of the blockgroup as shown on equation 2:

$$y_i = \frac{crash\ density_{blockgroup}}{population_{blockgroup}} \qquad (2)$$

The histogram for the dependent variables used such as crash density and crash density per population (y_i) are plotted to ensure that the variables follow a normal distribution. It can be seen from the histograms that the variables are skewed to the left. As an example figure 3.3 shows the histogram for crash density. Therefore, the natural log of the variable is computed to normalize the distribution as shown in figure 3.4. The natural log of the crash variable is therefore used as the dependent variable for the models.

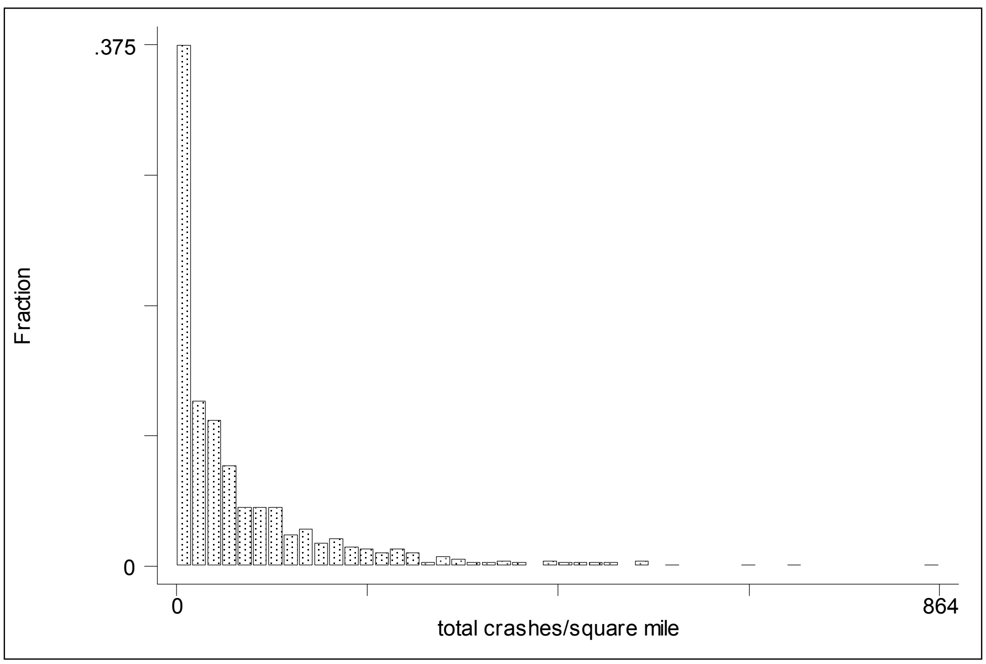


Figure 3.3: Histogram for Crashes per Square Mile Variable

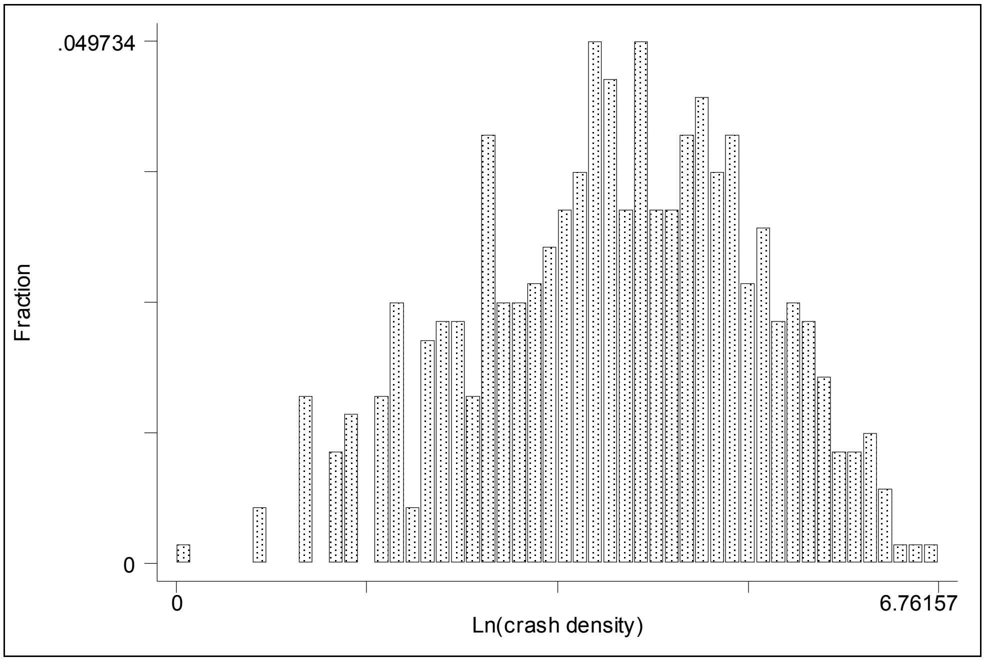


Figure 3.4: Histogram for Ln(crashes per square mile) Variable

The regression analysis is performed as a function of location characteristics and the land use policy variables. The location characteristics include total number of households, total employment, race, median income, education and percentage of vehicle ownership. The variables of interest that can have an impact of future policy for land use design are pedestrian activity, commercial accessibility, transit accessibility and roadway density.

STATA is used to run the linear regression models. Based on the descriptive statistics for the variables available and models used in previous literature, the variables for the final model was chosen as shown on equation 3:

$$ln(y_i)=\beta_o+ \beta_1(number\ of\ households)_i + \beta_2(employment)_i + \beta_3(pedestrian\ activity)_i+\beta_4(income)_i + \beta_5(parkland)_i +\beta_6(commercial\ accessibility)_i +\beta_7(transit\ accessibility)_i+\beta_8(population\ density)_i + \beta_9(children)_i + \beta_{10}(race)_i+ \beta_{11}(education)_i + \beta_{12}(vehicle\ ownership)_i + \beta_{13}(roadway\ density)_i +\beta_{14}(mixed\text{-}use)_i +\xi_i \qquad (3)$$

Correlation tables were computed for all variables in the model to screen for multicollinearity. If two parameters in a regression are highly correlated, it is said there is multicollinearity present in model and therefore the standard errors are overestimated. There is high correlation between two variables if the correlation coefficient, ρ, is greater than 0.8. Correlation coefficients for each pair of regressors are obtained.

In addition, these models are corrected for heteroskedasticity. If the models show heteroskedasticity, then it violates distributional assumptions so that the t-statistics and the standard errors may be incorrect. A test is performed for each regression model; if heteroskedasticity is found to be present, then the robust standard errors are reported in the results.

3.2.6 Demand Analysis

The crash analysis thus far is performed solely on the crash variables and environmental characteristics for the location of the pedestrian-vehicular crash without true consideration for pedestrian demand. Confounded with the crashes is the number of pedestrians that each of the intersections generate. Pedestrian collisions are likely correlated with environments that are conducive to walking. For better results, the analysis of crashes should include controls for pedestrian demand. There are few comprehensive databases available on pedestrian counts, especially over large geographical areas. One source of pedestrian exposure data is the Clifton et al (2005) report on pedestrian flow modeling. These data are estimated 24 hour pedestrian counts at every intersection in an area of approximately ten square miles in Central Baltimore City. The area included in this model can be seen in Figure 3.5. These data are used in this analysis of crashes to control for pedestrian exposure. Pedestrian exposure is defined as the number of crashes per number of pedestrians.

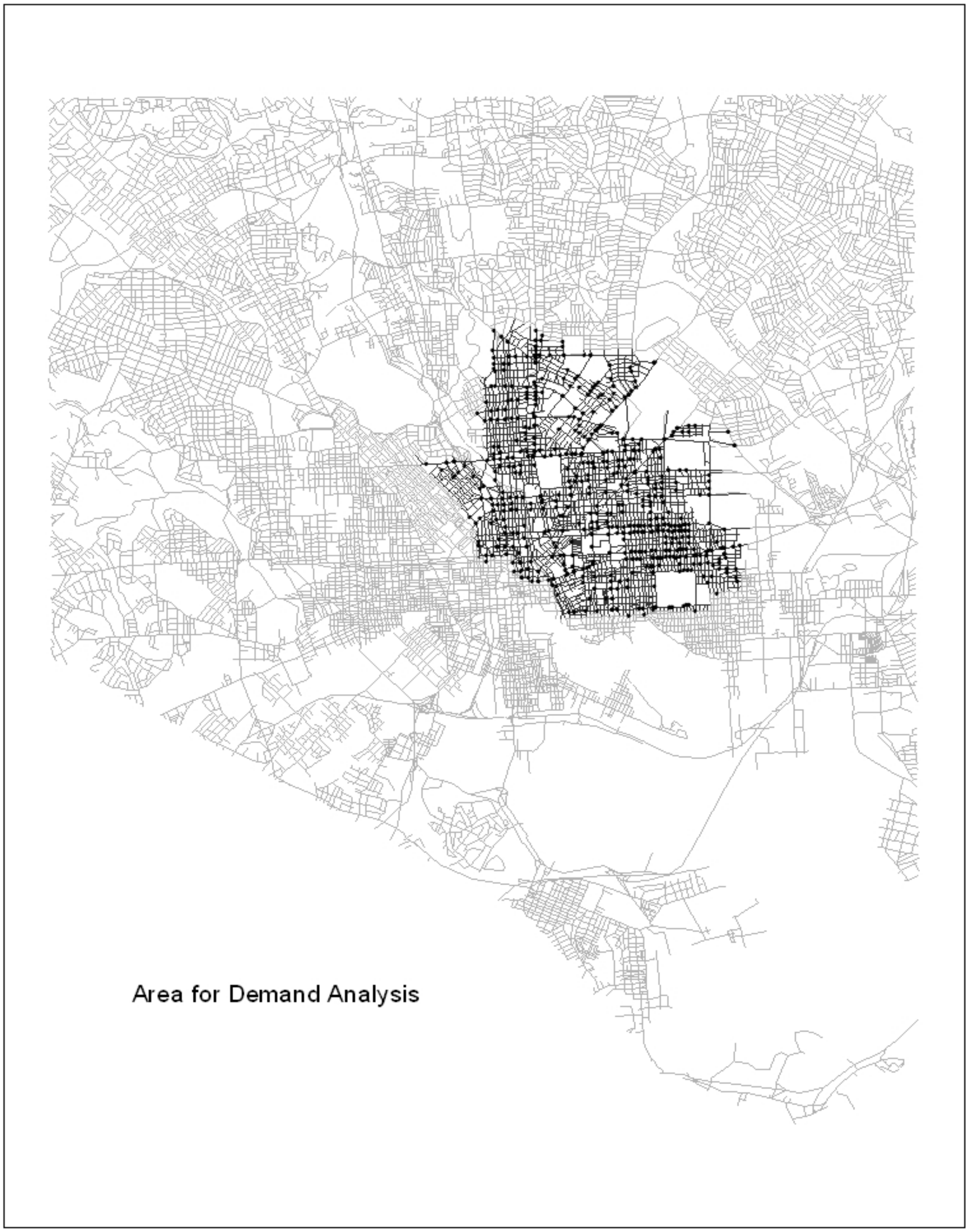


Figure 3.5: Area for estimated 24-hour pedestrian counts

The linear regression models described in Section 3.2.5 are recomputed with a control for pedestrian demand. The new models are restricted to the crashes occurring in the area described in figure 3.5. The analysis is performed at the intersection level. The dataset used for analysis is constructed using GIS and intersecting the crashes with the pedestrian demand layer. The dataset lists the intersections within the ten square mile area with attributes of the corresponding blockgroup. The dataset also includes an aggregate number of crashes for each intersection. The independent variables in the model remain unchanged from equation 3. However, the dependent variable used for exposure is the natural log of crashes per pedestrian count for each intersection. The dependent variable is then multiplied by 1000 for ease of interpretation. The final dependant variable used in the model is shown with equation 4.

$$Exposure = 1000 * \ln\left(\frac{\# crashes_{\text{intersection}}}{demand_{\text{intersection}}}\right) \tag{4}$$

Therefore, the final linear regression analysis model used in shown with equation 5.

$$\begin{aligned} Exposure = {} & \beta_o + \beta_1(\text{number of households})_i + \beta_2(\text{employment})_i + \beta_3(\text{pedestrian} \\ & \text{activity})_i + \beta_4(\text{income})_i + \beta_5(\text{parkland})_i + \beta_6(\text{commercial accessibility})_i \\ & + \beta_7(\text{transit accessibility})_i + \beta_8(\text{population density})_i + \beta_9(\text{children})_i \\ & + \beta_{10}(\text{race})_i + \beta_{11}(\text{education})_i + \beta_{12}(\text{vehicle ownership})_i + \beta_{13}(\text{roadway} \\ & \text{density})_i + \beta_{14}(\text{mixed-use})_i + \xi_i \end{aligned} \tag{5}$$

Chapter 4: Results

4.1 General Statistics on Crashes

The data are analyzed based on age, sex, severity of injury, presence of alcohol or other substance, pedestrian obedience to traffic signals, time of day, weather at time of crash, and land use variables. There are 3009 pedestrian-vehicular crashes in Baltimore City for a period of 3 years analyzed in this study. Prior to evaluating the impact of pedestrian or land use characteristics on the crashes, the basic conditions of the roadway and weather must be considered. Ninety-seven percent of these crashes occurred in roadways with no defect and 87% of the crashes occurred in clear or cloudy weather. In addition over 80% of the crashes occurred in roadways where the surface condition was dry at the time of the crash. These statistics imply that the majority of crashes are not a result of poor road conditions or bad weather and therefore may be attributed to other external factors.

Table 4.1 shows general descriptive statistics for the population of crashes being analyzed in this study. Out of the 3009 crashes from 2000-2002, 40% of the crashes involved female pedestrians and 60% involved male pedestrians. The distribution of the pedestrian-vehicular crash data among different age groups is 36% children, 60% adults and 4% elderly. Over 50% of these crashes resulted in injuries or fatalities. Almost 70% of all crashes occurred during daylight hours. Over 10% of crash victims were found to have a substance (alcohol or drugs) present or contributing to the crash. It is interesting to note that 19% of the victims were at crosswalks and nearly 55% of these crashes occurred with pedestrians that were on the roadway but not at a crosswalk.

Table 4.1: Summary of Crash Dataset

Characteristics of pedestrian	Total %	Characteristics of crash	Total %
Number of total crashes	3009	**Location of Pedestrian**	
Gender		On Roadway at Crosswalk	19.1%
Male	59.6%	On Roadway not at Crosswalk	54.5%
Female	40.4%	Other Locations	26.4%
Age		**Time of Day**	
Child	35.8%	Daylight	69.2%
Adult	59.9%	Dawn or Dusk	4.8%
Elderly	4.2%	Dark	26.0%
Substance Use		**Severity of injury**	
None	90.7%	No injury	49.4%
Substance Present	7.0%	Non-fatal injury	49.2%
Substance Contributed	2.3%	Fatality	1.5%

4.2 Pedestrian and Crash Characteristics

To understand the impact of land use characteristics on pedestrian-vehicular crashes, it is important to first evaluate the influence of the characteristics of the pedestrians involved in the crash and of the crashes themselves. For the purposes of this study, crashes are evaluated by two main categories: age and gender. Table 4.2 and 4.3 show the pedestrian characteristics variables by age and gender respectively.

Table 4.2: Pedestrian Crashes by Pedestrian Characteristics for Age Groups

Characteristics of pedestrian or crash	Total N	Total %	Child	Adult	Elderly	
Number of crashes	3009	100.0%				
Traffic Law Obedience						***
No or malfunctioning signal	1912	78.1%	84.4%	74.8%	67.3%	
Obeyed signal	225	9.2%	3.9%	11.7%	20.2%	
Disobeyed signal	311	12.7%	11.6%	13.4%	12.5%	
Total	2448	100.0%	100.0%	100.0%	100.0%	
Severity of injury						***
No injury	1376	48.8%	48.2%	49.7%	41.2%	
Non-fatal injury	1399	49.6%	50.8%	48.9%	50.4%	
Fatality	43	1.5%	1.0%	1.4%	8.4%	
Total	2818	100.0%	100.0%	100.0%	100.0%	
Substance Abuse						***
None	1151	90.9%	99.5%	86.2%	91.8%	
Substance present	89	7.0%	0.5%	10.4%	8.2%	
Substance contributed	26	2.1%	0.0%	3.3%	0.0%	
Total	1266	100.0%	100.0%	100.0%	100.0%	
Time of Day						***
Daylight	1948	69.4%	79.7%	62.5%	81.5%	
Dawn or Dusk	136	4.8%	5.6%	4.5%	5.5%	
Dark	721	25.7%	14.7%	33.0%	15.1%	
Total	2805	100.0%	100.0%	100.0%	100.0%	

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

Table 4.3: Pedestrian Crashes by Pedestrian Characteristics for Gender

Characteristics of pedestrian or crash	Total N	Total %	Male	Female	
Number of crashes	3009	100.0%	1778	1207	
Age					
Child (0 to 15 years of age)	1001	35.7%	37.1%	33.7%	
Adult (16 to 64 years of age)	1685	60.1%	58.8%	62.0%	
Elderly (over 65 years of age)	118	4.2%	4.1%	4.3%	
Total	2804	100.0%	100.0%	100.0%	
Traffic Law Obedience					***
No or malfunctioning signal	1971	78.1%	79.6%	75.8%	
Obeyed signal	229	9.1%	6.2%	13.3%	
Disobeyed signal	324	12.8%	14.2%	10.9%	
Total	2524	100.0%	100.0%	100.0%	
Severity of injury					*
No injury	1463	49.0%	47.3%	51.5%	
Non-fatal injury	1478	49.5%	51.1%	47.2%	
Fatality	44	1.5%	1.6%	1.2%	
Total	2985	100.0%	100.0%	100.0%	
Substance Abuse					***
None	1200	90.7%	87.4%	95.7%	
Substance present	93	7.0%	9.3%	3.6%	
Substance contributed	30	2.3%	3.3%	0.8%	
Total	1323	100.0%	100.0%	100.0%	
Time of Day					**
Daylight	2051	69.1%	67.4%	71.6%	
Dawn or Dusk	145	4.9%	5.5%	4.0%	
Dark	772	26.0%	27.1%	24.4%	
Total	2968	100.0%	100.0%	100.0%	

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

4.2.1 Age and Gender

Figure 4.1 shows the spatial distribution of the pedestrian-vehicle crashes differentiated by the age group of the pedestrian involved in the crash: child, adult, and elderly. The majority of crashes, nearly 60%, involved adults. Less than 5% of the crashes involved the elderly. However, over 35% of pedestrian crashes occurred with children. The 2000 U.S. Census shows that children comprise 22% of the population of Baltimore City, therefore there is a disproportionably higher percentage of children that are involved in crashes. Although Baltimore City has a high number of crashes involving children, it is interesting to see in figure 4.1 that the center cluster of crashes include very few children. This may be the result of the downtown area being less residential with fewer school routes and therefore having fewer pedestrians under the age of 16 years.

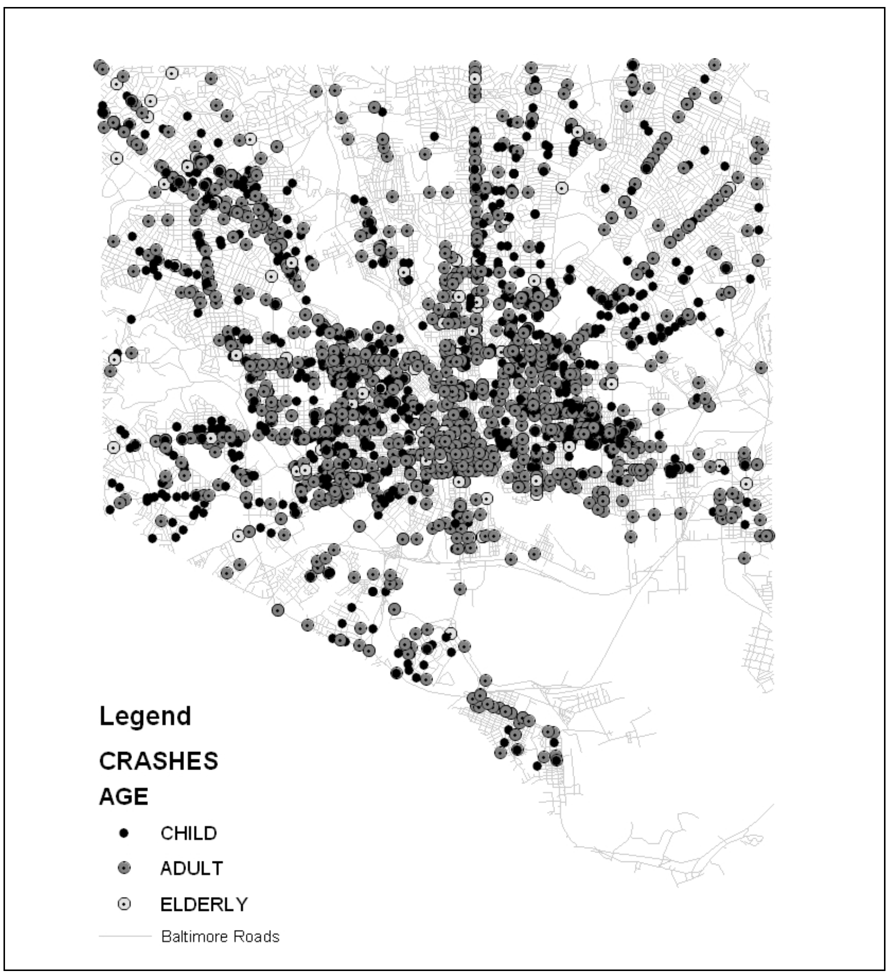


Figure 4.1: Pedestrian-Vehicular Crashes by Age Group

The U.S. 2000 Census shows that the City of Baltimore has a population distribution of 46.6% men and 53.4% women. However, nearly 60% of the pedestrian-vehicular crashes involved male pedestrians and approximately 40% involved female pedestrians. These statistics may indicate that male pedestrians may be involved in more risk taking behavior than women.

Although these results are not statistically significant, it is notable that there was a greater percentage of adult and elderly female pedestrians (66%) involved in crashes than adult and elderly male pedestrians (63%). However, the converse was true for children. Thirty-seven percent of females involved in crashes were children, whereas 34% of male victims were children.

An analysis of the crashes is performed by age and gender based on characteristics of the pedestrian and the crash such as pedestrian obedience to traffic signals, severity of injury, presence of alcohol or other substance and time of day.

4.2.2 Pedestrian Obedience

Seventy-eight percent of the pedestrian crashes were reported to have occurred at locations with malfunctioning or no signalization. In addition, almost 80% of fatalities occured in areas with no traffic signal. In the cases where a pedestrian signal was available, pedestrians who did not obey a traffic signal were slightly more likely to sustain injuries than those who crossed with a signal: 51.5% for the former and 56.5% for the latter, however these results are not statistically significant. However, the relationship between traffic law obedience with gender and age of pedestrian is statically significant. Of those pedestrians involved in a pedestrian-vehicle collision who were crossing where a pedestrian traffic signal was available, men were significantly more likely to disobey the pedestrian signal than women (70% of men, compared to 45% of women). These statistics are consistent with previous research that suggests that men seem to engage in more risk taking behavior than women.

With regard to pedestrian signal obedience, the behavior of pedestrians involved in crashes seems to be influenced by age. Among children involved in crashes, a much higher percentage of pedestrians disobeyed a signal (75%) than obeyed a pedestrian signal (25%). For the adult category, the difference between the pedestrians involved in collisions who obeyed traffic signals compared to the ones who disobeyed them is less apparent: 47% for the former and 53% for the latter. In contrast, among the elderly population more often than not pedestrians involved in crashes obeyed the traffic signal when available (62% obeyed pedestrian signal compared to 38% who did not). These results suggest that obedience to traffic law may increase with age.

4.2.3 Severity of Injury

Regarding severity of injury, approximately 50% of the crashes resulted in injuries and 1.5% of reported crashes resulted in pedestrian fatalities. The relationship between severity of injury and age is statistically significant as shown on table 4.2. Adults were slightly less likely to be injured in a crash, with nearly 50% sustaining no injuries. However, only 48% of children and 41% of elderly pedestrians involved in crashes sustained no injury. The percentage of fatalities in each age group increased with age. Only one percent of children involved in crashes sustained fatal injuries, whereas 1.4% of crashes involving adults resulted in fatalities. There is an even greater increase in the percentage of fatalities when looking at the elderly age group, 8.4% of the crashes involving the elderly were fatal to the pedestrian.

Regarding the relationship between gender and severity of injury, women were slightly less likely to be injured in a crash with 51% sustaining no injuries as compared to 47% of men. Consistent with previous research, men were more likely to be injured in a crash (51% compared to 47% of women) and slightly more likely to die as a result (1.6% compared to 1.2% of women). However, these gender differences in severity of injury are small and only significant for a 90% confidence interval.

4.2.4 Substance Use

For the pedestrians involved in the crash, the effect of the presence of alcohol, medication and illegal substances on crashes is evaluated. In over 90% of the pedestrians involved in crashes, no substance was detected. Seven percent of pedestrian crashes involved pedestrians that had medication, alcohol or illegal substances present in their system and 2% of the pedestrians had substances in their system that contributed to the crash. As expected, there was a significantly greater percentage of adults involved in crashes (14%) that tested positive for alcohol or drugs than elderly (8%) or children (<1%).

Regarding gender and substance use, the results are found to be statistically significant. Men involved in collisions proved more likely to test positive for some substance than women: 13% for the former and 4% for the latter. In over 3% of all male crashes, alcohol, medication or illegal substances contributed to the crash, compared to less than 1% of females. In addition, pedestrian men involved in collisions are more likely to have alcohol present at the scene of the accident and were more likely to have alcohol found to be a contributing factor to the crash.

Pedestrian substance use is also found to have a statistically significant effect on severity of injury. The data show that the severity of injury increased with presence of substance as shown in table 4.4. Only 2% of crashes involving pedestrians who were found to have no substance detected at the time of crash resulted in fatalities. In contrast, over 6.5% of crashes where a substance was present or contributing to the crash resulted in the pedestrian's death. In addition over 70% of pedestrian crashes where a substance contributed to the crash resulted to injuries or fatalities.

Table 4.4: Severity of Injury by Substance Use

Substance Use	No Injury	Non-fatal Injury	Fatality	Total
None	46.5%	51.5%	2.0%	100%
Substance Present or Contributed	29.3%	64.2%	6.5%	100%
Statistically Significant at the 99% confidence level				

4.2.5 Time of Day

Almost 70% of all pedestrian crashes occurred during daylight hours and account for about 61% of fatalities. In contrast, 26% of pedestrian crashes occurred after dark but account for nearly 40% of the fatalities. These results indicate that nighttime pedestrian crashes are more likely to result in more severe injuries. Sixty-two percent of adults were victims of crashes during daylight hours. In contrast, over 80% of crashes involving children and elderly victims occurred during daylight hours. Men and women had a similar distribution of crashes by time of day; women (72%) are involved in slightly more crashes during daylight then men (67%). Although these differences were small, the analysis shows that the relationships between time of day and gender and age of the pedestrian are statistically significant.

4.2.6 Functional Classification

The functional classification of the roadways is also used to evaluate the characteristics of the location of the crashes. The City of Baltimore has over 6000 roadway segments classified by their functionality; their distribution is as follows: 3.3% primary unseparated, 0.6% primary separated, 96% local unseparated and 0.1% local separated. It is very interesting that when the classification of the roadways is evaluated for the crash dataset, the results are considerably different, as shown on Table 4.5. The majority of the crashes, nearly 65%, occurred on primary roadways, whereas just over 35% occurred in local streets. It can be seen in Figure 4.1 that a large number of accidents occurred in major arterials leading towards the center of the city. There is a large discrepancy between the percentage of crashes that occurred in primary roadways and the percentage of roadways that are of the primary classification. This inconsistency may signify that the primary roadways are more dangerous to pedestrians because of the high volume and speed of vehicles and the less pedestrian-friendly amenities such as sidewalks and crosswalks.

Table 4.5: Distribution of Functional Classification of Roadways

Functional Classification	Percentage of Roadways in Baltimore City	Percentage of Crashes in each type of
Primary - Unseparated	3.31%	61.5%
Primary - Separated	0.63%	3.1%
Local – Unseparated	95.97%	33.4%
Local - Separated	0.08%	1.9%
Total	100.0%	100.0%
Statistically Significant at the 99% confidence level		

This is further supported by examining the severity of injury among crashes in different roadway classifications. Although there is a higher percentage of crashes occurring in primary roadways, there is an even greater percentage of fatalities as compared to local streets: 82% for the former and 18% for the latter.

There are no significant differences for crashes and classification of roads by gender. However the analysis of functional classification of roadways shows a statistically significant difference between age groups. Table 4.6 shows that there is a lesser percentage of children (53%) involved in crashes on primary roadways than adults or elderly (over 70%). These results are not surprising because children are more likely to be playing or crossing the street in local or neighborhood streets than in primary highways.

Table 4.6: Functional Classification of Roadways by Age Group

Functional Classification	Total	Children	Adult	Elderly
Primary roadways	64.6%	52.9%	70.6%	74.8%
Local roadways	35.3%	47.2%	29.4%	25.2%
Total	100.0%	100.0%	100.0%	100.0%
Statistically Significant at the 99% confidence level				

4.2.7 Summary of Descriptive Statistics for Pedestrian and Crash Characteristics

To summarize the important results for the pedestrian and crash characteristics, Figures 4.2 and 4.3 show the distribution of crashes by the pedestrian characteristics by gender and age group respectively.

Figure 4.2: Main Findings for Pedestrian Characteristics by Gender

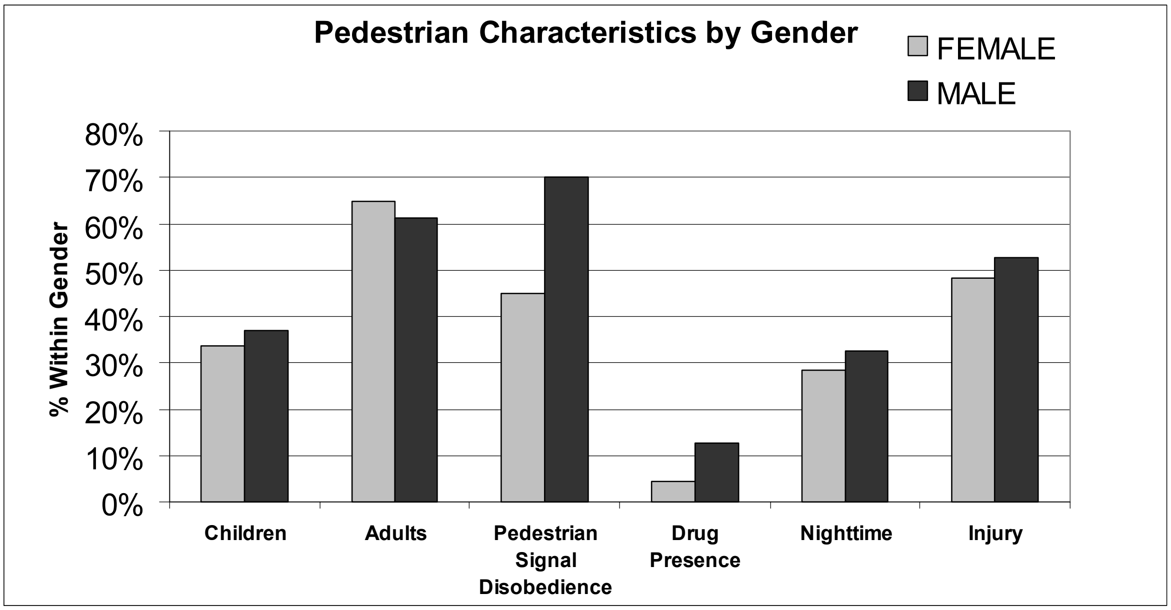
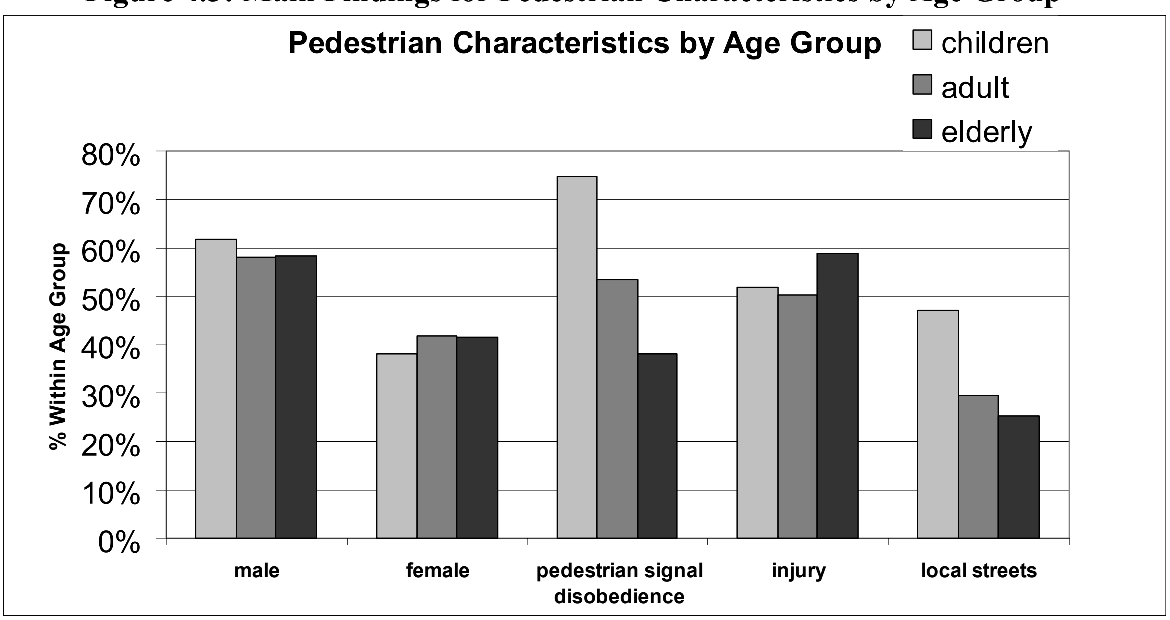


Figure 4.3: Main Findings for Pedestrian Characteristics by Age Group



4.3 Land Use Characteristics

Additional analysis of the crashes was performed using land use characteristics for the location of the pedestrian crash. Tables 4.7a and 4.7b show the land use attributes at the crash location by age groups. Tables 4.8a and 4.8b show the land use attributes at the crash location by gender. However, these results are based on descriptive statistics and are not controlled for area or population of each blockgroup. The result of multivariate analysis is shown in the subsequent section.

Table 4.7a: Pedestrian Crashes by Land Use for Age Groups

Characteristics of blockgroup where crash is located	Total N	Total %	Child	Adult	Elderly	
Pedestrian Activity						***
Low (less than 10%)	2040	72.4%	81.7%	67.4%	63.9%	
Medium (10% to 27%)	465	16.5%	15.9%	16.9%	15.1%	
High (> 27%)	313	11.1%	2.4%	15.6%	21.0%	
Total	2818	100.0%	100.0%	100.0%	100.0%	
% Parkland						***
Low (0% to 10%)	2397	85.1%	80.8%	87.6%	85.7%	
Medium (10% to 40%)	302	10.7%	13.4%	9.1%	10.9%	
High (> 40%)	119	4.2%	5.8%	3.3%	3.4%	
Total	2818	100.0%	100.0%	100.0%	100.0%	
Population density						***
Low (0 to 10,000 persons/ mile^2)	1097	38.9%	32.7%	42.0%	48.7%	
Medium (10,001 to 20,000 persons/ mile^2)	1208	42.9%	46.2%	40.9%	42.0%	
High (> 20,001 persons/mile^2)	513	18.2%	21.1%	17.1%	9.2%	
Total	2818	100.0%	100.0%	100.0%	100.0%	

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

Table 4.7b: Pedestrian Crashes by Land Use for Age Groups - continued

Characteristics of blockgroup where crash is located	Total N	Total %	Child	Adult	Elderly	
Median Household Income						
Low (< $25,000)	1130	40.1%	41.0%	39.8%	37.0%	
Medium ($25,000 to $40,000)	1392	49.4%	49.2%	49.5%	49.6%	
High (>$40,000)	296	10.5%	9.8%	10.7%	13.4%	
Total	2818	100.0%	100.0%	100.0%	100.0%	
% of SFRDU within ¼ mile from commercial parcel						***
Low (0% to 55%)	155	5.5%	7.9%	4.1%	5.0%	
Medium (55% to 85%)	289	10.3%	11.4%	9.6%	10.3%	
High (>85%)	2374	84.2%	80.7%	86.3%	84.2%	
Total	2818	100.0%	100.0%	100.0%	100.0%	
% of SFRDU within ¼ mile from bus stop						**
Low (40% to 80%)	89	4.4%	3.4%	4.8%	8.0%	
Medium (80% to 95%)	352	17.3%	15.2%	19.0%	14.8%	
High (95% to 100%)	1589	78.3%	81.3%	76.3%	77.3%	
Total	2030	100.0%	100.0%	100.0%	100.0%	
% Mixed-Use						***
Low (40% to 80%)	546	19.4%	23.7%	17.1%	15.1%	
Medium (80% to 95%)	547	16.2%	19.9%	14.4%	10.1%	
High (95% to 100%)	1815	64.4%	56.4%	68.4%	74.8%	
Total	2818	100.0%	100.0%	100.0%	100.0%	
% Density of Roadways						***
Low (40% to 80%)	598	21.3%	25.8%	18.8%	16.9%	
Medium (80% to 95%)	855	30.4%	30.8%	29.8%	35.6%	
High (95% to 100%)	1360	48.3%	43.3%	51.4%	47.5%	
Total	2813	100.0%	100.0%	100.0%	100.0%	

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

Table 4.8a: Pedestrian Crashes by Land Use for Gender

Characteristics of blockgroup where crash is located	Total N	Total %	Male	Female	
Pedestrian Activity					*
Low (less than 10%)	2143	71.8%	73.1%	69.8%	
Medium (10% to 27%)	506	17%	16.6%	17.5%	
High (> 27%)	336	11.3%	10.3%	12.7%	
Total	2985	100.0%	100.0%	100.0%	
% Parkland					
Low (0% to 10%)	2534	84.9%	84.4%	85.7%	
Medium (10% to 40%)	322	10.8%	10.8%	10.8%	
High (> 40%)	129	4.3%	4.8%	3.6%	
Total	2985	100.0%	100.0%	100.1%	
Population density					
Low (0 to 10,000 persons/ mile^2)	1160	38.9%	39.7%	37.6%	
Medium (10,001 to 20,000 persons/ mile^2)	1271	42.6%	42.4%	42.9%	
High (> 20,001 persons/mile^2)	554	18.6%	17.9%	19.5%	
Total	2985	100.0%	100.0%	100.0%	
Median Household Income					
Low (< $25,000)	1201	40.2%	40.6%	39.8%	
Medium ($25,000 to $40,000)	1474	49.4%	49.1%	49.8%	
High (>$40,000)	310	10.4%	10.3%	10.4%	
Total	2985	100.0%	100.0%	100.0%	
% within ¼ mile from commercial parcel					
Low (0% to 55%)	164	5.5%	5.7%	5.1%	
Medium (55% to 85%)	299	10.0%	10.2%	9.8%	
High (>85%)	2522	84.5%	84.1%	85.1%	
Total	2985	100.0%	100.0%	100.0%	
% within ¼ mile from bus stop					
Low (40% to 80%)	93	4.3%	4.4%	4.2%	
Medium (80% to 95%)	374	17.4%	17.7%	17.1%	
High (95% to 100%)	1681	78.3%	77.9%	78.8%	
Total	2148	100.0%	100.0%	100.1%	

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

Table 4.8b: Pedestrian Crashes by Land Use for Gender

Characteristics of blockgroup where crash is located	Total N	Total %	Male	Female
% Mixed-Use				
Low (40% to 80%)	576	19.3%	20.2%	18.0%
Medium (80% to 95%)	476	15.9%	15.7%	16.3%
High (95% to 100%)	1933	648%	64.1%	65.7%
Total	2985	100.0%	100.0%	100.0%
% Density of Roadways				
Low (40% to 80%)	623	20.9%	21.2%	20.5%
Medium (80% to 95%)	905	30.4%	30.8%	29.7%
High (95% to 100%)	1451	48.7%	48.0%	49.8%
Total	2979	100.0%	100.0%	100.0%

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

4.3.1 Pedestrian Activity

Journey to work data from the U.S. Census 2000 are used as the indicator for pedestrian activity at the blockgroup level. The results show that crashes have a negative correlation with levels of pedestrian activity. Areas with a low percentage of people who walk to work had the highest percentage (72%) of pedestrian vehicle collisions. These results may indicate that neighborhoods with lower pedestrian activity are more prone to have pedestrian crashes occur because motorists are not expecting pedestrians to be present.

However it is important to note that journey to work is only one measure of pedestrian activity. Section 4.5 will discuss the patterns resulting from the addition of pedestrian demand analysis. Figure 4.4 shows the spatial distribution of the pedestrian crashes by percentage of workers who walk to work. In addition, subsequent multivariate analysis will show the results of crashes in these areas when controlled for density and population of the area where the collision occurred.

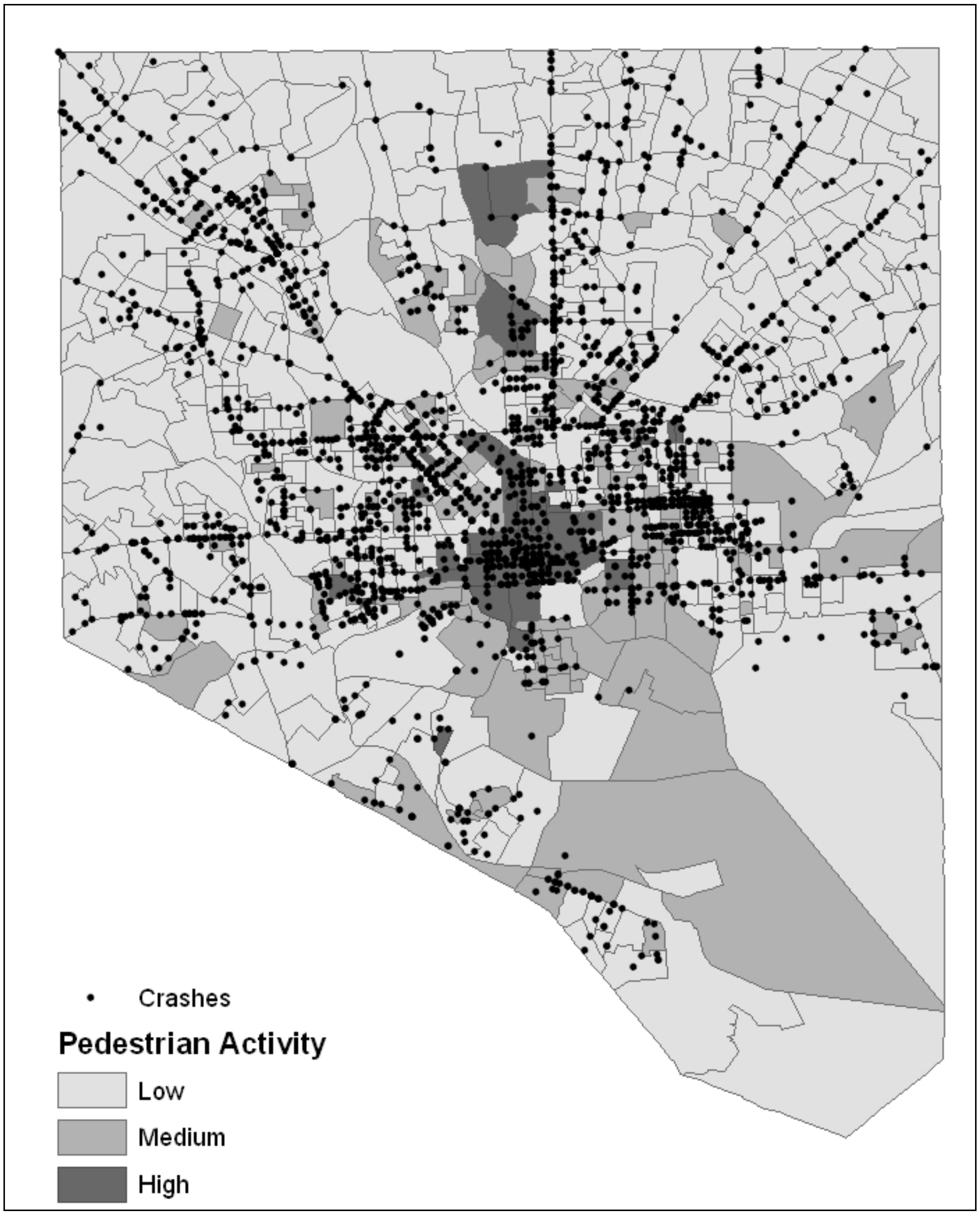


Figure 4.4: Pedestrian-Vehicular Crashes by Pedestrian Activity

The different age groups show different percentages of crashes in each category of pedestrian activity. Nearly 82% of children involved in crashes were victims in locations with low pedestrian activity in contrast with adults (67%) and elderly (64%). Twenty-one percent of pedestrian crashes involving the elderly occurred in locations with high pedestrian activity whereas 15.5% of adults and 2.5% of children involved in crashes were victims in areas with high pedestrian activity. These differences are statistically significant.

Men were involved in a slightly higher percentage of pedestrian crashes than women in locations with low pedestrian activity (73% for the former and 70% for the latter). However, women were involved in a higher percentage of crashes in medium and high pedestrian activity areas than men. It is important to note that these differences are small and are statistically significant at the 10% level.

4.3.2 Park area

The relationship between park land and crashes is considered. Blockgroup areas are categorized by the amount of area dedicated to parks. The results show that areas with the least amount of park land have the highest number of pedestrian-vehicular crashes.

Eighty-five percent of crashes occurred in areas with few parks and only 4% of crashes occurred in areas with high levels of parkland. These results are rather surprising because one might infer that locations with high levels of parkland would generate more pedestrians and therefore more opportunities for crashes. Again, these results are not controlled for demand. The effect of demand analysis will be discussed in Section 4.5. In addition, it may be the case that motorists driving near park areas may be more attentive to pedestrian traffic and therefore drive slower and avoid possible collisions with pedestrians.

The gender of the pedestrian involved in crashes does not appear to have a significant relationship with park area. However, the age of the pedestrian involved in crashes does have a statistically significant relationship with amount of park area. A low percentage, just over 3%, of adults and elderly are involved in crashes in areas with high provision of parks. In contrast, nearly twice as many, almost 6% of pedestrian crashes involving children occur at locations with high density of parkland. These numbers may be explained by children running across the street to reach the park facilities.

4.3.3 Population Density

Population density of neighborhoods has a statistically significant effect on the frequency of pedestrian crashes. Areas with medium levels of population density have the highest percentage (43%) of pedestrian crashes, compared to 39% in low density and 18% in high density areas. Population density does not appear to have a significant relationship with gender of the pedestrian involved in the crash. However, there is a statistically significant difference between the age groups. It is notable that for areas with high population density, a significantly higher percentage of crashes involve children at 21% than adults (17%) and elderly (9%).

4.3.4 Medium Household Income

Median household income does not have a statistically significant effect on crashes. However the distribution of crashes shows that areas with high median income comprise a much lower percentage of accidents than areas with low or medium income. This may be a result of areas with higher income having a higher rate of vehicle ownership and therefore fewer pedestrians.

4.3.5 Commercial Accessibility

The relationship between pedestrian-vehicular crashes and pedestrian accessibility to commercial land uses is examined. As stated in the methodology, this measure of commercial accessibility is defined as the percentage of single-family residential units in the area within ¼ mile of commercial uses. The results show that 84% of all pedestrian crashes occurred in areas with high levels of pedestrian accessibility to commercial sites.

These statistics may be the result of commercial sites being clustered in the downtown area and therefore drawing increased pedestrian volume. There is no significant relationship between commercial accessibility and gender. However, analysis of commercial accessibility impact on crashes shows that although it is small, there is a statistically significant difference among age groups. It is not surprising that the adults category has a greater percentage of crashes in high commercial areas (over 86%), compared to 81% for children and 84% for elderly.

4.3.6 Transit Accessibility

The relationship between transit accessibility and pedestrian crashes was another land use variable analyzed in this study. The transit access measure was calculated as the percentage of single-family residential units in a tract that have a transit stop within ¼ mile. Areas with high transit access show the highest percentage of crashes, over 78%, compared to 17% and 4% for areas with medium and low transit accessibility respectively. These results may indicate that areas with good accessibility to bus stops have more pedestrians and therefore more opportunity for pedestrian crashes. Gender also did not have a significant affect on the relationship of crashes and transit accessibility. However, it is interesting that a higher percentage of crashes involving children occur in areas of high transit access (81%), than adults (76%) or elderly (77%). The effect of age is statistically significant to the 90% level of confidence.

4.3.7 Mixed-Use

The percentage of mixed use property in neighborhoods has a statistically significant effect on the frequency of pedestrian crashes. Areas with high levels of mixed use have the highest percentage (64.5%) of pedestrian crashes, compared to 16% in medium density and 19.5% in low density areas. The high number of crashes in locations with high levels of mixed use may be explained by the fact that locations with more commercial entities tend to generate more pedestrian traffic and therefore may create increased opportunities for crashes. The mixed use variable does not appear to have a significant relationship with the gender of the pedestrian involved in the crash. However, there is a statistically significant difference between the age groups.

It seems that for the low levels of mixed use, there are a significantly higher percentage of crashes involving children (24%) than adults (17%) and elderly pedestrians (15%). Conversely, there are greater percentages of crashes involving the elderly (75%) and adults (68%), than children (56%) for the areas with high levels of mixed use. This is likely the result of a greater percentage of adults and elderly walking in areas with higher levels of commercial use. Children, on the other hand, are more likely to be playing in the streets or walking to school on residential neighborhoods where the levels of mixed use are lower.

4.3.8 Density of Roads

The relationship between density of roadways in each blockgroup and crashes is considered. The results show that the higher the density of roadways in an area, the higher the number of pedestrian-vehicular crashes. Over 48% of crashes occurred in areas with high roadway density, compared to 30.5% in medium roadway density areas and 21% in low roadway density areas. These results are not surprising because one might deduce that locations with high density of roadways would generate more pedestrians and therefore more opportunity for crashes. Again, these results are not controlled for demand. The analysis on the effect of demand is imperative to understand the occurrence of crashes.

The gender of the pedestrian involved in crashes does not appear to have a significant relationship with density of roadways. However, the age of the pedestrian involved in crashes does have a statistically significant relationship with density of roadways. A higher percentage of children (26%) are involved in crashes in areas with high density of roadways, compared to adults (19%) and elderly (17%) populations.

4.3.9 Summary of Descriptive Statistics for Land Use Variables

Figure 4.5 shows the distribution of crashes based on the characteristics of the area where the crash occurred. These characteristics include the policy variables of interest: pedestrian activity, mixed-use, density of roadways, commercial accessibility and transit accessibility.

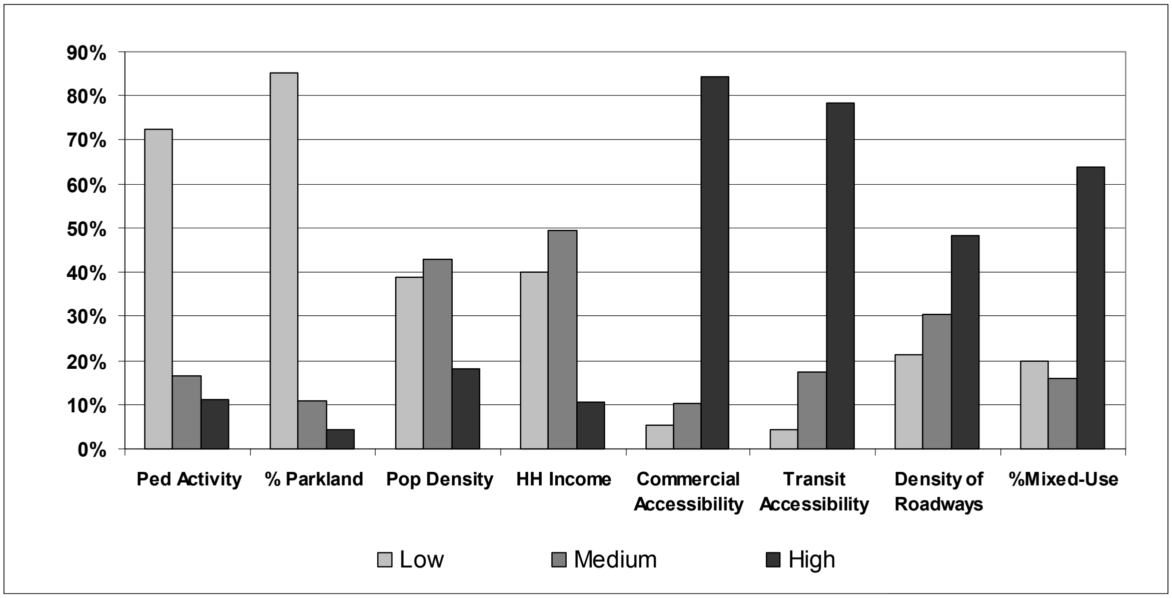


Figure 4.5: Summary of Land Use Characteristics

4.4 Multivariate Analysis

Table 4.9 shows two linear regression models using different dependent variables: natural log of crash density and natural log of crash density per population. The results show that the variables are not highly correlated and therefore there is no significant multicollinearity among the variables in the model. However, both models are corrected for heteroskedasticity and therefore robust standard errors are reported. Although both models shown have different coefficients for each variable, they are consistent in the interpretation of nearly all the independent variables. Population density was removed as an independent variable in Model 2 because the dependent variable is controlled for population as shown in Table 4.9.

Table 4.9: Linear Regression Models

Regressors	**Model 1: ln (crash density)**	**Model 2: ln (crash density/population)**
Dependent Variable	**2.855** *** (7.48)	**-2.123** *** (-4.32)
Population Density	**-0.0501*** (9.59)	
Total Households	**-0.0010*** (-1.79)	**-0.0026***** (-3.29)
Employment	-0.0001 (0.23)	-0.0003 (0.33)
Income (in $1000)	-0.0056 (-1.18)	**-0.0107**** (-2.01)
Education	-0.2654 (-0.61)	0.3403 (0.92)
Race	**0.4025***** (2.57)	**0.6091***** (3.41)
Vehicles ownership	-0.2376 (-0.67)	**-0.895**** (-1.94)
% Children	-0.0217 (-0.04)	-0.8164 (-1.19)
Density of Roads	**0.0231***** (8.00)	**0.0328***** (10.02)
Pedestrian Activity	**-0.0029**** (-0.59)	-0.0023 (-0.40)
% Park	-0.0020 (-0.40)	-0.0091 (-1.60)
Commercial Accessibility	**0.0041**** (2.25)	**0.0058***** (2.71)
Transit Accessibility	**-0.0051**** (-2.44)	**-0.0085***** **(-3.35)**
Mixed Use	**1.1636***** (6.01)	**0.7471***** (3.80)
N	474	474
R2	0.541	0.542
Note: t-statistic is shown in parentheses Corrected for heteroskedasticity No multicollinearity present		
* Statistically significant at the 10% level ** Statistically significant at the 5% level *** Statistically significant at the 1% level		

The models show that there is a negative relationship between number of households and crashes. In model 2, this relationship is significant at the 99% level of confidence. This result may indicate that the areas with greater numbers of households have a tendency to have less pedestrian-vehicle crashes occur.

Statistically significant differences in Model 2 seem to indicate that fewer crashes are happening in locations with higher income levels. This may be explained by the fact that these higher income areas may have a greater percentage of the population that can afford cars. Therefore these areas may have fewer pedestrians and thus less opportunity for crashes. This statistic is further supported when looking at the relationship of vehicle ownership and crashes in Model 2. Vehicle ownership also has a negative statistically significant effect on crashes, which may be a result of there being fewer pedestrians in that particular area and therefore fewer opportunities for a pedestrian-vehicular crash.

The race variable is defined as the percentage of non-white population in the area. This variable is positively related to crashes. This statistically significant result shows that locations with a greater percentage of minority groups are more likely to have pedestrian crashes.

The models show that there is a negative relationship between crashes and percentage of parks. Parks are an attractor for pedestrian activity. Although this relationship was not statistically significant in either model, this negative relationship may be indicative of drivers being more careful around areas where there would be a higher chance of pedestrians, specifically children crossing the streets.

The interest of this thesis is to study the relationship between the land use variables and pedestrian crashes. The land use variables of interest in the model are those that show the different levels of population and infrastructure density: pedestrian activity, density of roadways, mixed use, commercial accessibility, and transit accessibility. Model 1 shows statistically significant results for all five of these variables and Model 2 shows significant results to all except pedestrian activity. Model 2 shows that the propensity of crashes increases with an increase in density of roads (3%), commercial accessibility (0.6%) and mixed use. The reasons for this may be related to the greater number of pedestrians in that area and therefore greater opportunity for a pedestrian-vehicular crash. In the more suburban areas, with lower housing densities, there is usually less pedestrian infrastructure, more vehicle ownership, and therefore less pedestrian activity. The percentage for mixed use is not reported because the coefficient shown in the model is very large and therefore may be endogenous with crashes.

In addition, the coefficient for the variable for percentage of children in the area was negative and not statistically significant in the model. However, the coefficient for the relationship between percentage of children and crashes is very large. The coefficient for the variable for vehicle ownership is also surprisingly high. These coefficients are indicative that there may be a problem of endogeneity in these models. Particularly, the percentage of children and the vehicle ownership variables are likely to be endogenous. Families with children may choose not to move to a location where there are a high number of pedestrian-vehicular crashes. Similarly, if the location of residence is unsafe for walking, residents may be more likely to own a vehicle. Subsequent work will address the endogeneity issue.

It is interesting that transit accessibility and pedestrian activity have statistically significant negative impacts on crashes. Model 2 shows that a one unit increase in transit accessibility results in a decrease of crashes by approximately 0.8% and a one unit increase in pedestrian activity results in a decrease of crashes by approximately 0.2%. This result may indicate that in areas with high transit accessibility, more people tend to walk and less tend to drive, and thus there may be a decrease in vehicles on the road. In addition, drivers may be more careful when driving through an area where there is a high opportunity for pedestrian activity. Figure 4.6 shows the results of the policy variables for Model 2.

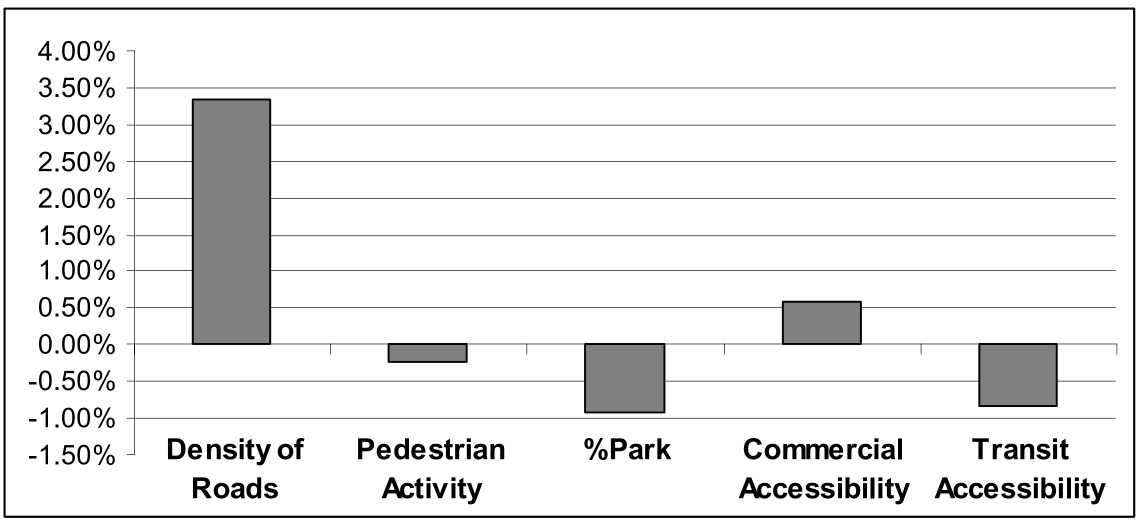


Figure 4.6: Summary of Policy Variables

Model 1 shows a highly statistical significant negative relationship between population density and crashes. The model is then improved by controlling for the population of each blockgroup (Model 2). The results of these regressions suggest the importance of controlling for demand.

4.5 Demand Analysis

A linear regression model was performed using pedestrian exposure as the dependent variable. Pedestrian exposure (crashes per demand) is regressed to control for the demand of pedestrians in each of those areas. Although the demand data used are estimated, the results indicate the importance of controlling crash models for the number of pedestrians exposed to potential crashes. Table 4.10 shows the result for the pedestrian exposure model.

It is interesting that nearly all location and land use characteristics are now statistically significant. However, many of the variables have the opposite sign than in the previous models. These results suggest than when controlling for demand, crashes are less likely to occur in denser urban areas. Percentage of crashes decreases in locations with greater number of households, higher density of roadways, higher pedestrian activity and higher commercial accessibility.

Conversely, the variables for percent park and transit accessibility have a positive effect on crashes. The City of Baltimore has good bus service coverage and therefore few homes are not within ¼ mile of a bus stop. The variable for transit may not be representing the true effect of transit on crashes because there is little variation of transit accessibility among blockgroups in the data set. Similarly, the percent park variable may be problematic. The Clifton et al (2004) model used for the demand data does not consider parks as a generator for pedestrian activity. Their model is based on travel diaries and does not account for leisure trips. Therefore, the percent park variable may be underestimating the demand variable and therefore overestimating pedestrian exposure.

Table 4.10: Linear Regression Model controlled for Demand

Regressors	1000*ln (Pedestrian Exposure)
Dependent Variable	**1.545 ***** (3.169)
Population Density	0.003 (0.608)
Total Households	**-0.002***** (-3.449)
Employment	**0.002**** (2.498)
Income (in $1000)	**-0.022***** (-2.677)
Education	**-0.901*** (-1.770)
Race	**0.837**** (2.400)
Vehicles ownership	**1.298***** (3.314)
% Children	**-1.314**** (-2.028)
Density of Roads	**-0.016***** (-5.608)
Pedestrian Activity	**-0.008*** (-1.717)
% Park	**0.015***** (3.395)
Commercial Accessibility	**-0.059**** (-3.287)
Transit Accessibility	**0.060***** (3.371)
Mixed Use	**0.574***** (3.244)
N	481
R2	0.291
Note: t-statistic is shown in parentheses	
*, **, *** Statistically significant at the 10%, 5% and 1% levels respectively	

Income and education appear to have negative effects on pedestrian exposure. This result may indicate that areas with residents of higher income levels may have pedestrian environments more conducive to walking. In addition, there is a positive relationship between race and crashes. Census data show that in Baltimore City, blockgroups with greater percentages of minority groups (non-white populations) tend to have lower medium incomes. Therefore it follows that the effect of the race variable should correspond to income. Residents of neighborhoods with higher education and income status tend to have more lobbying power to ensure that their streets have better sidewalks and decreased traffic speeds.

The coefficients for percent children and vehicle ownership are very high thus supporting the possibility of endogeneity. The percent children variable is negatively correlated to pedestrian exposure. This result in not surprising because motorists tend to slow down and be more attentive to pedestrians in locations where there is a greater demand of children such as school zones and parks. Conversely, vehicle ownership has a positive effect on crashes. This may be explained by the presence of more cars leading to greater opportunity of pedestrian-vehicular crashes.

The most surprising result is the positive relationship between mixed-use and crashes. Notable is the fact that the correlation between mixed-use and crashes remained unchanged after controlling for demand. This result is contradictory to the effect of the other land-use variables on crashes and may indicate that in fact the mixed-use variable is confounded with demand.

Chapter 5: Conclusion

5.1 Contribution

Many studies have focused on the relationship between pedestrian activity and land use and on the correlation between personal characteristics and crashes. However, less research has been devoted to the effect of land use on pedestrian-vehicular crashes and even fewer have linked land use to pedestrian exposure. The main contribution of this study is the analysis of the relationship between land use and pedestrian exposure. The results show that controls for demand are essential in understanding the interactions between crashes and land use. Table 5.1 shows the differences in relationship between the pedestrian crash and the pedestrian exposure models.

Table 5.1: Comparison of Linear Regression Models

Regressors	Crash Density/ Population	Pedestrian Exposure
Dependent Variable	- ***	+ ***
Population Density	+	+
Total Households	- ***	- ***
Employment	-	+ **
Income (in $1000)	+ **	- ***
Education	+	- *
Race	+ ***	+ **
Vehicles ownership	- **	+ ***
% Children	-	- **
Density of Roads	+ ***	- ***
Pedestrian Activity	-	- *
% Park	-	+ ***
Commercial Accessibility	+ ***	- **
Transit Accessibility	- ***	+ ***
Mixed Use	+ ***	+ ***
*, **, *** Statistically significant at the 10%, 5% and 1% levels respectively		

5.2 Findings

The findings in this research project suggest that there is a high number of pedestrian-vehicular crashes among urban areas with high population and roadway densities and good commercial accessibility. However, these areas also generated the highest percentage of pedestrian demand. When controlling for demand, these same areas are found to have negative relationships to pedestrian exposure. The results may justify the promotion of more urban neighborhood designs with higher population densities and good accessibility to commercial areas. The findings also advocate the need to guide safety policy investments to these urban areas with high pedestrian activity. These policy investments may include better sidewalks, better lighting, pedestrian education, and lower vehicular speeds in locations of high pedestrian activity. The results also showed that a large number of accidents happen where there are no pedestrian signals. Better signalization of intersections may be necessary to promote pedestrian safety. In addition, areas with high risk exposure are the areas that may have low pedestrian activity but high crash rates. Therefore it is also important to address safety policy geared towards those areas.

The disproportionately high number of children involved in crashes is also a concern and indicates a need to guide pedestrian safety policy towards children. This may be accomplished by decreasing vehicular speeds, improving sidewalk conditions and buffers to the streets and better signalizing locations where there is a greater propensity of children being present, such as school zones and tourist attractions. Better education campaigns on traffic safety geared towards children may also be warranted.

The evidence from the analysis presented here suggests that in general, there are few significant gender effects in the majority of pedestrian crashes. Women tend to be involved in fewer pedestrian crashes overall and when they are involved, appear to exhibit less risk-taking behaviors, such as violating traffic laws and consuming alcohol or drugs. Women were slightly less likely to be injured in a crash and less likely to die as a result. The effects of land use on pedestrian crash rates did not show significant effects by gender. However, a higher percentage of women's crashes occur in high pedestrian activity areas, which may be reflective of the distribution of areas where women walk. These findings are not new, but do confirm the evidence presented in previous research.

5.3 Shortcomings and Future Research

The models discussed in the previous chapter show some important results regarding the relationship between land use and pedestrian-vehicular crashes. However, there may be some omitted variables in these models that can be critical when determining safety policy.

As an improvement to this study, future research should focus on acquiring additional control variables including pedestrian demand variables, microscale design measures and traffic information. Actual pedestrian counts as a measure of pedestrian demand would significantly improve the model. Additional traffic information should also be added as control variables including vehicular speed and number of vehicles per day.

Analysis of sub samples of the data may also yield interesting results. In particular, segregate models for different age groups, times of day, and specific minority populations could be performed to observe the variability of crashes among these groups.

The questions of where, why and when pedestrians walk are difficult to answer but essential when determining safety policy. Microscale design measures, such as distance of buildings to roadway, sidewalk width and condition and presence of buffer zones between roadway and sidewalks are important and would improve the model substantially.

References

Cervero, R. (2002). Built Environments and Mode Choice: Toward a Normative Framework. Transportation Research – Part D, 265-284.

Clifton, K. J., Davies, G., Allen, W. G., Raford, N. (2005). Pedestrian Flow Modeling for Prototypical Maryland Cities. Project report prepared for Maryland Department of Transportation, Maryland Highway Safety Office, Hanover, MD. Available at: http://www.smartgrowth.umd.edu.

Dieleman, F., Dijst, M., Burghouwt, G. (2002) Urban Form and Travel Behavior: Micro-level Household Attributes and Residential Context. Urban Studies, Vol. 39 Issue 3, 507-528.

Handy, S. (1996). Urban Form and Pedestrian Choices: Study of Austin Neighborhoods. Transportation Research Record 1552, 135-144.

Hanson, S. (1982) The Determinants of Daily Travel-activity Patterns: Relative Location and Sociodemographic Factors, Urban Geography, 3, 179–202.

Hebert-Martinez, K. L. and Porter, B.E. (2004). The likelihood of becoming a pedestrian fatality and drivers' knowledge of pedestrian rights and responsibilities in the Commonwealth of Virginia. Transportation Research - Part F. Vol. 7, 43-58

Jang, Tae Youn (2003). Causal Relationship Among Travel Mode, Activity, and Travel Patterns. Journal of Transportation Engineering, Vol. 129 , 16-22.

National Highway Traffic Safety Administration. (2001). Traffic safety facts 2000 (DOT HS 809 337). Washington, DC: US Department of Transportation. Available at: http://www-nrd.nhtsa.dot.gov/pdf/nrd-30/NCSA/TSFAnn/TSF2000.pdf. Accessed: March 2004.

National Highway Traffic Safety Administration. (2004). Traffic safety facts 2002-pedestrians. Available at: http://www-nrd.nhtsa.dot.gov/pdf/nrd-30/NCSA/TSF2002/2002pedfacts.pdf. Accessed: March 2004.

Preusser, D.F., Wells, J.K., Williams, A.F. and Weinstein, H.B. (2002). Pedestrian crashes in Washington, DC and Baltimore. Accident Analysis and Prevention, 34, 703-10.

Rajamani, J., C. Bhat, S. Handy, G. Knaap and Y. Song (2003). Assessing Impact of Urban Form Measures on Nonwork Trip Mode Choice After Controlling for Demographic and Level-of-Service Effects. Transportation Research Record 1831, 158-165.

Retting, R., Fergunson, S. and McCartt, A.(2003). A Review of Evidence- Based Traffic Engineering Measures Designed to Reduce Pedestrian-Motor Vehicle Crashes. American Journal of Public Health Vol. 93, 1456-1463.

Rosenbloom, S. (1998). Trends in Women's Travel Patterns, in "Women's Travel Issues, Proceedings from the Second National Conference," Report FHWA-PL-97-024, FHWA, U.S. Department of Transportation.

Stutts, J., Hunter, W. and Pein, W. (1996). Pedestrian Crash Types: 1990s Update. Transportation Research Record 1538, 68-74.

Thom, B. (2003). Risk-taking behavior in men: substance use and gender. Prepared for The Health Development Agency, London, UK.

U.S. Census Bureau (2000). Census 2000 Summary File 3. Available at: www.census.gov. Accessed from 2004-2005.