OKLAHOMA TRANSPORTATION CENTER

Economic Enhancement through Infrastructure Stewardship

Analysis of FARS Data on State Highways in Oklahoma

Jonathan C. Comer, Ph.D.
Leonard S. Bombom, M.Sc.
Nicholas J. Rose, M.Sc.

OTCREOS11.1-12-F

Oklahoma Transportation Center
2601 Liberty Parkway, Suite 110
Midwest City, Oklahoma 73110

Phone: 405.732.6580
Fax: 405.732.6586
www.oktc.org

DISCLAIMER

TECHNICAL REPORT DOCUMENTATION PAGE

1. REPORT NO.	2. GOVERNMENT ACCESSION NO.	3. RECIPIENTS CATALOG NO.
OTCREOS11.1-12-F		

4. TITLE AND SUBTITLE Analysis of FARS Data on State Highways in Oklahoma	5. REPORT DATE November 2012 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Jonathan C. Comer, Leonard S. Bombom, and Nicholas J. Rose	8. PERFORMING ORGANIZATION REPORT
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Geography Oklahoma State University 337 Murray Hall Stillwater, OK 74078	10. WORK UNIT NO. 11. CONTRACT OR GRANT NO. DTRT06-G-0016
12. SPONSORING AGENCY NAME AND ADDRESS Oklahoma Transportation Center (Fiscal) 201 ATRC, Stillwater, OK 74078 (Technical) 2601 Liberty Parkway, Suite 110 Midwest City, OK 73110	13. TYPE OF REPORT AND PERIOD COVERED Final September 2011-November 2012 14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES
University Transportation Center

16. ABSTRACT

Analysis of fatality automobile accident data can be challenging in rural areas where a relatively small number of such accidents occurs on specific sections of highways. Combining crash data for 1998 to 2011 from the Fatality Analysis Reporting System (FARS) and highway networks and design specifications from the Oklahoma Department of Transportation (ODOT), this project employs Poisson regression analysis to determine what roadway characteristics (e.g., grade, geometry, and design) are most associated with fatal crashes on predominantly rural segments of highways in Oklahoma.

A state-wide model found that median width and type, terrain type, surface type and thickness, grade, and outside shoulder width were all significant predictors of fatality crashes. Vertical grade was the most significant variable in the state model, with uphill grades contributing to significantly more fatal crashes while flat roads modestly reduce the risk. Terrain type was also very significant, both flat and rolling terrain categories having higher contributions to fatality crash rates than baseline levels in residential areas. Outside shoulder width significantly negatively correlated with fatalities, as wider shoulders mean more room for drivers to recover or at least shed speed. Median width was also significant with the expected negative sign as wider medians result in fewer fatalities. Surprisingly, median type was only moderately significant, though as expected most other median types did contribute to higher crash rates compared to cable barriers.

The results provide information about what combinations of highway design traits have contributed most to past crashes and therefore can identify potentially dangerous road segments system-wide. This information will help transportation engineers evaluate current construction practice and seek ways to address design issues that are shown to contribute significantly to serious crashes.

17. KEY WORDS	18. DISTRIBUTION STATEMENT		
FARS, Highways, Fatalities, Rural Oklahoma	No restrictions. This publication is available at www.oktc.org and from the NTIS.		
19. SECURITY CLASSIF. (OF THIS REPORT) Unclassified	20. SECURITY CLASSIF. (OF THIS PAGE) Unclassified	21. NO. OF PAGES 49 + covers	22. PRICE

SI (METRIC) CONVERSION FACTORS

Approximate Conversions to SI Units

Symbol	When you know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	25.40	millimeters	mm
ft	feet	0.3048	meters	m
yd	yards	0.9144	meters	m
mi	miles	1.609	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm
ft^2	square feet	0.0929	square meters	m^2
yd^2	square yards	0.8361	square meters	m^2
ac	acres	0.4047	hectares	ha
mi^2	square miles	2.590	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.0283	cubic meters	m^3
yd^3	cubic yards	0.7645	cubic meters	m^3
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.4536	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams	Mg
		TEMPERATURE (exact)		
°F	degrees Fahrenheit	(°F-32)/1.8	degrees Celsius	°C
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.448	Newtons	N
lbf/in^2	poundforce per square inch	6.895	kilopascals	kPa

Approximate Conversions from SI Units

Symbol	When you know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.0394	inches	in
m	meters	3.281	feet	ft
m	meters	1.094	yards	yd
km	kilometers	0.6214	miles	mi
		AREA		
mm^2	square millimeters	0.00155	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.196	square yards	yd^2
ha	hectares	2.471	acres	ac
km^2	square kilometers	0.3861	square miles	mi^2
		VOLUME		
mL	milliliters	0.0338	fluid ounces	fl oz
L	liters	0.2642	gallons	gal
m^3	cubic meters	35.315	cubic feet	ft^3
m^3	cubic meters	1.308	cubic yards	yd^3
		MASS		
g	grams	0.0353	ounces	oz
kg	kilograms	2.205	pounds	lb
Mg	megagrams	1.1023	short tons (2000 lb)	T
		TEMPERATURE (exact)		
°C	degrees Celsius	9/5+32	degrees Fahrenheit	°F
		FORCE and PRESSURE or STRESS		
N	Newtons	0.2248	poundforce	lbf
kPa	kilopascals	0.1450	poundforce per square inch	lbf/in^2

ANALYSIS OF FARS DATA ON STATE HIGHWAYS IN OKLAHOMA

Final Report

November 2012

Jonathan C. Comer, Ph.D.
Professor, Department of Geography

Leonard S. Bombom, M.Sc.
Graduate Research Assistant, Department of Geography

Nicholas J. Rose, M.Sc.
Graduate Research Assistant, Department of Geography

Oklahoma State University
Stillwater, OK 74078

Oklahoma Transportation Center
2601 Liberty Parkway, Suite 110
Midwest City, Oklahoma 73110

ACKNOWLEDGMENTS

The authors would like to thank Mark Brown (GIS Development) and Justin Calvarese (Pavement Management Branch) in the Oklahoma Department of Transportation (ODOT) Planning and Research Division for assistance in providing and clarifying the road network and condition data obtained from ODOT. Also, Joseph Havlicek, Rob Barnes, and Patrick Adrian Campbell at the University of Oklahoma Intelligent Transportation Systems (ITS) Laboratory provided invaluable assistance and data support. Most importantly, this project would not have been possible without the support of Jim Self and the Oklahoma State University Center for Local Government Technology, from which the majority of the cost-share funds for this project were provided via the Tribal Technical Assistance Program (TTAP). This support is greatly appreciated.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AADT	Annual Average Daily Traffic
AASHTO	Association of American State Highway and Transportation Officials
BIA	Bureau of Indian Affairs
FARS	Fatality Analysis Reporting System
FHWA	Federal Highway Administration
FLHP	Federal Lands Highway Program
GIS	Geographic Information Systems
HMVMT	Hundred Million Vehicle Miles Travelled
HRRRP	High Risk Rural Roads Program
HSIP	Highway Safety Improvement Program
IRI	International Roughness Index
NB	Negative Binomial (regression)
NHTSA	National Highway Traffic Safety Administration
ODOT	Oklahoma Department of Transportation
OkTC	Oklahoma Transportation Center
ROR	Run Off Road (accidents)
SAFETEA-LU	Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users
SHSP	Strategic Highway Safety Plan
TRB	Transportation Research Board

EXECUTIVE SUMMARY

This project studied the locations of highway fatalities in Oklahoma with a goal of identifying the physical characteristics of the state's highways (i.e., grade, geometry, and design) that contributed to higher rates of fatality crashes. While agencies such as the Oklahoma Department of Transportation (ODOT) can promote driver education and awareness initiatives that encourage safer driving and hence reduce fatalities, ODOT has far more control over highway safety via design, construction, and maintenance activities. This project used statistical analysis to uncover relationships between fatality crashes and road characteristics along predominantly rural (including tribal) segments of the state's roadways. This information can help transportation engineers evaluate current construction practice and seek ways to address design issues that are shown to contribute significantly to serious crashes. This was accomplished with fatality data from 1998 to 2011 for Oklahoma from the national Fatality Analysis Reporting System (FARS) database, Oklahoma road inventory data from ODOT, and other related data sources such as the International Roughness Index (IRI). Data management and integration was accomplished using Geographic Information Systems (GIS), though statistical analysis was conducted with the statistical package SPSS.

Crash analysis has historically focused on high-crash locations, a technique also known as "black spot" analysis. Because fatality and serious injury accidents are so rare and spread out geographically, black spot analysis usually includes crashes of all severities recorded by the state or local agency responsible for such data. However, this approach biases the results towards less severe crashes at urban intersections simply due to the large amounts of traffic that pass through those locations, and safety treatments arising from black spot analysis have not been very effective at reducing fatal and serious injury crashes system-wide. Notably, run-off-road (ROR) crashes comprise over half of all fatal crashes in the United States and typically occur in rural areas, at high-speed, on two-lane highways, and involve a single vehicle. Black spot analysis rarely identifies clusters of these types of accidents or offers direction on remediation.

Thus, research is needed that isolates specific design *elements* rather than specific *locations* (or intersections) that contribute to elevated serious accident rates, helping transportation engineers to identify corrective measures for existing highways and to develop new designs for future (re)construction. This approach is called "systems" analysis because both the analysis and the resulting safety enhancements are done system-wide. There is strong empirical evidence that relatively low-cost system-wide treatments, such as shoulder rumble strips, are very beneficial in reducing ROR crashes and hence fatality accidents. The results of this research should lead to the identification and implementation of similar methods that can be broadly beneficial in the state.

The nature of the crash data used here, especially the use of discrete fatality counts along highway segments, necessitated more advanced methods than simple linear regression. Shortcomings of linear regression have been frequently documented in the literature and more advanced models like Poisson regression and negative binomial (or Poisson-gamma) regression are advocated, among other methods. Both methods were employed in this study, though ultimately end results proved very similar and Poisson regression was deemed adequate to the task.

Three different modeling frameworks were used. First, thirteen counties in northeastern Oklahoma were analyzed and results between Poisson and negative binomial regression were compared. Surface roughness, thickness, and (pavement) type, median width and type, and elevation were significantly related to fatality crashes, while curvature was only moderately related and vertical grade was insignificant.

A state-wide Poisson model was then developed which found that median width and type, terrain type, surface type and thickness, grade, and outside shoulder width were all significant predictors of fatality accidents. Vertical grade was the most significant variable in the state model, with uphill grades contributing to significantly more fatal crashes while flat roads modestly reduce the risk, relative to the rate that downhill grades cause fatal crashes. Terrain type was also very significant, both flat and rolling terrain categories having higher contributions to fatality crash rates than baseline levels in residential areas. Surface type and surface thickness were both very significant, but these were harder to interpret in a meaningful way. Outside shoulder width significantly negatively correlates with fatalities, as wider shoulders mean more room for drivers to recover or at least shed speed. Median width was also significant with the expected negative sign as wider medians result in fewer fatalities. Surprisingly, median type was only moderately significant, though most other median types did contribute to higher fatal crash rates compared to cable barriers as might be expected.

Finally, models were constructed individually for each of ODOT's eight field divisions. Perhaps due to fewer numbers of observations, no more than two independent variables were significant in any one field division and a few field divisions had no significant predictors. However, field divisions 1 (east central), 5 (southwest), and 8 (northeast) had extremely significant overall models while field division 3 (central) achieved moderate significance. In contrast, field divisions 2 (southeast), 4 (north central), and 7 (south central) had very weak models and field division 6 (northwest, including Panhandle) achieved virtually no goodness of fit or significance. These differential results can help transportation personnel modify possible system-wide improvements at the field division level, as the great breadth of Oklahoma's terrain, elevations, and traffic levels means few true system-wide treatments will probably be as effective as rumble strips or cable barriers have been.

1.0 INTRODUCTION

1.1 Problem Statement

This project examined Oklahoma highway fatalities because the Oklahoma Department of Transportation (ODOT) seeks system-wide treatments that can further reduce highway accidents, injuries, and fatalities. Relatively cheap and effective treatments, like shoulder rumble strips, are highly desirable because they help reduce accident rates system-wide without waiting for hot spots to emerge from accident data. Though this research was neither funded nor supervised by ODOT, the origins of this topic filtered down from the state agency through the FY11.1 OkTC Funding Competition as a “pull” topic under the **Traffic** subheading. As such, specific information about ODOT’s interest in analyzing data from the Fatality Analysis Reporting System (FARS) was obtained through personal conversations with several ODOT personnel in the summer of 2011.

“Hot spot” (or black spot) analysis is the most common form of accident analysis in the literature and has been useful in highlighting high crash incidence locations. Due to data limitations, however, all crashes are typically included in such research, which then usually isolates high-frequency, low-severity crashes at urban intersections. While an important area of research that can help reduce property damage, this approach overlooks infrequent, spatially dispersed, but more serious highway crashes that more often have fatalities. Hot spot analysis has therefore been less useful in generating system-wide improvements that could reduce severe crashes. This is particularly critical in rural areas, where run-off-road (ROR) crashes account for more than 50% of all automotive fatalities. Rural ROR crashes often involve high speeds, a single vehicle, and occur on two-lane highways [1]. However, cross-over, head-on accidents are also a risk on two-lane highways since median cable barriers being installed on four-lane divided highways are not feasible on these roads.

Additional research is therefore needed that uncovers specific design *elements*, rather than specific *locations*, that contribute to elevated accident rates, helping transportation engineers identify corrective measures for existing highways and develop new designs for future (re)construction. This approach is called “systems” analysis because the safety enhancements are done system-wide, rather than targeting the unique traits of specific locations (such as intersections) identified as hot spots. There is evidence that low-cost system-wide treatments, such as shoulder rumble strips, are very beneficial in reducing ROR crashes and fatality crashes, and similar treatments are sought to further reduce crash rates.

1.2 Background to Problem

While a precise value cannot be attached to human life or the life-long impacts of traffic injuries, one study [2] estimated that the total costs of auto crashes, including the lost quality of life, averaged 2.5% of gross national product and was over 5% in some countries. A federal study in 2000 calculated the overall economic cost of motor vehicle crashes at $230.6 billion in the United States, including future work loss costs (i.e., lost wages), travel delays, medical care, property damage, legal costs, and emergency services costs, among others [3]. That study also cited past research that sought to quantify a multitude of costs associated with automobile accidents. Another study calculated that expected injury costs averaged over $2.00 per hour travelled by vehicle [4].

Worldwide, the United Nations has declared 2011-2020 as the “Global Decade of Action for Road Safety”, citing annual road-related deaths of 1.3 million globally and economic consequences of accidents at between 1% and 3% of gross national product, or $500 billion [5]. Thus, reducing these costs is a global priority no matter what benchmark goals are set.

Due to the high costs of accidents, attention and funding focused on U.S. highway safety has increased notably over the past two decades. In 1996, numerous organizations involved with highway safety met to develop a comprehensive national Strategic Highway Safety Plan (SHSP). These organizations included the Association of American State Highway and Transportation Officials (AASHTO), the Federal Highway Administration (FHWA), the National Highway Traffic Safety Administration (NHTSA), and the Transportation Research Board (TRB). This meeting identified twenty-two specific focus areas grouped under six major headings: drivers, special users, vehicles, highways, emergency medical services (EMS), and management. These various areas of concern have been instrumental in guiding research and safety efforts as well as focusing attention on funding needs to address each area [6].

Subsequently, in 2005 the President signed into law the “Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users” (SAFETEA-LU), guaranteeing nearly a quarter of a trillion dollars for highways, highway safety, and public transportation. A mandate of SAFETEA-LU was that states develop individual SHSPs to guide investment decisions towards achieving significant reductions in highway fatalities and serious injuries. Oklahoma consequently developed a plan identifying its own state-specific goals, based on problems identified on all public roads, which was completed in September 2007. Among the goals of the state SHSP is a 20% reduction of both the 2004 fatality and serious injury rates by 2015 [7].

1.3 Scope of Study

Highway fatalities in most developed countries have been decreasing since reaching a peak in the early 1970s, but total fatalities tend to trend up and down over time, degrading the quality of statistical trend analyses that attempt to predict highway fatalities [8]. The United States exemplifies this trend, though there has been strong downward trend since 2005 [9] and a preliminary tally of 32,310 deaths in 2011 represents a decrease of over 10,000 fatalities compared to 2006. Despite this progress, a 2003 goal to reduce the U.S. highway fatality rate to below 1.00 per hundred million vehicle miles traveled (HMVMT) by the year 2008 was not achieved, as the rate was still 1.09 fatalities per HMVMT in 2011 [10].

Also, while there were over 500 fewer fatalities nationwide in 2011 than 2010 and eight NHTSA regions experienced declines, Region 9 (Arizona, California, and Hawaii) saw a 3.3% increase while Region 6 (Louisiana, Mississippi, New Mexico, Oklahoma, and Texas) had no change between 2010 and 2011 [10]. Thus, Oklahoma is in a region that lags the majority of the country with respect to crash fatality reductions and so analysis in this project is confined to fatality crashes recorded in the FARS database that occurred in the state between 1998 and 2011. In Oklahoma, annual reductions have been achieved recently; both fatalities and serious injuries in crashes have decreased by 1% and the number of crash fatalities per HMVMT dropped by 0.01. Even so, there were still 668 fatalities, over 36,500 injuries, and a 1.40 fatality rate per HMVMT in 2010 [11], the latter figure being notably higher than the national rate of 1.09 in 2011 (see Figure 1).

Rural areas are disproportionately represented in the FARS database. Nationally in 2007, 57% of traffic fatalities occurred on rural roads [12] even though only 23% of the U.S. population was rural. The fatality rate for rural roads was 2.21 per HMVMT compared to 0.88 in urban areas [13]. Crashes in rural areas typically involve higher speeds on highways that often lack paved shoulders, curbs, or other safety features [12]. Also, 60% of all drivers who died en route to hospitals in 2007 had crashed in rural areas [13]; longer wait times for the arrival of first responders and longer ambulance rides to local or regional hospitals are likely part of the reason, but more severe crashes probably also affect this statistic. Of note, Oklahoma's territory is largely rural but about two-thirds of its population is urban. FARS data for 2007 reveal that Oklahoma was one of twenty states where rural highways accounted for 70% or more of all highway deaths. That same year 69% of all fatal crashes took place on state highways [14]. Thus, both nationally and state-wide important reductions have occurred but there will always be room for improvement, and many goals remain unmet.

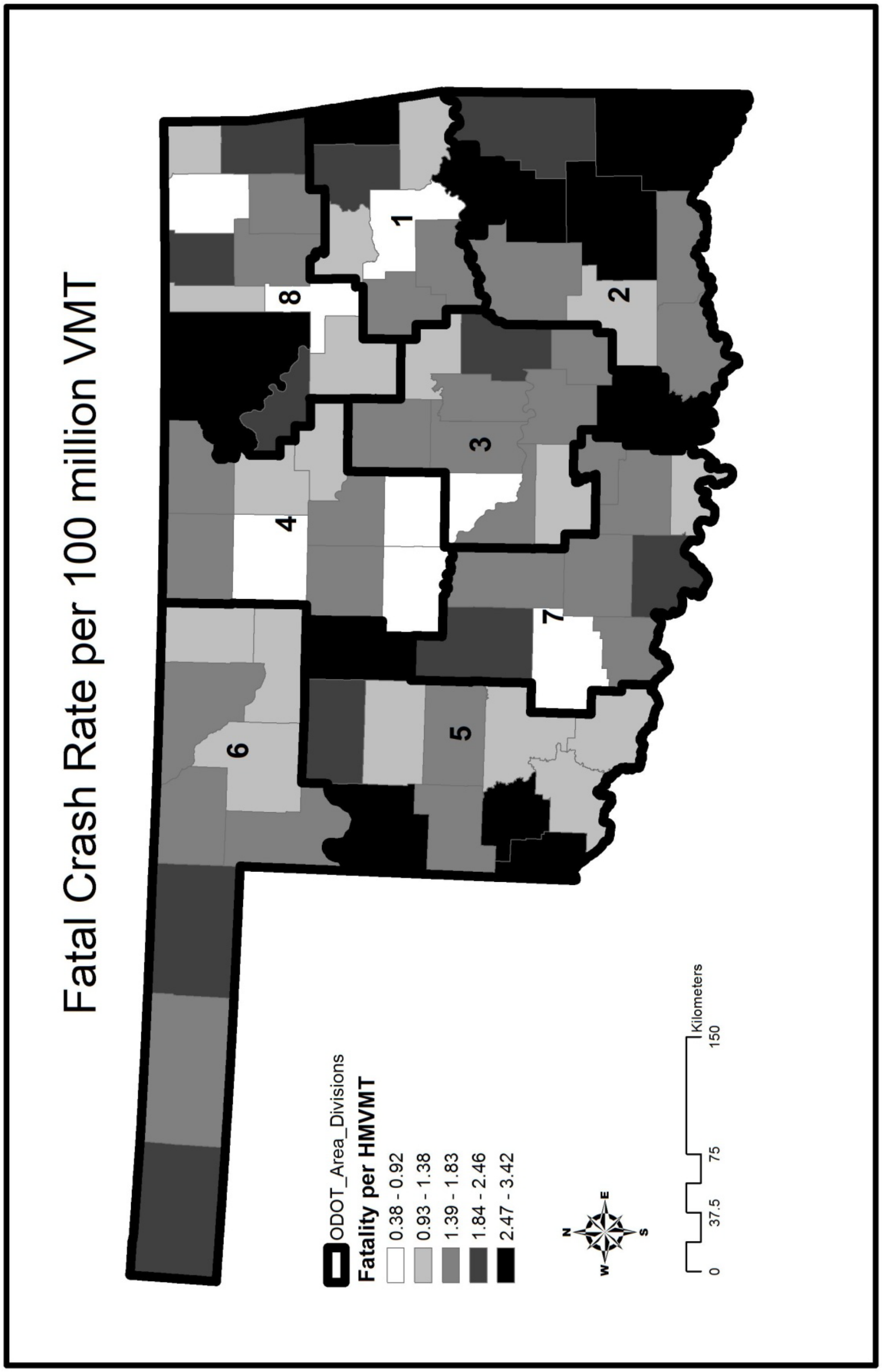

Figure 1. Fatality Rates per HMVMT by County with ODOT Field Divisions

1.4 Study Objectives

This project can provide state, county, and tribal transportation planners with information about the common characteristics of highways that contribute to fatal crashes. Once identified, said officials can make better cost-benefit decisions and identify low-cost solutions that can be applied to specific segments of highways or system-wide. The ultimate goal is to identify strategies for deploying relatively low-cost safety measures across the entire highway system. These measures can be better programmed if policy and decision-makers know the likely locations of severe or fatal highway crashes. All this will aid in cost-benefit analysis of competing safety treatments with a goal of improving highway safety in the most efficient manner.

As outlined in the project proposal, three deliverables were listed. First, an integrated geo-database with the state's highways linked to fatality crash data (obtained from FARS) and roadway characteristics (obtained from ODOT and others) would be constructed. This was a necessary, and very lengthy, component of the project given the incredibly diverse sources and formats of the needed data. The steps for this process are outlined in Section 3 later in this report. Ultimately, however, this database is not in a format that is publicly useful and is simply the platform within which all data were spatially cross-referenced for the purposes of analysis. Ultimately, statistical analysis had to be performed in a computer statistics package (SPSS) and so the pertinent variables from the geo-database were exported in a spreadsheet-type format for importation into the statistics package. Thus, there is no tangible manifestation of this deliverable, but the databases are available should interested parties request them. They are in ESRI's ArcGIS 10.0 format.

Second, a web page providing materials and information gathered and generated for this study is intended. This deliverable is in progress, though ultimately the nature of the project, especially the analysis phase, went in different directions than originally envisioned and very few maps were generated, for example.

Third and finally, the most important deliverable is this comprehensive final report that contains an extensive literature review of past work (Section 2), a description of the data and methods used in this project (Section 3), an analysis of the causes of highway fatality crashes on rural roadways in Oklahoma (Section 4), and a summary evaluation of the results (Section 5).

It is important to mention what is beyond the scope of this project. This project neither proposed nor included a cost analysis component. The project also does not involve technology transfer; though the methods and procedures are replicable by others with the appropriate expertise they are not novel in the sense of being newly developed nor are they patentable. As a result, these sections are absent from this report.

This research project also does not include information about the drivers themselves and the behaviors that can lead to crashes and fatalities [15,16]. Abundant research exists on driver behavior *(*e.g., drunk, distracted, or fatigued driving); the reader is directed to a large body of literature as exemplified by recent research on cell phone use in particular, but a sample of this work is provided for context especially in terms of the role of mobile devices as driver distractors. McCartt, Hellinga, and Bratiman [17] reviewed 125 studies on cell phones and driving, finding that cell phone use in autos has been increasing over time but that laws banning such activity have had little effect and were unevenly enforced, while more effective alternatives were unknown at that time and remain elusive still.

McCartt, Hellinga, Strouse, and Farmer [18] analyzed the effects of cell phone laws on driver usage patterns, finding that states with laws proscribing cell phone use while driving had sustainably lower usage rates but also that police citation rates were low. However, because the research was conducted with video evidence, they were unable to determine if the decrease was simply due to drivers switching to hands-free devices. Atchley, Hadlock, and Lane [19] evaluated the slow pace at which social and behavioral norms affect driver behavior, finding that drunken driving laws have heavily influenced the public's perception that this behavior is wrong but that similar attitudes about cell phone use while driving have yet to achieve the same impact, especially among younger drivers.

This research also does not focus on temporal trends [20, 21, 22] or weather-related incidents [23, 24, 25]. These are areas that could likewise inform transportation safety decision making and which have also been studied as noted by the representative citations given. However, it was the mandate of the topic generated by ODOT that roadway characteristics and especially geometrics be the focus of this project.

2.0 BACKGROUND

This section places the research problem in context and establishes a basis for the research conducted in this study. The first section below reviews initiatives, laws, mandates, and programs in the area of highway safety at the federal, state, and tribal levels, in particular issues pertinent to rural areas. The second section summarizes notable academic/applied research as published in peer-reviewed journals on the topic. The two sections necessarily overlap and intertwine, as the results of academic work have informed policy/practice and vice-versa. However, the first section largely contains a review of public initiatives as well as statistics on the nature of the problem, whereas the second section summarizes the methodological approaches and statistical results found in academic research.

2.1 Public Policy Initiatives and Studies

Since public roadways are almost by definition built by governmental entities, and furthermore most crucial datasets pertaining to crashes are similarly collected and maintained, it is important to first review the initiatives and statistics originating from these governmental agencies. These statistics and initiatives help provide context for the research undertaken in this project.

2.1.1 Federal Agencies

The current framework for addressing highway safety in the United States was established in 1997 with the publication of the Strategic Highway Safety Plan (SHSP), identifying twenty-two emphasis areas in six categories. One of the six categories is “Highways”, under which are listed the following emphasis areas relating to crashes: trees, run-off-road (ROR), horizontal curves, utility poles, unsignalized intersections, head-on collisions, head-on collisions on freeways, and work zones [6]. As part of this initiative, the National Cooperative Highway Research Program produced volumes for each emphasis area as part of its Report 500 series. Particularly pertinent to this research are several volumes that address roadway design elements, the physical properties of roads indicated earlier that are the focus of the research.

Volume 3 addresses collisions involving trees in hazardous areas, noting that trees are the most common object struck, comprising about 8% of all fatal crashes and about one-third of all crashes involving fixed objects. Furthermore, of all fatal tree crashes, 90% occur on two-lane roads and 77% occur in rural areas [26]. Another subset of fixed objects is utility poles, the subject of Volume 8, which notes that there are 88 million utility poles on highway rights of way in the United States, and in 2002 there were over 1,000 fatal crashes involving utility poles [27].

Both types of crashes above are a subset of the larger problem of ROR crashes, the topic of Volume 6. Reducing ROR fatality crashes involves a three-step objective: first, keeping vehicles from leaving the road; second, minimizing the likelihood of crashing into a fixed object or overturning once a vehicle travels off the shoulder; and third, reducing the severity of those crashes not prevented by steps one or two [1]. Leaving the roadway often results in one-vehicle crashes, but another subset is that of head-on collisions, addressed in Volume 4. Statistics from the late 1990s indicated that 75% of head-on crashes occurred on rural roads and also that 75% occurred on undivided two-lane roads (85% aggregate). Surprisingly, passing situations only accounted for slightly over 4% of head-on crashes. Equally counterintuitive, less than one-third of crashes occurred on curves [28]. In 2002, horizontal curves accounted for about 25% of fatal crashes, mostly on rural two-lane highways that are not part of state DOT systems. Of the curve-related crashes, around 76% were single vehicle, ROR crashes striking fixed objects, and about 11% were head-on crashes. With an estimated 10 million horizontal curves in the United States on two-lane highways and few highway agencies linking highway geometrics to accident data [29], this could be an area of significant contribution of this research.

Given these statistics and emphasis areas, a key document for transportation planners is Volume 21, which addresses safety data and analysis. This guide reviews state and local crash data sources, FARS, state inventory data, and also discusses the role of Geographic Information Systems (GIS) [30], and was thus a critical resource for this research. Of note, Section IV of this report (Roadway Segment Programs) discusses two different approaches to targeting areas for safety improvements, spot (aka “black spot” or “high-crash location”) analysis and systematic approaches. Though black spot approaches have historically been more common, the FHWA has pushed for inclusion (if not a focus) of system approaches in SHSPs. More recently, the TRB has published a research results digest (#345) focusing on black spot and systematic methods based on a survey in which 25 out of 50 state traffic safety engineers responded (Oklahoma did not). The results indicate that most states still target safety funds at HCLs [14], but many states are shifting some of their funding towards systematic approaches. Conceptual and methodological differences in these two approaches, and the results found by applying both methods, will be reviewed later in the academic and applied research section (Section 2.2).

In 2005, the SAFETEA-LU legislation provided over $244 billion of guaranteed funding for highways, highway safety, and public transportation, with $5.1 billion allocated between 2006 and 2009 for states to address their needs. Of the roughly $1.2 billion annual allotment, $90 million was to be set aside annually for the High Risk Rural Roads Program (HRRRP) as part of the Highway Safety Improvement Program (HSIP) [31]. Through four years of the HRRRP (as of 2009), states had under-requested funds based on FHWA expectations, raising concerns that there were impediments to implementing the program. A study revealed that the most common problems at the state level were: collecting crash data on locally-owned roads; determining the criteria for selecting the best projects, soliciting proposals, and choosing projects to fund; coordinating with federal, state, and local agencies; and working within state law to administer HRRRP funds [12]. Thus, a myriad of challenges exists to improve rural highways and their crash fatality rates, even with secure sources of funding, so a review of state-level efforts is given next.

2.1.2 State Agencies

The HSIP required all states to submit SHSPs by October 1, 2007 in order to make full use of these funds; Oklahoma published its SHSP in September, 2007 [7]. Motivating and informing Oklahoma's SHSP was the fact that traffic fatalities had increased 14.1% from 2000 to 2005 and the fatality rate per HMVMT had risen to 1.71, compared to a U.S. rate that had declined to 1.45. A large number of stakeholder and constituencies were involved in the development of the state SHSP [7] which aspired to a vision of "…zero deaths, zero injuries" but more practical goals of achieving 20% reductions in both fatalities and serious injuries per HMVMT, by 2015. The participants reached consensus on four emphasis areas for the state: unsafe driver behavior, intersection crashes, crashes involving young drivers, and lane departure crashes [7]. While the intersections emphasis area relates to this research since it might reveal that rural highway intersections are a problem area in the state, the lane departure emphasis area very directly impacts this research.

Following the strategies listed in Report 500: Volume 6 [1], the state SHSP proposes the same three-step objective function listed earlier for preventing lane departure crashes, and state initiatives seem to be working. After experiencing 803 roadway fatalities in 2005, Oklahoma deaths dropped steadily to 766 in 2007 and to 668 in 2010, well below the target goal of 726 set for 2011. Likewise, serious injuries fell from 17,663 in 2007 to 16,077 in 2009, though they increased to 16,557 in 2010 [11].

However, Oklahoma still lags national averages in several areas. Fatalities for drivers with blood-alcohol content (BAC) of 0.08% or more have slightly risen, motorcycle fatalities have risen, and 2008 saw a modest rise in unrestrained passenger fatalities [32]. These factors are less controllable by state (or any other) agencies as they involve driver and passenger choices, but nonetheless said agencies have education, awareness, and safety programs in place to encourage safer driving behaviors. However, in part motivated by the research emphasis promulgated by ODOT for this research effort, and in part because roadway design is something that ODOT and others can directly affect, the research conducted here focuses on physical characteristics of the roadways and their correlations with various types of crashes.

Because Oklahoma possesses unique traits, a tailored approach to this research is necessary. Oklahoma's territory is largely rural, but 65% of the population in the 2000 Census was classified as urban. However, based on 2007 FARS data, Oklahoma was one of twenty states in which 70% or more of highway deaths were on rural roads. Also that year, 69% of fatal crashes took place on the state highway system [14]. Hence, rural, state-roadway deaths remain a chronic problem in the state, and the state has a large territory to manage. Oklahoma's roadway system consists of 113,147 total miles. In 2000 mileages were calculated as: 669 miles of interstate highway, 559 miles of turnpike, 11,601 miles of state and federal highways, 284 miles of state park roads, 86,665 miles of county roads, and 12,865 miles of local city streets [32].

Another trait that distinguishes Oklahoma from many other states is the strong presence of Native American population and tribal governing entities. Oklahoma ranks second in the nation (behind Alaska) in its citizens declaring Native American ancestry in the Census, and 38 different tribes administer federal and tribal lands through the Bureau of Indian Affairs (BIA) or Tribal Governments. Overall, including National Park, U.S. Forest, and U.S. Fish and Wildlife Service Units, about 4.5% of the total state area is federal land acreage, including 9,675 Federal Lands Highway Program (FLHP) road miles [33]. Both nationally and statewide, these tribal areas likewise adopted SAFETEA-LU-related safety plans and focus areas.

2.1.3 Tribal Agencies

While traffic fatalities in the United States declined by 2.2% between 1975 and 2002, they increased by 52.5% on reservations [34], with a further caveat that "underreporting is highly likely and that the statistics may be considerably worse" [35], a finding confirmed by Bailey and Huft [36]. While Oklahoma is not listed as one of the top five states (Arizona, Montana, South Dakota, New Mexico, and Idaho) that account for over 80% of Indian reservation fatalities, this could be due to different definitions of "reservation" in the various states and the fact that Oklahoma's situation is unique regarding tribal jurisdictions and recognition. Nonetheless, the problems and strategies outlines for reservations nationwide are applicable here. The traits cited in government statistics and academic research indicates that occupant restraint levels are lower, alcohol/drug-impaired driving levels are higher, and unsafe driving is higher for Native Americans than the national average [15, 35].

Numerous emphasis areas are highlighted in the Indian Lands SHSP, but two that directly relate to this research are the need for better data collection and the prevention of ROR crashes [35], which is clearly a cross-cutting problem and safety strategy at all levels. In Oklahoma, a Tribal Transportation Safety Summit was held in April 2010 to address state-level concerns and strategies. This summit worked towards crafting a process to facilitate collaboration among tribal, local, state, and federal entities with a goal of solving transportation safety issues. Within the "4-E" approach (engineering, education, emergency response, and enforcement), engineering challenges included maximizing data collection and developing a data sharing system. This research integrates well into these areas.

In summary, many federal, state, and tribal initiatives are in place to address numerous dimensions of safety improvement on the nation's roadways, with an overall goal of reducing fatalities and serious injuries by ameliorating the many factors that cause automobile (and motorcycle) crashes in the United States. The initiatives and statistics reviewed above, however, only tell half the story – what the problem is and how it is being addressed. Informing these policies and "best practices" is a large body of academic research into the root causes, statistical relations, and best methods for uncovering correlations between crash types and roadway features, predicting likely future crash locations based on roadway design, and making recommendations for design and engineering solutions. Pertinent literature in these areas is reviewed next.

2.2 Academic and Applied Research

2.2.1 Data Limitations

Out the outset, collecting, obtaining, aggregating, geocoding, and even defining the data needed for this type of research occupies most transportation researchers in their pursuit of specific research questions. Ultimately, this research seeks to identify hazardous locations on Oklahoma's rural roadways, so identifying similar research and understanding Oklahoma-specific limitations will be an important area of background research. A recent study of eight European countries documented state-of-the-art approaches to the identification and management of hazardous locations. Notable features of current methods includes: applying a sliding window approach (similar to moving averages in time series analysis), use of the empirical Bayes (EB) method for generating estimates of expected accidents at sites, and use of an appropriate period of data for accident estimation [37]; between three to five years is an oft-cited figure [38, 39, 40]. As noted earlier, an unfortunate trait of Indian lands in the U.S. is underreporting of crash and fatality data. Even so, in South Dakota from 2001 to 2005 Native Americans accounted for 26% of traffic fatalities (in FARS) in the state, and a study estimated that in one county (Shannon) reported crashes represented, at best, only about half of the real number of crashes that occurred in 2003 [36]. Though the research described here cannot encompass data quality or quantity improvement techniques, it is critical to understand the difficulties associated with research of this type and that no standard practices exist. This research has undertaken the collection and use of data in ways that seemed best suited to the overall research goals and are described later in Section 3.

2.2.2 Methodological Approaches

High risk location identification methods (i.e., black spot, hot spot, high-crash location analysis) remain firmly entrenched in the literature, largely because of past policy and the relatively easier identification of such locations with existing data. For a long time in the United States, this was the *modus operandi* because of a focus on reducing the overall number of crashes of all types, under the two assumptions that (1) there were few differences in the factors contributing to the various types of crashes, and (2) if total crashes were reduced, proportional reductions in fatalities and serious injuries would follow. While true through the 1980s, eventually the drop in fatalities leveled out at about 42,000 per year, though the fatality rate per HMVMT did continue to drop (TRB 2010). Black spot analysis has not been effective in reducing serious crashes and fatalities in rural areas where a majority of fatalities occur. As a result, the national SHSP sought to refocus efforts on higher-speed, rural, single-vehicle crashes, as those tend to be more severe and account for a disproportionate share of fatalities. Also, a majority of fatalities occur on local, not state, roads [6]. These statistics have focused attention on the need for system-wide methods for identifying locations at risk for severe crashes where traditional methods of calculating crash rates cannot distinguish safe from dangerous locations [14].

Spatial analysis techniques have also been applied to the problem of black spot analysis. Songchitruksa and Zeng [41] applied spatial autocorrelation statistics black spot analysis in Houston, demonstrating the need for additional care in conducting spatial analysis on crash location data. In a similar vein, Flahaut, Mouchart, San Martin, and Thomas [42] compared two different spatial autocorrelation measures, a kernel method and local Moran's I_i, to study black spots, finding merits for each method depending on the local setting and safety needs. Flahaut et al. were motivated by an earlier study by Thomas [43] that examined the influence of road segment *length* on statistical results. This was not a spatial analysis approach *per se*, but represented early recognition that explicitly geographic traits of the crash analysis framework are of critical importance, and that conclusions made at one level of spatial aggregation may not hold at other levels. Also related to spatial scale, Quddus [44] and Wang, Quddus, and Ison [45] examined crash data for British census wards in London and across all of England to determine the various effects of infrastructure, traffic flow, speed, and road curvature on crash frequencies. Furthermore, Quddus used a global spatial autocorrelation statistic to account for spatial dependence of the various influences on crashes [44]. Erdogan [46] studied road mortality in Turkey using geographically-weighted regression (GWR) to account for spatial non-stationarity in crash data, while Aguero-Valverde and Jovanis [47] found no spatial correlation in fatality data but did find it present in injury crash data. Hence, spatial analysis methods are improving research in this area.

The overriding issue when studying rural areas is statistical; crashes are relatively infrequent but more often severe, so innovative methods are needed that work with sparse data to tease out subtle differences in crash locations. Research methods must contend with the fact that many highway segments will have no crashes or fatalities during the study period, skewing the data and biasing many statistical methods. Cafiso, Di Graziano, Di Silvestro, La Cava, and Persaud [48] addressed this topic by studying different ways of partitioning highways into usable analysis “segments”, finding that smaller homogeneous segments of between one-half and five kilometers in length provide for better analysis of rural crash data vis-à-vis exposure and geometry variables. Using a “left-censored” dependent variable in a Tobit regression with data for Indiana interstates, Anastasopoulos, Tarko, and Mannering [38] found a variety of significant influences on crash rates relating to pavement conditions, highway geometrics, and annual average daily traffic (AADT). In the southeast United States, Zhu, Dixon, Washington, and Jared [49] found that lane width, horizontal curvature, and lighting were the only variables significantly and consistently associated with single-vehicle fatal crashes on two-lane rural highways. Similarly, Deng, Ivan, and Gårder [50] studied head-on collisions on rural, two-lane Connecticut highways and identified pavement width as the most consistent factor influencing the severity of such crashes. In Ohio, Schneider, Savolainen, and Moore [51] identified horizontal curvature as a significant influence on rural motorcycle crashes on two-lane highways, along with shoulder width and AADT. Likewise, Karlaftis and Golias [52] found differences between two-lane and multi-lane roads in Indiana; overall, geometric variables and pavement condition variables most significantly affected accident rates.

A large body of literature has analyzed crash rates and crash locations [53, 54, 55]. As can be inferred from the discussion above, there are many potential contributing factors in automobile crashes. As a result, most studies use some form of multiple regression to relate crash characteristics to factors such as weather, gradient, and intersection characteristics [56, 57, 58]. Standard linear regression, however, suffers from certain limitations. Linear regression relies on the assumptions of the normal distribution and lacks the “distributional property necessary to describe adequately the random and discrete” nature of crash events [59, p. 471]. Specifically, the dependent variable is usually a discrete, non-negative integer (crashes, fatalities, etc.) while regression models predict continuous values of the dependent variable that can be negative. Crashes, however, are random events that occur independently over time and are usually positively skewed. Consequently, Poisson and negative binomial (NB) regression models are frequently used to model and predict the relationship between road characteristics and crash frequencies or severity [52, 59, 60, 61].

The Poisson model is well-suited to modeling crash data because the number of crashes in a given space-time region can be considered as a random variable with probabilities that are Poisson-distributed [60]. A standard formulation of a Poisson regression model is:

$$P(y_i) = \frac{EXP(-\lambda_i)\lambda_i^{y_i}}{y_i!}$$

where $P(y_i)$ represents the probability of roadway segment i experiencing y_i crashes per time period (year) and λ_i is the Poisson parameter for segment i (equal to the segment's expected number of crashes per year). The Poisson parameter λ_i is typically estimated from a linear regression in which any number of explanatory (independent) variables (X_i) that represent highway traits are used:

$$\lambda_i = EXP(\beta X_i)$$

The slope coefficients β from this regression (in exponential form) provide important information on both the direction and influence of each independent variable X_i on the number of crashes [62]. Negative slope values for β indicate a variable lowers the risk of crashes relative to other variables and positive values indicate the variable raises the risk.

The Poisson model has been used in many crash studies. Miaou [59], Miaou, Hu, Wright, Rathi, and Davis [63], and Miaou and Lum [64] employed Poisson regression to estimate truck crash rates using traffic and geometric characteristics of roads and model relationship between truck accidents and geometric designs of roads. Both Miaou [59] and Miaou et al. [63] found that AADT per lane, curvature, and mean absolute grade or vertical alignment significantly correlated with truck crash rates. Saccomanno, Grossi, Greco, and Mehmood [65] developed a Poisson model to estimate expected crash frequency along homogeneous segments of highway sections in southern Italy using crash and road geometric data from 1993 to 1999. Since the AADT was uniform for entire road sections, length of road segments was used to measure crash exposure and the study found that the length of the section, number of private driveways, number of major intersections, and the change in 85th percentile speed from the previous road section were significant determinants of crash frequency.

One of the major criticisms of the Poisson regression model is that it assumes the variance and mean of the data are equal (isodispersion); overdispersion means the variance exceeds the mean while underdispersion means the opposite. Many studies have found the assumption of isodispersion does not hold, as crash data especially tend to be overdispersed [52, 61, 62]. The negative binomial (NB) regression model overcomes the problem of overdispersion by allowing the variance to exceed the mean.

The negative binomial (NB) model is a variant of the Poisson model and assumes that the Poisson parameter follows a gamma probability distribution; as a result the NB model is also known as the Poisson-gamma model and overdispersion in crash data do not negatively affect model validity [61]. Here the Poisson parameter is found by:

$$\lambda_i = EXP(\beta X_i + \varepsilon_i)$$

where $EXP(\varepsilon_i)$ is a gamma-distributed error term that has a mean of 1 and a variance of α. This term, known as the overdispersion parameter, allows the variance to differ from the mean [62].

Hadi, Aruldhas, Chow, and Wattleworth [66] employed NB regression to estimate crash rates on various types of rural and urban highways of different traffic levels. Results suggested that higher crash frequencies were influenced by AADT and intersections and wider road shoulders and lanes were effective crash reduction measures. Shankar, Mannering, and Barfield [67] used NB regression to study how roadway geometrics and environmental factors influence crash frequency. They found significant correlations between fewer (well-spaced) numbers of curves per mile of road section and an increased number of severe overturning vehicle crashes. Brown, Labi, Tarko, and Fricker [68] obtained and used crash and road inventory data for Indiana from 1991 to 1995 to develop crash prediction models for crash rates on road segments based on geometric and access control characteristics, using NB regression to develop crash prediction models for all crashes, for property-damage only crashes, and for fatal and injury crashes. Results of the study indicated that the increased access density and proportion of signalized access led to an increase in the number of crashes on roadway segments, whereas lower crash frequencies were associated with the presence of outside shoulders on roadways, two-way left-turn lanes, and medians without openings between signals [60].

Generally, the Poisson model is initially used to analyze the relationship between crashes and roadway characteristics. The NB regression model and other variants are then explored when over- or underdispersion is detected in the model following the Poisson regression analysis [59]. However, Berk and MacDonald [69] argued that poor model performance is just as likely to result from omitted variables or from an incorrect systematic part of the Poisson regression (as opposed to an incorrect stochastic part of the model, i.e., over- or underdispersion). Other limitations of Poisson and NB regression models have been identified, including their sensitivity to outliers, inadequate handling of missing data, inability to deal well with multicollinearity in independent variables, and the fact that they are both parametric procedures that require the functional form of the models to be known in advance [52]. Nonetheless, both techniques are still widely used to model relationships between crashes and the geometric and environmental factors that may influence them, while numerous alternatives also exist as summarized by Lord and Mannering [62].

2.2.3 Engineering, Design, and Management Studies

In addition to identifying the traits of existing roadways that contribute to severe crashes, several recent studies evaluated the impacts of specific treatments. Using data from Arkansas, California, Illinois, and North Carolina, Lyon, Persaud, Lefler, Carter, and Eccles [70] employed the EB method to study the effects of installing two-way left-turn lanes (TWLTL) on two-lane roads and found a significant reduction in crashes, further finding that TWLTLs were more effective in rural areas. In Oklahoma and Texas, Eisele and Frawley [71] studied test corridors where TWLTLs were replaced by raised medians and determined that the increased inconvenience caused by longer travel times was offset by increased safety and fewer conflict points. Though conducted in Spain, Pérez [72] used the EB method and found that upgrading highways (via a variety of engineering and safety treatments) significantly improved safety while improving traffic signage, repainting road markings, and repaving highway surfaces did not.

Though tangential to this study, management and enforcement policy studies can provide useful insights. Malyshkina and Mannering [73] studied the issue of roadway design from the perspective of whether approved design exceptions had an impact on the frequency or severity of crashes. While overall design exceptions were not found to contribute to significantly different levels of traffic safety, they cautioned that notable differences existed between rural and urban areas and also that horizontal curvature and pavement roughness (as measured by IRI) was very critical and deserved greater scrutiny when design exceptions are requested. Using Ohio data for 1973 to 2000, Welki and Zlatoper [74] found that drunk-driving enforcement, higher speed limits, and rural driving all significantly increased fatalities.

Given the higher level of alcohol-related fatalities cited earlier in both rural locations and on Indian lands and the higher speed limits on rural roads, these results can reinforce and inform current efforts by ODOT, OHSO, and others in coordinating their efforts to tackle the highway fatality problem in the state. In a similar study for Ohio using 1975 to 2000 data that disaggregated six types of vehicular fatalities, Welki and Zlatoper [75] found that the positive effects of speed limit enforcement and drunk-driving arrests cited in their 2007 paper did not affect all socioeconomic groups equally, motorcyclists in particular. With a 52% increase in motorcycle registrations in Oklahoma between 2005 and 2009 [32], this is a potentially overlooked area needing attention as motorcycle fatalities have been on the rise in Oklahoma in this same period [11].

3.0 DATA

3.1 Basic Datasets

This research integrates fatality data and roadway data. Fatality data were obtained from the federal FARS dataset (http://www.nhtsa.gov/FARS) from 1998 to 2011, including 5,635 crashes resulting in 6,667 total fatalities in Oklahoma during that period. The FARS database includes dozens of fields relating to the vehicles involved, crash severity, location, weather, driver and passenger demographics, etc. Crashes are also geocoded with latitude and longitude as well as highway segment identifiers (e.g., Interstate 35, U.S. Highway 177).

Highway information obtained from ODOT included road condition data and road inventory data. The road condition dataset has information from sensor data measuring roughness, rutting, and faulting, as well as observed distress information like cracks and potholes. The road condition data are collected on a two-year cycle and are complete from 2001 to 2010. Road inventory features are of three types: design, geometry, and other. Design features include number of lanes and the type and width of the road surface, shoulders, and medians. Geometric features include up/down and left/right grade, and curve radius. Other features include Annual Average Daily Traffic (AADT), elevation, and the type of terrain in the surrounding area.

3.2 Data Integration

Road geometry and condition data are generally reported for road segments (control sections), a format that requires conversion into a functional road network topology. The latitude and longitude of each fatal crash is matched to the highway control segment where that crash occurred so each crash can be assigned the design, geometry, and condition attributes of the control section. In addition to having to assume consistent condition and design traits along each control segment, this also places limitations on the type of analysis possible. The observations are discrete points with various highway attributes matched to those points, and the dependent variable is a discrete tally of fatalities at each location where fatalities occurred. As road traits at a location can change over time, each crash record is treated as an independent event to distinguish it from other fatal crashes on the same control segment but at a different time.

The following provides a concise list of specific steps undertaken to assemble the geo-database:

1. Download the GIS shapefile of crashes (from 1998 to 2011) from the Oklahoma SAFE-T website using the query (Crash98_11).

2. Create a personal geo-database and save the shapefiles: Oklahoma crashes, control sections, highways, counties, and cities. Note: the Feature Dataset is projected to Albers_USGS Version.
3. Open ArcMap document and add the crash and fatal crashes feature classes along with the Oklahoma highway feature class.
4. Create a new field in both fatal crashes and highways feature classes. New field is control sections.
5. Right-click on the fatal crash layer > Joins & Relates > Join (a Join Data window opens). In the first drop down box, click the drop down arrow and select “Join attributes from table.”

    This join operation joins the highways layer to the fatal crash layer based on control section field. Important attributes of both layers are joined together to allow for easier query of the database.

6. Right click of fatal crash layer [joined with highways], navigate to data, then export, choose personal database, name [fatalities_final], save.

    The new join output (point) layer now has the important fields of the highway layer combined with that of the fatal crash layer. Now all crash records contain information on the road sections where they occurred.

7. Load the new layer [fatalities_final] in ArcMap.
8. Open attribute table and exported it as DBF to the personal geo-database.
9. Open DBF in Excel and edit out repeated fields (e.g., surface type/width, median type/width) which were in both highways and fatal crash layers. Also at this stage, edit out fields that are not used (e.g., mileposts, cities, counties).
10. Save file as an Excel file [Complete_Dataset_Final_original].
11. Import Excel file into SPSS for statistical analysis.

The aggregation of road inventory and road condition databases as described above results in dozens of potential independent variables. However, many variables are discarded, mostly due to low recording incidence in the datasets. Also, the geometry variables are converted from ratio to categorical due to the way those variables are measured. Table 1 summarizes the dataset by class of variable, general inventory characteristic, data type, and the numbers of subdivisions for categorical variables. While the independent variables include a mix of categorical and ratio data types, the dependent variable is the number of fatalities at a fatal crash location/section of highway, so advanced analysis methods like Poisson regression are needed to correlate crashes to highway conditions.

Table 1. Independent Variable Categories

Class	Characteristic	Type	Number
Condition	Roughness (IRI)	Ratio	1
Condition	Rutting	Ratio	1
Design	In/outside shoulder type	Categorical	7
Design	In/outside shoulder width	Ratio	1
Design	Intersection type	Categorical	4
Design	Median type	Categorical	7
Design	Median width	Ratio	1
Design	Number of lanes	Ratio	1
Design	Right-of-Way (ROW) width	Ratio	1
Design	Surface thickness	Categorical	9
Design	Surface type	Categorical	3
Design	Surface width	Ratio	1
Geometry	Curve radius	Categorical	3
Geometry	Vertical grade	Categorical	3
Geometry	Horizontal grade	Categorical	3
Other	AADT	Ratio	1
Other	Elevation	Ratio	1
Other	Terrain type	Categorical	7

3.3 Descriptive Statistics

Categorical variable distributions are given in Tables 2-5, while descriptive statistics for select ratio variables are given in Table 6. Categorical shares do not sum to 100% in some cases due to rounding and in other cases due to the omission of some categories.

Table 2. Distributions of Geometry Variables (Categories)

Geometry	0	1	2
Curve radius [a]	73%	14%	14%
Vertical grade [b]	3%	42%	54%
Horizontal grade [c]	1%	86%	13%

a. Curve radius: 0 = straight 1 = left curve 2 = right curve
b. Vertical grade: 0 = flat 1 = uphill 2 = downhill
c. Horizontal grade: 0 = level 1 = left edge higher 2 = right edge higher

Table 3. Distribution of Median Type

Median type	Undivided	Open w/ shoulders	Open w/ curb	Cont. left turn lane	Concrete barrier	Cable barrier
Share	54%	34%	1%	3%	5%	2%

Table 4. Distribution of Terrain Type

Terrain type	Flat (rural)	Rolling (rural)	Mountain (rural)	CBD 1[a] (urban)	CBD 2[b] (urban)	CBD 3[c] (urban)	Residential (urban)
Share	28%	40%	2%	0%	3%	18%	8%

a. CBD 1 = central business district

b. CBD 2 = fringe of CBD

c. CBD3 = outlying business district

Table 5. Distribution of Pavement Type

Pavement type*	Armor coat	Asphalt concrete 'B'	Asphalt concrete 'C'	P.C. concrete	All other types
Share	11%	41%	15%	18%	15%

* Pavement type definitions from ODOT *Road Inventory Manual*, 7th Edition [76].

Table 6. Descriptive Statistics for Ratio Variables

Variable	Minimum	Mean	Maximum	St. dev.	Records	Units
IRI	25	108.2	600	64.1	5338	none
Rutting	0.00	0.18	1.02	0.09	5558	inches
Shoulder width	0	7.3	19	3.2	5593	feet
Median width	0	18.4	99	25.7	5616	feet
Surface width	18	28.5	96	10.6	5616	feet
R.O.W. width	33	187.2	460	93.0	4488	feet
Lanes	2	3.1	8	1.3	5616	count
Base thickness	0	6.2	19	3.2	5588	inches
AADT	50	15,804	165,200	23,033	5616	count
Elevation	300	933	4,744	461.1	5490	feet

4.0 ANALYSIS AND RESULTS

Analysis was conducted in three stages. Initially, a trial analysis was conducted for thirteen counties in northeastern Oklahoma (Figure 2) as a means of developing experience with Poisson and NB regression as well as to study the portion of the state that has the highest percentages of Native American citizenry. While 12.9% of Oklahoma's total 2010 population of nearly 3.8 million claimed some type of Native American ancestry, collectively these thirteen counties are 26.0% Native American with Adair County being the highest in the state at 53.4% [77]. These counties are primarily rural and cover the Native American tribal territories of the Cherokee, Eastern Shawnee, Miami, Miami-Peoria, Modoc, Peoria, Ottawa, Quapaw, Seneca-Cayuga, and Wyandotte nations.

Subsequently, a Poisson regression model was run using data for the entire state. Explorations of both Poisson and NB regression models revealed that (1) very few differences resulted between both models when run for the same dataset, and (2) very few variables of interest exhibited either over- or underdispersion. Thus, only Poisson regression is discussed in Section 4.2. Finally, Poisson models were run for each of ODOT's eight field divisions. Though not requested or mandated, this seemed like a logical means of disaggregating the state. Any regionalization scheme would have sufficed, but since ODOT is the most likely agency to act on any findings, region-specific outcomes have the highest potential of being useful to ODOT. These results are discussed in Section 4.3 along with comparisons to the state-wide model reviewed in Section 4.2 as the independent variable set was the same for the state and field division models.

On a methodological note, both Poisson and NB models can be run unweighted or weighted. Weighted models were used here because of the temporal and spatial nature of the data. Many highway sections during the study period (1998-2011) have undergone deterioration, repair, and/or redesign, so the road inventory and design attributes could have changed over time. To account for changing road conditions over time, each crash was treated as an individual record and the roadway characteristics at the time of the crash were connected to that crash record. While this approach was necessary to account for changing road conditions, treating each crash as an independent event made it difficult to identify control segments with higher fatality crash rates. Therefore, the number of crash fatalities at a location served as a weight so that "dangerous" locations (i.e., those with more deaths) were more heavily weighted that those with fewer deaths.

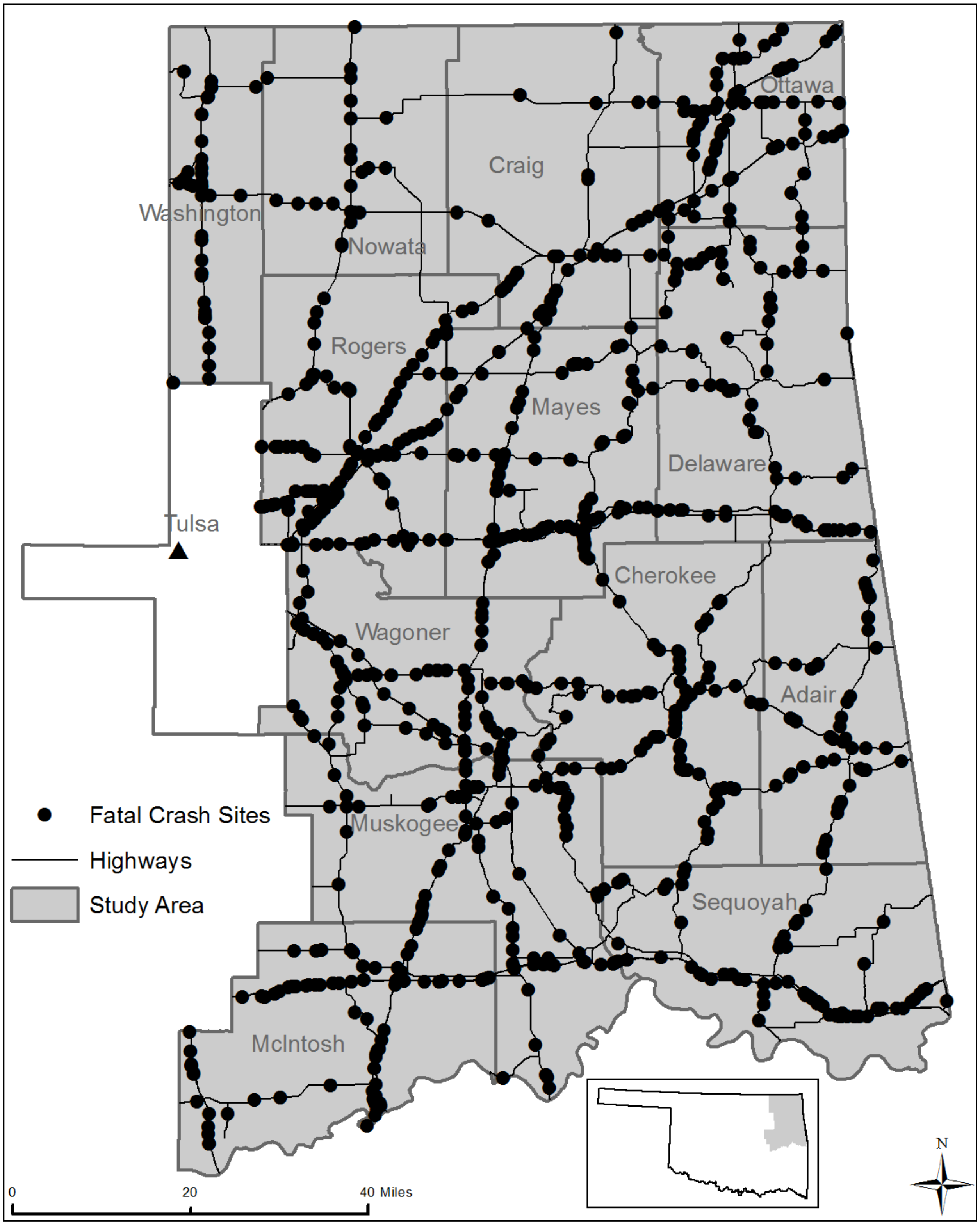

Figure 2. Northeastern Oklahoma Study Area

4.1 Northeastern Oklahoma Poisson and NB Models

As demonstrated in the literature [62, 69], Poisson regression is the logical starting point for regressions involving a dependent variable that is a non-negative integer (here, fatality counts). Due to the possibility of overdispersion in the data as well as for comparative purposes, an NB regression model is also constructed.

The Poisson regression model fits the crash data well (Table 7). The model has a standardized deviance (deviance/df) of 0.419, easily meeting the rule of thumb that the size of the raw deviance (524.96) should not be larger than twice the number of degrees of freedom (1,252). Furthermore, the omnibus test of goodness-of-fit, the chi-square test of the likelihood ratio, is significant to at least the $\alpha = 0.001$ level. This test compares the fitted model against an intercept-only model, which tests the null hypothesis of no effect of the independent variables.

Table 7. Poisson and NB Regression Results, Northeastern Oklahoma

Model parameters	Poisson	NB
Deviance/df	0.419	0.146
AICc	3,471	4,289
Chi-squared/p-value	170/0.000	43.9/0.000
Observations (n)	1,284	1,277

Descriptive statistics were examined for evidence of over- or underdispersion. The lack of isodispersion was not chronic, but when it existed it tended towards overdispersion. Since no underdispersed variables were factors in the Poisson model (e.g., number of lanes, pavement rutting), the NB model is pursued next as a hedge against violations of the Poisson model's assumptions as well as for comparative purposes.

Results for the NB regression model are also shown in Table 7 and the determination of which model is "better" depends on which model fit statistics are deemed most important. The Poisson model has a lower Akaike Information Criterion (AICc) value, indicating less information loss in the Poisson regression model, and the Poisson omnibus test has a smaller p-value, though both models are significant to at least the $\alpha = 0.001$ level. However, the NB model produces a smaller standardized deviance, which is perhaps a function of the optimization goals of that method. Ultimately, both models are significant, have a very good fit, and meet the basic assumptions of the models. This outcome provides confidence that either approach is sufficient for analysis.

Numerous independent variables are significant predictors of fatality crashes in both models (Table 8), though individual components of categorical variables (e.g., surface, intersection, and terrain types) have varying levels of significance for their parameter estimates. The surface type results are somewhat misleading, because despite a wide variety of possible pavement surfaces based on standard coding, just three main types dominate the highway surfaces in northeast Oklahoma. This result is largely a reflection of the fact that nearly all highway fatalities in the study area/period occurred on these three surface types (asphalt concrete types C (S5), B (S4), and Portland Cement Concrete) [76].

Table 8. Variable Model Effects Significance, Poisson and NB Models

Model Variables	Poisson	NB
Intersection type	0.000	0.118
AADT	0.003	0.000
Roughness (IRI)	0.003	0.002
Median Type	0.004	0.000
Surface Thickness	0.008	0.009
Elevation	0.013	0.082
Surface Type	0.016	0.011
Median Width	0.018	0.000
Radius	0.047	0.069
Terrain Type	0.075	0.000
Grade	0.525	0.973

Unsurprisingly, AADT, pavement roughness (IRI), and several design characteristics significantly correlate to locations of highway crashes resulting in fatalities. Median width is inversely related to fatality counts, which is logical as wider medians will provide more run off room and lessen the chances of a head-on collision. Among the various median types, “open with a combination of curbs and shoulders” (type 3) has the highest impact on fatality counts while “open with curbs” (type 2) has the only inverse relationship with fatality counts. Surface thickness is a variable that is harder to directly correlate to crash rates, though most individual thickness levels are not as significant as the overall effect of this trait in the model (thickness depths are reported in one inch (ordinal) ranges rather than as continuously measured depths). Radius barely meets the standard $\alpha = 0.05$ significance threshold while grade proves unrelated. The appearance of elevation as a fairly significant variable in both models is curious, as elevation alone should not directly impact driving safety as related to the risk of a fatal accident. It is likely that this variable is indirectly measuring other influences (perhaps regional/cultural differences across the state) on fatality crashes relating to location in the study area, given that grade is not significant (i.e., if elevation and grade were both significant one might expect a lot of fatalities on steep, hilly highways).

Most variables achieve approximately the same significance values across the two models, which is not surprising given that both methods “are responding to the same expected mean function” [69, emphasis in original]. While moderate differences exist, especially for individual variable parameters and significances, overall there is little substantive difference between the two models. While this result could be considered to render the NB model redundant and therefore unnecessary, the consistency of the two models increases the confidence that the highway traits found to significantly predict fatality crashes in northeastern Oklahoma are useful areas for traffic planners and engineers to evaluate and develop safety treatments for system-wide application.

4.2 State-Wide Poisson Model

Results for the entire state highway network reveal a handful of significant predictors of fatal crash sites. Table 9 presents the results for the state-wide model, with seven variables achieving 0.05 significance (α) or better: vertical grade, terrain type, outside shoulder width, surface type, surface thickness, median width, and median type.

Table 9. State-Wide Poisson Model Results

	Omnibus test				
Indicator	Likelihood ratio χ^2	Model significance	Deviance/df	AICc	Observations
State model	254.9	0.000	0.379	17,127	6,536

The omnibus chi-square test results indicate an overall strong fit of the model, rejecting the null hypothesis of no effect of the independent variables, while the ratio of the deviance to degrees of freedom (df) permits comparisons between models. This ratio permits comparisons between all models run. Generally, a model is considered to have a good fit if the ratio is less than 2, which is the case here. In contrast, the AICc is sensitive to the magnitudes of the models and is used primarily to compare Poisson and NB models (not shown), rather than to compare between Poisson models. One statistic that does not result from running Poisson regression is an R^2 value as occurs with standard regression analysis, so the statistics above must be used to evaluate overall model performance and, later, to compare performance across the eight field divisions.

Evaluation of independent variables is also different with Poisson models, as categorical variables can have significant model effects as a group while few individual divisions (e.g., the nine total median types or the seven total terrain types) actually prove to be significant. Also, categorical variables are treated in the same way as dummy variables in ordinary regression, in that one category must be left out to avoid redundancy (multicollinearity). This means that interpretations of slope coefficients of categorical variables, in exponential form (*EXP(β)* or e^{β}), are expressed relative to the omitted, final category. Raw negative slope coefficients (less than 1.0 in exponential form) indicate that a given category type contributes to proportionally fewer fatal crashes than does the redundant category, while raw positive coefficients (greater than 1.0 in exponential form) imply the reverse. However, slope coefficients for ratio variables can be evaluated in the usual way, with negative slopes indicating an inverse relationship with the dependent variable and positive slopes indicating direct relationships.

Vertical grade is the most significant (α = 0.000) variable in the state model. Downhill grade is the redundant category (parameter e^{β} = 1.0); the exponential parameter for uphill grades is 1.67 and flat highways is 0.95. Hence, uphill grades contribute to significantly more fatal crashes while flat roads modestly reduce the risk, relative to the rate that downhill grades cause fatal crashes. It may be that downhill travel affords drivers better views ahead to anticipate changing road geometries and oncoming traffic. While not a surprising finding in general, this does create challenges in the context of this study, as leveling hills and filling in valleys is costly, often unsightly, and thus does not qualify as a cost-effective treatment. In comparison, horizontal grade (“x-fall”) is marginally significant (α = 0.079, higher than a default level of 0.050). Right edge higher is the omitted category, and the parameters for both level roads and roads with a higher left edge are both larger than 1.0 and thus contribute more to fatal accidents than roads with higher right edges. However, given that 86% of all crashes occurred on highways with higher left edges, this result may simply be due to the overwhelming presence of this design aspect that is used to drain rainwater to the outside edges of the road.

Terrain type is very significant (α = 0.001), with the residential (urban) type as the omitted redundant category. Both the flat and rolling terrain categories have an exponential parameter larger than 1.0 indicating higher contributions to fatality crash rates than baseline levels in residential areas. Mountainous terrain logically should be significant, but only two percent of Oklahoma highways are classified as such and drivers may practice more cautious driving in those areas. The other four urban categories, not surprisingly, have lower crash rates due to the generally lower speeds and less severe crashes that result, though none of the urban categories were individually significant. Overall, these results reinforce national-level studies [12, 13] that higher speeds and fewer safety treatments increase fatality rates in rural areas.

Surface type (α =0.002) and surface thickness (α = 0.005) are both significant, but these are harder to interpret in a meaningful way. Many surface type categories are defined but just five have notable representation in the dataset, as shown in Table 5. Asphalt type 'E' was the default category and only P.C. concrete among the types in Table 5 has a parameter above 1.0. With respect to surface thickness, in which the highest thickness category (over nine inches in depth) is the redundant category (out of ten total), thickness categories 2 (two to three inches), 6 (six to seven), 7 (seven to eight), and 8 (eight to nine) had slope parameters greater than 1.0 in exponential form. The authors, unfortunately, lack the pavement engineering expertise to address these findings but hopefully pavement engineers can learn from these results and determine their significance in terms of reducing highway fatality crashes.

Other highway design characteristics contributed significant effects in the state-wide model. Outside shoulder width negatively correlates with fatalities, as wider shoulders mean more room for drivers to recover or at least shed speed, and this variable is highly significant (α = 0.001). Median width is also significant (α = 0.008) with the expected negative sign or exponential value below 1.0, as wider medians result in fewer fatalities.

Finally, median type is moderately significant (α = 0.051). It is important to examine this class of design feature due to the ways that median design can mitigate crashes. Cable barriers are the category type omitted due to redundancy, and relative to cable barriers most other median types contribute to higher crash rates, as might be expected. However, two median types have exponential coefficients below 1.0: "flush brick" and concrete barriers. Concrete barriers are even more effective than cable barriers since vehicles can sometimes jump over or occasionally cross through cable barriers, while flush brick type has no obvious explanation. However, only eight fatalities were recorded at locations having this type of median compared to nearly 100 with cable barriers (bearing in mind that this simply means that the highway had cable barriers, not that the fatal crash necessarily struck the cable barrier). This result is likely just a function of the low presence of flush brick medians on highways in the state.

Summarizing the state-wide model, fairly logical findings are confirmed, though the data available limit our ability to pinpoint specific safety treatments that might greatly reduce fatalities on rural highways. Rural areas are much more likely to experience fatal highway accidents, and wider medians (especially with concrete barriers), wider outside shoulders, and flatter roadways all significantly contribute to lower fatality rates. Of equal interest, perhaps, are the variables that are not significant contributors to fatal accidents, even though they could be addressed through road construction practice and/or maintenance: outside shoulder type (α = 0.386), number of lanes (α = 0.396), overall road surface width (α = 0.624), curve radius (α = 0.751), pavement rutting (α = 0.958), and pavement roughness or IRI (α =0.995).

4.3 Field Division Poisson Models

Given the wide variety of landforms, terrains, climates, and traffic patterns across the state, Poisson regression models are developed for each ODOT field division to determine if regional variations exist among significant contributors to fatality accidents. Comparative results are provided in Table 10 for the eight field divisions along with the state model statistics from Table 9. Figure 1, given earlier in this report on page 4, shows the territories of the field divisions. Field divisions consist of either nine counties (divisions 2, 4, 6, and 7) or eleven counties (3, 5, and 8) except division 1 which covers just eight counties. Field divisions adhere to county boundaries, though this scheme is arbitrary and any regionalization scheme could be undertaken for sub-state analysis.

Table 10. Field Division Poisson Model Results

Field Division	Omnibus test		Deviance/df	AICc	Observations
	Likelihood ratio χ^2	Model significance			
1	124.011	0.000	0.317	2,062	754
2	64.176	0.831	0.316	2,311	859
3	105.798	0.046	0.363	2,776	1,002
4	69.164	0.801	0.240	2,809	1,088
5	183.872	0.000	0.404	1,418	459
6	37.497	0.999	0.247	804	256
7	74.533	0.681	0.224	1,746	660
8	190.144	0.000	0.318	3,819	1,458
State model	254.947	0.000	0.379	17,127	6,536

As demonstrated in Table 10, field divisions 1 (east central), 5 (southwest), and 8 (northeast) have extremely significant models while field division 3 (central) achieves moderate significance; these are the only areas where significant models occurred based on overall omnibus tests of goodness-of-fit. In contrast, field divisions 2, 4, and 7 have very weak models and field division 6 achieves virtually no goodness of fit. With barely half as many observations as the field division (5) with the next fewest number of fatalities, field division 6 may simply not have enough data points (due to small populations and low densities) for a good model fit. Conversely, there may simply be no consistent determinants of fatal crashes in this part of the state, containing the far northwestern counties including the Oklahoma Panhandle.

It is also interesting to compare field divisions to each other (and the state model) in terms of individually significant variables. Table 11 lists each variable that had significant model effects in at least one of the nine models, while variables from Table 1 that never proved significant are omitted.

Table 11. Significant Variables in Field Division Models

Variable	Field Divisions								State Model
	1	2	3	4	5	6	7	8	
Median type									**
Median width				*					***
Outside shoulder type							***		
Outside shoulder width		***					**		***
Surface type					***				***
Surface thickness			*		***				***
Surface width					*				
Terrain type					**			**	***
Number of lanes				**					
Curve radius					*				
Horizontal grade			***						*
Vertical grade	***							***	***

*significant ≤ 0.10 level **significant ≤ 0.05 level ***significant ≤ 0.01 level

Table 11 shows some correlation between the number of significant variables in a field division and the overall model significance, but the trend is not perfect. Field division 6 has an utterly worthless regression model and hence no significant variables. However, field divisions 2, 4, and 8 all have poor aggregate models but each has at least one significant predictor of highway fatalities. In contrast, field division 1 had a very significant overall model but just one (very) significant predictor, vertical grade. Field division 5, with the second-highest likelihood ratio value (good) but the highest standardized deviance (bad), has the most significant variables. While this implies that fatal crash prediction should be more accurate in the southwest part of the state, the two most significant predictors were surface type and thickness, two variables whose influence on crashes is uncertain based on the (lack of) expertise of the researchers. The strength of this relationship in field division 5 was likely a significant contributor to the surface variables being significant in the overall state model. Further investigation is warranted to examine field division 5 in more detail to determine why surface type and thickness is so correlated to fatal accidents there.

An examination across the rows of Table 11 reveals that no variable is significant in more than two field divisions. The paucity of significant variables in each region challenges the ability to identify localized areas of improvement in the state's highways, though the overarching goal is to identify system-wide improvements. The seven variables significant in the state model can thus provide some guidance to transportation agency personnel who are better positioned to evaluate costs versus benefits of the safety treatments that might be considered.

5.0 CONCLUSIONS

This project sought to identify characteristics of fatal crash locations on Oklahoma highways. Using ODOT data on highway networks and pavement conditions in tandem with FARS data on recorded characteristics of fatality crashes, this study employed advanced regression (Poisson and negative binomial) techniques to model fatality occurrences across the entire state as well as in sub-regions defined by both ODOT (field divisions) and an arbitrary thirteen county test area in northeast Oklahoma. There was only modest consistency in the results, with median and outside shoulder design consistently proving influential on the likelihood of crashes at given locations. Some other variables were also consistently significant but were harder to explain, such as surface thickness and surface type, the latter especially challenging because only a few pavement surface types dominate the state's highways. The research also discovered that for Oklahoma at least the problems often associated with using Poisson regression for analyzing fatality counts, over- and underdispersion, were not so much an issue and thus both Poisson and negative binomial regression performed quite similarly.

However, there were some large operational challenges, mostly related to the nature of the raw data. Many variable definitions were not clear (at least to non-engineers) and we were not always able to obtain complete explanations of some variables. As a result, some measured variables (such as curvature) had to be converted into categories in order to conduct some type of analysis that incorporated those traits. Additionally, the variable of interest, the locations of fatal crashes, was in point form (latitude and longitude in FARS) while all the roadway data were in vector form. This meant that many widely-used and understood methods of spatial analysis and regression, such as the use of spatial autocorrelation statistics and/or geographically weighted regression, were not deemed useful to this research, though literature reviewed earlier demonstrates that some analysis of this type has been applied to the analysis of crash locations. However, geostatistical analyses are typically employed to identify spatial clusters of high crash locations whereas this research sought to understand the physical traits at individual crash locations and subsequently model those traits.

Furthermore, the discrete nature of the dependent variable (both in variable and physical space) limited analysis approaches, and the Poisson regression method is the most commonly used in the literature but it is not a simple method to employ or describe. This report has endeavored to narrate the results in such a way that readers familiar with basic regression can still understand the overall results and implications. The large number of categorical variables also resulted in the common outcome that an overall category of a given highway characteristic might prove significant in the model but few if any of the individual category tallies are likewise significant.

These various challenges and obstacles were overcome with a lengthy, thorough, and iterative process of data evaluation, clean up, (when necessary) categorization, geo-referencing, integration, and finally analysis. As noted earlier, this process took much longer than anticipated due to the diverse nature of the disparate datasets that were used in this study despite the assistance of both ODOT personnel and researchers at the University of Oklahoma Intelligent Transportation Systems (ITS) lab with whom we collaborated on various tasks.

The ultimate goal of this research is to provide traffic engineers with information about what factors influence the occurrence of fatality crashes across the state. The engineers can focus on the design and condition characteristics over which the transportation agency has authority in order to identify, develop, and apply safety treatments to reduce highway fatalities. Beyond Oklahoma, these methods have broad applicability since data collection and highway construction methods are relatively consistent across the United States. Nonetheless, the diverse agencies that collect these data use different formats best suited to their own needs, and integrating the data to permit meaningful analysis remains a challenge that is also evident in the literature. There are myriad approaches to preparing datasets for this type of analysis, and no single method seems dominant and furthermore every year more new approaches emerge [62].

However, we do not necessarily advocate for standardization of the research, since different researchers and agencies have different goals, but it does seem that there is far less consistency in this area of investigation than other areas of transportation geography. Such diversity is both healthy but also frustrating, as a great deal of time, effort, and expertise is necessary to determine the best way to proceed with research in a given location and with specific datasets available.

In response to the great breadth of techniques and approaches available, this report reviewed in an extensive, but by no means exhaustive, fashion the primary literature on accident analysis techniques. Like most basic, academic research in the social sciences, this project does not analyze or recommend specific accident prevention construction techniques; such research lies in the realm of civil and mechanical engineers, and ODOT. Instead, we proposed and were tasked with identifying characteristics of the state's highways the correlate with fatality crash locations and to report our findings. Oklahoma and the United States have both made significant strides in reducing crashes, fatalities, and other negative outcomes of travel on the state's roadways, and this research will hopefully contribute further to this trend.

6.0 REFERENCES

1. Transportation Research Board (TRB). 2003. *NCHRP Report 500, Volume 6: A Guide for Addressing Run-Off-Road Collisions*. Washington, D.C.: Transportation Research Board.

2. Elvik, R. 2000. How much do road accidents cost the national economy? *Accident Analysis and Prevention*, *32*(6), 849-851.

3. Blincoe, L.J., Seay, A.G., Zaloshnja, E., Miller, T.R., Romano, E.O., Luchter, S., and Spicer, R.S. 2002. *The Economic Impact of Motor Vehicle Crashes, 2000* (Report No. DOT HS 809 446). Washington, D.C.: National Highway Traffic Safety Administration.

4. Chu, X. 2009. Measuring injury risks from motor vehicle crashes with an integrated approach. *Journal of Transportation Safety & Security*, *1*(3), 190-202.

5. World Health Organization (WHO). 2010. *Global plan for the decade of action for road safety 2011-2020*. Geneva, Switzerland: World Health Organization.

6. American Association of State Highway and Transportation Officials (AASHTO). 2005. *AASHTO Strategic Highway Safety Plan*. Washington, D.C.: American Association of State Highway and Transportation Officials.

7. Oklahoma Department of Transportation (ODOT). 2007. *Oklahoma Strategic Highway Safety Plan*. URL: http://www.okladot.state.ok.us/oshsp/pdfs/OSHS-Plan_Final.pdf. Last accessed November 30, 2012.

8. Elvik, R. 2010. The stability of long-term trends in the number of traffic fatalities in a sample of highly motorised countries. *Accident Analysis and Prevention*, *42*(1), 245-260.

9. Sivak, M., and Schoettle, B. 2010. Toward understanding the recent large reductions in U.S. road fatalities. *Traffic Injury Prevention*, *11*(6), 561-566.

10. National Highway Traffic Safety Administration (NHTSA). 2012. *Traffic Safety Facts: Crash•Stats* (Report No. 811 604). Washington, D.C.: National Highway Traffic Safety Administration's National Center for Statistics and Analysis.

11. Oklahoma Highway Safety Office (OHSO). 2011. *2011 Annual Report*. URL: http://www.ok.gov/ohso/documents/2011 Annual Report.pdf. Last accessed November 30, 2012.

12. Chandler, B., and Anderson, R. 2010. *Implementing the High Risk Rural Roads Program* (Report No. FHWA-SA-10-012). Washington, D.C.: Federal Highway Administration.

13. National Highway Traffic Safety Administration (NHTSA). 2007. *Traffic Safety Facts: Rural/Urban Comparison* (Report No. 810 996). Washington, D.C.: National Highway Traffic Safety Administration's National Center for Statistics and Analysis.

14. Transportation Research Board (TRB). 2010. *Research Results Digest 345: Alternate Strategies for Safety Improvement Investments*. Washington, D.C.: Transportation Research Board.

15. Romano, E.O., Tippetts, A.S., and Voas, R.B. 2006. Language, income, education, and alcohol-related fatal motor vehicle crashes. *Journal of Ethnicity in Substance Abuse*, *5*(2), 119-137.

16. Stamatiadis, N., and Puccini, G. 2000. Socioeconomic descriptors of fatal crash rates in the Southeast USA. *International Journal of Injury Control and Safety Promotion*, *7*(3), 165-173.

17. McCartt, A.T., Hellinga, L.A., and Bratiman, K.A. 2006. Cell phones and driving: Review of Research. *Traffic Injury Prevention*, *7*(2), 89-106.

18. McCartt, A.T., Hellinga, L.A., Strouse, L.M., and Farmer, C.M. 2010. Long-term effects of handheld cell phone laws on driver handheld cell phone use. *Traffic Injury Prevention*, *11*(2), 133-141.

19. Atchley, P., Hadlock, C., and Lane, S. 2012. Stuck in the 70s: The role of social norms in distracted driving. *Accident Analysis and Prevention*, *48*, 279-284.

20. Levine, N., Kim, K.E., and Nitz, L.H. 1995. Spatial analysis of Honolulu motor vehicle crashes: I. Spatial Patterns. *Accident Analysis and Prevention*, *27*(5), 663-674.

21. Levine, N., Kim, K.E., and Nitz, L.H. 1995. Spatial analysis of Honolulu motor vehicle crashes: II. Zonal Generators. *Accident Analysis and Prevention*, *27*(5), 675-685.

22. Levine, N., Kim, K.E., and Nitz, L.H. 1995. Daily fluctuations in Honolulu motor vehicle accidents. *Accident Analysis and Prevention*, *27*(6), 785-796.

23. Andrey, J. 2010. Long-term trends in weather-related crash risks. *Journal of Transport Geography*, *18*(2), 247-258.

24. Edwards, J.B. 2002. Motorway speeds in wet weather: The comparative influence of porous and conventional asphalt surfacings. *Journal of Transport Geography*, *10*(4), 303-311.

25. Maze, T.H., Agarwal, M., and Burchett, G. 2006. Whether weather matters to traffic demand, traffic safety, and traffic operations and flow. *Transportation Research Record 1948*, 170-176.

26. Transportation Research Board (TRB). 2003. *NCHRP Report 500, Volume 3: A Guide for Addressing Collisions with Trees in Hazardous Locations*. Washington, D.C.: Transportation Research Board.

27. Transportation Research Board (TRB). 2004. *NCHRP Report 500, Volume 8: A Guide for Reducing Collisions Involving Utility Poles*. Washington, D.C.: Transportation Research Board.

28. Transportation Research Board (TRB). 2003. *NCHRP Report 500, Volume 4: A Guide for Addressing Head-On Collisions*. Washington, D.C.: Transportation Research Board.

29. Transportation Research Board (TRB). 2004. *NCHRP Report 500, Volume 7: A Guide for Reducing Collisions on Horizontal Curves*. Washington, D.C.: Transportation Research Board.

30. Transportation Research Board (TRB). 2008. *NCHRP Report 500, Volume 21: Safety Data and Analysis in Developing Emphasis Area Plans*. Washington, D.C.: Transportation Research Board.

31. Federal Highway Administration (FHWA). 2005. *SAFETEA-LU: A Summary of Highway Provisions*. URL: http://www.fhwa.dot.gov/safetealu/safetea-lu_summary.pdf. Last accessed November 30, 2012.

32. Oklahoma Highway Safety Office (OHSO). 2012. *FY 2012 Performance Plan and Highway Safety Plan*. URL: http://www.ok.gov/ohso/documents/HSP 2012.pdf. Last accessed November 30, 2012.

33. Federal Lands Highway Program (FLHP). 2009. *Oklahoma FLHP*. URL: http://flh.fhwa.dot.gov/about/states/documents/ok.pdf. Last accessed November 30, 2012.

34. Poindexter, K. 2004. *Fatal Motor Vehicle Crashes on Indian Reservations: 1975-2002* (Report No. 809 727). Washington, D.C.: National Highway Traffic Safety Administration.

35. Federal Highway Administration (FHWA). 2009. *Strategic Highway Safety Plan for Indian Lands.* URL: http://flh.fhwa.dot.gov/programs/irr/safety/documents/strategic-hsp.pdf. Last accessed November 30, 2012.

36. Bailey, L., and Huft, D. 2008. Improving crash reporting: Study of crash reporting practice on nine Indian reservations. *Transportation Research Record 2078*, 72-79.

37. Elvik, R. 2008. A survey of operational definitions of hazardous road locations in some European countries. *Accident Analysis and Prevention*, *40*(6), 1830-1835.

38. Anastasopoulos, P.C., Tarko, A.P., and Mannering, F.L. 2008. Tobit analysis of vehicle accident rates on interstate highways. *Accident Analysis and Prevention*, *40*(2), 768-775.

39. Cheng, W., and Washington, S.P. 2005. Experimental evaluation of hotspot identification methods. *Accident Analysis and Prevention*, *37*(5), 870-881.

40. Montella, A. 2010. A comparative analysis of hotspot identification methods. *Accident Analysis and Prevention*, *42*(2), 571-581.

41. Songchitruksa, P., and Zeng, X. 2010. Getis-Ord spatial statistics to identify hot spots by using incident management data. *Transportation Research Record 2165*, 42-51.

42. Flahaut, B., Mouchart, M., San Martin, E., and Thomas, I. 2003. The local spatial autocorrelation and the kernel method for identifying black zones: A comparative approach. *Accident Analysis and Prevention*, *35*(6), 991-1004.

43. Thomas, I. 1996. Spatial data aggregation: Exploratory analysis of road accidents. *Accident Analysis and Prevention*, *28*(2), 251-264.

44. Quddus, M.A. 2008. Modelling area-wide count outcomes with spatial correlation and heterogeneity: An analysis of London crash data. *Accident Analysis and Prevention*, *40*(4), 1486-1497.

45. Wang, C., Quddus, M., and Ison, S. 2009. The effects of area-wide road speed and curvature on traffic casualties in England. *Journal of Transport Geography*, *17*(5), 385-395.

46. Erdogan, S. 2009. Explorative spatial analysis of traffic accident statistics and road mortality among the provinces of Turkey. *Journal of Safety Research*, *40*(5), 341-351.

47. Aguero-Valverde, J., and Jovanis, P.P. 2006. Spatial analysis of fatal and injury crashes in Pennsylvania. *Accident Analysis and Prevention*, *38*(3), 618-625.

48. Cafiso, S., Di Graziano, A., Di Silvestro, G., La Cava, G., and Persaud, B. 2010. Development of comprehensive accident models for two-lane rural highways using exposure, geometry, consistency and context variables. *Accident Analysis and Prevention*, *42*(4), 1072-1079.

49. Zhu, H., Dixon, K.K., Washington, S., and Jared, D.M. 2010. Predicting single-vehicle fatal crashes for two-lane rural highways in Southeastern United States. *Transportation Research Record 2147*, 88-96.

50. Deng, Z., Ivan, J.N., and Gårder, P. 2006. Analysis of factors affecting the severity of head-on crashes. *Transportation Research Record 1953*,137-146.

51. Schneider IV, W.H., Savolainen, P.T., and Moore, D.N. 2010. Effects of horizontal curvature on single-vehicle motorcycle crashes along rural two-lane highways. *Transportation Research Record 2194*, 91-98.

52. Karlaftis, M.G., and Golias, I. 2002. Effects of road geometry and traffic volumes on rural roadway accident rates. *Accident Analysis and Prevention*, *34*(3), 357-365.

53. Lord, D., Manar, A., and Vizioli, A. 2005. Modeling crash-flow-density and crash-flow-V/C ratio for rural and urban freeway segments. *Accident Analysis and Prevention*, *37*(1), 185-199.

54. Miaou, S-P., and Lord, D. 2003. Modeling traffic crash-flow relationships for intersections: Dispersion parameter, functional form, and Bayes versus empirical Bayes methods. *Transportation Research Record*, *1840*, 31–40.

55. Miaou, S-P., and Song, J.J. 2005. Bayesian ranking of sites for engineering safety improvements: Decision parameter, treatability concept, statistical criterion, and spatial dependence. *Accident Analysis and Prevention*, *37*(4), 699-720.

56. Fu, R., Guo, Y., Yuan, W, Feng, H., and Ma, Y. 2011. The correlation between gradients of descending roads and accident rates. *Safety Science*, *49*(3), 416–423.

57. Milton, J., and Mannering, F. 1998. The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies. *Transportation*, *25*(4), 395-413.

58. Chen, J-S., and Wang, S-C. 1999. Statistically modelling relationships between accident types and highway features. *Civil Engineering and Environmental Systems*, *16*(1), 51-65.

59. Miaou, S-P. 1994. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accident Analysis and Prevention*, *26*(4), 471-482.

60. Labi, S. 2006. *Effects of Geometric Characteristics of Rural Two-Lane Roads on Safety*. Publication FHWA/ IN/JTRP-2005/02. West Lafayette, IN: Joint Transportation Research Program, Indiana Department of Transportation and Purdue University.

61. Lord, D. 2006. Modeling motor vehicle crashes using Poisson-gamma models: Examining the effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter. *Accident Analysis and Prevention*, *38*(4), 751-766.

62. Lord, C., and Mannering, F. 2010. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice*, *44*(5), 291-305.

63. Miaou, S-P., Hu, P.S., Wright, T., Rathi, A.K., and Davis, S.C. 1992. Relationship between truck accidents and highway geometry design: A Poisson regression approach. *Transportation Research Record*, *1376*, 10–18.

64. Miaou, S-P., and Lum, H. 1993. Modeling vehicle accidents and highway geometric design relationships. *Accident Analysis and Prevention*, 25(6), 689-709.

65. Saccomanno, F.F., Grossi, R., Greco, D., and Mehmood, A. 2001. Identifying black spots along Highway SS107 in Southern Italy using two models. *Journal of Transportation Engineering*, *127*(6), 515-522.

66. Hadi, M.A., Aruldhas, J., Chow, L-F., and Wattleworth, J.A. 1993. Estimating safety effects of cross-section design for various highway types using negative binomial regression. *Transportation Research Record*, *1500*, 169–177.

67. Shankar, V., Mannering, F., and Barfield, W. 1996. Statistical analysis of accident severity on rural freeways. *Accident Analysis and Prevention*, *28*(3), 391-401.

68. Brown, H.C., Labi, S., Tarko, A.P., and Fricker, J.D. 1998. *A Tool for Evaluating Access Control on High Speed Urban Arterials, Part I.* Publication FHWA/IN/JTRP-98/07-I. West Lafayette, IN: Joint Transportation Research Program, Indiana Department of Transportation and Purdue University.

69. Berk, R., and MacDonald, J.M. 2008. Overdispersion and Poisson regression. *Journal of Quantitative Criminology*, *24*(3), 269-284.

70. Lyon, C., Persaud, B., Lefler, N., Carter, D., and Eccles, K.A. 2008. Safety evaluation of installing center two-way left-turn lanes on two-lane roads. *Transportation Research Record 2075*, 34-41.

71. Eisele, W.L., and Frawley, W.E. 2005. Estimating the safety and operational impact of raised medians and driveway density. *Transportation Research Record 1931*, 108-116.

72. Pérez, I. 2006. Safety impact of engineering treatments on undivided rural roads. *Accident Analysis and Prevention*, *38*(1), 192-200.

73. Malyshkina, N.V., and Mannering, F.L. 2010. Empirical assessment of the impact of highway design exceptions on the frequency and severity of vehicle accidents. *Accident Analysis and Prevention*, *42*(1), 131-139.

74. Welki, A.M., and Zlatoper, T.J. 2007. The impact of highway safety regulation enforcement activities on motor vehicle fatalities. *Transportation Research Part E*, *43*(2), 208-217.

75. Welki, A.M., and Zlatoper, T.J. 2009. How highway safety regulations and enforcement activities affect subcategories of motor vehicle fatalities. *Transportation Research Part E*, *45*(6), 1030-1038.

76. Oklahoma Department of Transportation. 2010. *Road Inventory Manual, 7th Edition*. Oklahoma City: ODOT.

77. U.S. Census Bureau. 2010. *Census 2010*. URL: http://factfinder2.census.gov. Last accessed June 10, 2012.