THE IMPACT OF INDETERMINATE METHODS AND RESULTS ON DECISION-MAKING IN HIGHWAY SAFETY: SPATIAL FACTORS, MODEL SPECIFICATION AND MEASUREMENT ERROR

by

OLAYEMI ADEDIJI

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Planning & Public Policy

Written under the direction of

Robert B. Noland

And approved by

____

____

____

____

New Brunswick, New Jersey

MAY, 2019

ABSTRACT OF THE DISSERTATION

ON THE RELATIONSHIP BETWEEN THE EMERGING EXPERT PARADIGM IN ROAD SAFETY DECISION-MAKING AND THE IMPACT OF SPECIFICATION AND MEASUREMENT ERROR

By

OLAYEMI ADEDIJI

Dissertation Director:

Robert B. Noland

In highway safety practice, the conventional approach to improving safety by implementing road safety treatments is to use deterministic crash modification factors (CMF). The problem with these crash modification factors is that they are actually indeterminate. They are based on estimates of the associations of various road geometry attributes with crash frequency, which are derived from models that can be affected by many methodological and data problems. And since the methods employed are largely dependent on the analyst's discretion, results can vary quite widely. The areas of model specification and data are two areas that are very much subject to the analyst's discretion. Specifying crash models wrongly (introducing specification error) and using data with availability and quality problems (introducing measurement error) can cause erroneous inferences to be made about the results of safety countermeasures that are applied to highway segments, compounding the problem of crash occurrences and potentially creating inefficiency in road safety spending. This problem is exacerbated by the fact that they may be propagated as an industry standard through the existence of an expert manual. I have examined the specific specification error problem of omitted variable bias, where the associations of variables included in safety models are biased due to the omission of certain important variables, and the measurement error problems of data availability and quality. In examining these problems, I used a comparison method, where I compared the results of statistical models affected by the specification and measurement error problems to models where I have attempted to rectify the problems in order to see if there is an improvement in the results of the latter. My findings are mixed; results show no substantial change in the associations with crash frequency between models affected by specification and measurement error and models unaffected for certain variables and show notable change in association for other variables. I have also examined the implications of these problems in practice through interviews of safety practitioners and found that practical limitations preventing transportation agencies from addressing these problems exist.

ACKNOWLEDGEMENT AND DEDICATION

To the one through whom I have accomplished anything I have ever accomplished.

My husband, Oloyede, through your constancy, this too is your accomplishment.

My son Eri, my drive.

I acknowledge my parents Bode and Lara, who have always given me support, encouragement, and so much more. Lamide, Ope and Tiwa, have also been encouragement to me, in words and by example.

I am indebted to many friends and counselors who helped me along the way.

I thank Dean Weston and other EJB staff as well.

And finally, I thank my advisor, Bob Noland and my committee for their invaluable help in bringing this work to completion.

Contents

List of Tables

Introduction

Crash frequency is an established and frequently researched topic in transportation planning and engineering. One reason is that roadway crashes have a high social and economic impact. According to the National Highway Traffic Safety Administration (NHTSA), there were about 32,539 fatal crashes, 2.4 million injuries and 6.3 million police-reported crashes in 2015 (NHTSA 2018). The economic cost of crashes in 2010 (the most recent year for cost data availability) was $242 billion (NHTSA 2018). It is as a result of this high cost of crashes that decisions made to alter highway geometry are backed by research, usually undertaken by engineering consultants. Typically, this research is based on statistical models that predict the ways in which changes made to various roadway attributes, such as the widening of lanes or the addition of roadway medians affect the frequency of crash occurrences.

A wealth of studies examining the soundness of the methods used in making these predictions exists. Researchers have mostly given thought to the adequacy of different types of models, and to the nature of association, whether positive or negative, of some variables on crash frequency. The *Highway Safety Manual* (AASHTO 2010a), published by the American Association of State Highway and Transportation Officials (AASHTO), is currently one of the most important publications based on this kind of research. Many state transportation agencies routinely use it for safety analysis and in determining how to alter common roadway geometric attributes in order to decrease crash occurrences. There are state mandates for the use of the HSM through the Highway Safety Improvement Program (HSIP), under the Fixing America's Surface Transportation Act. The variables typically analyzed are geometric and traffic volume variables including lane width, roadway width, median width and type, shoulder width, horizontal curvature and annual average daily traffic (AADT). This wealth of analysis on the soundness of methods is in line with the general importance, and the high social and economic impact of crashes.

Not as much emphasis however, has been given to the ways in which decision makers use this kind of analysis. In addition, the ways in which specification and measurement error cause indeterminacy and affect the inferences used in making safety decision-making is largely unexplored. In this dissertation study, I examine the ways in which expert recommendations, such as those that make up the *Highway Safety Manual*, are used in making decisions to alter highway attributes to improve road safety. I also examine the ways that specification and measurement error in crash frequency models affect these decisions. I assess the ways by which public servants in the legal system, specifically judges create legal precedent that set an unofficial mandate for the use of the *Highway Safety Manual.* In *Street Level Bureaucracy* (Lipsky, 1980), street-level bureaucrats are described as executors of government policy who through the use of their discretion on a case-by-case basis also form policy. The sum of their individual responses add up to an agency behavior and in *the* case of the adjudication of injury liability cases, form an overall policy of the mandatory use of standards like the *Highway Safety Manual.*

Roadway design and conditions, vehicle attributes and human factors are thought to be important factors associated with the frequency of crashes. Studies that have examined the relative importance of these factors estimate human factors to be the main factor associated with about 90% of all crashes, with road and vehicle conditions being far less important (Petridou, E. 2000). Much of this is attributed to risk taking behavior such as speeding (Evans,L. 1996, Treat, J.R. 1980, Sabey, B.E. 1975). Many of these studies have influenced the creation of policy to influence driver behavior as a solution to address the problem of crashes. These studies have mixed conclusions about the efficacy of policies targeting driver behavior, with the finding that drivers can sometimes compensate for reduced risk by increasing risk-taking behavior (Smeed 1949, Taylor 1964, Näätänen and Summala 1974, Wilde 1982). It is crucial for efforts to reduce the frequency of crashes to be informed by the fact that factors associated with crashes are not equally important. In my dissertation I argue that the conventional approach is to alter roadway design. While studies show that policies to alter driver behavior are not always effective, the relatively marginal influence of roadway design on crash frequency raises questions about the focus on roadway design countermeasures.

An examination of safety decision-making and statistical errors affecting it

In examining safety decision-making, I focus on the *Highway Safety Manual* (HSM) because it is increasingly becoming the standard manual for such decisions. The *Highway Safety Manual* (AASHTO 2010a), published by the American Association of State Highway and Transportation Officials (AASHTO) has created a new “expert” paradigm in road safety decision-making, where the recommendations from the manual are given expert status by consultants and decision makers, and relied upon without much consideration for how they are derived. I discuss my reasons for this assertion in my literature review. The major issue with the HSM being conferred expert status is that it is possible that certain issues affecting the validity of findings, present in the process of model estimation, may result in the failure of road safety improvement decisions to yield expected results.

One of the more important issues possibly affecting the validity of these results is the presence of model specification error. This specification error, especially if caused by the omission of contextual variables or variables that have to do with the location of crash occurrence, will serve to limit the transferability of findings from the *Highway Safety Manual* to other contexts and lead to erroneous inferences. This kind of specification error is known as omitted variable bias. I propose that the omission of contextual variables, including various kinds of demographic and economic variables causes omitted variable bias, since these variables are correlated with crash frequency, and one or more geometric variables. Another important issue that can affect the validity of findings is measurement error. In this case, certain practices in data collection and processing have introduced error into the dataset, such that the dataset is not quite representative of reality. To test these hypotheses, I have examined two datasets using statistical analysis, and carried out a study of how such statistical analyses are used through in-depth interviews of transportation professionals.

The first dataset used contains the roadway network of the state of Pennsylvania. This network consists of roads of six functional classifications. I first examined pavement width, lane count, median width, vehicle miles traveled and sinuosity for each functional classification. These were the link-based models since they contained only road geometry variables. I then re-examined the same road segments for the associations of those variables while including the contextual variables of population density, employment density, and median income on crash frequency. These were the combined models since they combined the road geometry variables with contextual variables. Differences in the results between the link-based model and the combined model are taken to be due to the omission of contextual variables in the link-based model, showing that such models suffer from an internal validity problem (Mitra, Washington 2012).

This same procedure is repeated for another dataset of the roadway network of North Carolina to investigate the effect of using data with improved quality. I use data sourced from the *Highway Safety Information System* (HSIS), a database administered by the University of North Carolina Highway Safety Research Center, in partnership with the Federal Highways Administration (FHWA). I discuss the attributes of HSIS data that make it better than other data in detail in another section. For my North Carolina dataset, I analyze crashes using the same variables I used in my Pennsylvania dataset, along with the additional variables of age, precipitation and elevation.

Through my dissertation research, it is my goal to learn the benefits and challenges created by the emerging expert paradigm on road safety decision-making, with a specific look at how specification and measurement error impact it.

Research Questions

In this section I discuss the main research questions raised by my proposed research topic. Each research question has a sub-set of more specific questions that points to the data and methodology required for addressing the main question.

My first research question asks, *"Is there evidence to point to specification and measurement error in crash frequency modeling"?* The underlying assumption here is that specification error is introduced by analysts' choice to omit contextual variables from model specifications. I discuss the underlying theory as well as empirical research that supports this assumption in my review of pertinent literature. To be able to answer this research question, specification error must be detectable. Therefore, certain follow-up questions arise.

One arising question is *"to what extent can omitted variable bias be detected after accounting for various limitations to its detection?"* Detection is complicated by the fact that the outcomes from one kind of specification error might be the same as the outcomes from another kind of specification or other statistical error. In the case of omitted variable bias, the problem of biased coefficients, may be the same outcome for the conditions of autocorrelation, incorrect functional form, and measurement error. Spatial correlation can adversely affect the precision of parameter estimates (Washington et al.,2010). When an incorrect functional form is used, the parameter estimates for the explanatory variables will be biased (Lord, Mannering 2010). In the case of measurement error, there is an internal validity problem because the data has issues in representativeness. I therefore use methodology that can eliminate these sources of bias in my analyses.

Data availability and quality are two other areas that can affect the inferences made from crash frequency studies. Another question to answer is then *"how do data availability and quality affect the ability to make correct inferences from crash frequency analysis?"* There are many known data availability and quality issues in crash frequency analysis and it is important to understand how they might affect inferences from crash frequency models, since measurement error diminishes internal validity.

My second research question asks, "how might a better understanding of the impact of the problems of deterministic crash modification factors affect decision-making to improve road safety?" This question naturally derives from my first research question because the ultimate goal of crash frequency analysis is to make informed highway safety decisions and implement treatments. This second research question focuses on the decision-making aspects of highway safety. I am mainly interested in four main questions in safety decision-making. The first is *"how is the Highway Safety Manual used in road safety decision-making?"* One of the overall goals of my research is to learn how specification and measurement error can affect decision-making. Understanding how decision-making occurs under the current expert paradigm created by the emergence of the *Highway Safety Manual* is an important starting point, since I am proposing that such errors are introduced through the recommendations made by the *Highway Safety Manual*. The next question is *"Are decision makers aware of possible problems associated with the use of the Highway Safety Manual?"* This question aims to discover the extent of the awareness of problems arising from use of the manual among transportation officials. I am specifically interested in their awareness of the problem of indeterminacy, caused by the introduction of specification and measurement error.

Another question is *"How are transportation officials accounting for the possible problems with the use of the Highway Safety Manual?* It is also important to understand the level of importance that transportation officials place on the indeterminacy problem and what strategies they currently use to deal with it. The final question I considered is *"How can better modeling practices gain ground?"* This question aims to discover what the practical application of improved crash frequency modeling might be on road safety decision-making from the perspective of road safety decision makers. It is asking transportation officials and consultants to weigh the benefits of the *Highway Safety Manual* which include its authoritativeness and its simplification of decision-making, against an alternative that can minimize specification error and enhance the efficiency of road safety spending but is relatively more complex. The issue of efficiency in road safety spending is important as there is a limited amount of funds available for road safety improvements. Currently, the Fixing America's Surface Transportation (FAST) Act has authorized an average of $2.3 billion annually between 2015 and 2020 for highway safety improvement projects at state, municipal, and metropolitan planning organization levels (Federal Highway Administration 2016).

Each chapter of analysis in my dissertation directly responds to these research questions. With my Pennsylvania study, I address the first research question on the presence of specification error, while my North Carolina study is an attempt to confirm the results from my Pennsylvania chapter, in addition to addressing my research question on measurement error. My final chapter of analysis covers the result of interviews I conducted in order to understand safety decision-making, and addresses my second research question. In the next section, I discuss literature pertinent to the above research questions.

Literature Review

One of the key assumptions of my research topic is that specification and measurement errors are important problems that affect crash frequency analysis. This premise is crucial to the other important assumption that these specification and measurement errors impact road safety decision-making because they are transferred through the emerging expert paradigm that has been created by the publication of the *Highway Safety Manual*. The unquestioned application of the recommendations in the *Highway Safety Manual* is enabled by the nature of expert paradigms- they are usually readily applied because of their status as expert advice.

In this literature review, I discuss the *Highway Safety Manual* generally and address such questions as the reasons for its publication, followed by a discussion on how it is creating an expert paradigm and the problems that expert paradigms in turn, create. I look at specific problems with the *Highway Safety Manual,* including problems with its positivist approach and some internal validity problems, both of which are exacerbated by its wide acceptance. I briefly discuss the reasons for its wide acceptance, followed by a discussion of how various states are using it. I then review several published crash frequency studies in order to illustrate how methods used in many studies, which are examples of the studies that the *Highway Safety Manual* sources its recommendations from, may be affected by specification and measurement error. Having shown that most crash frequency studies are potentially affected by specification error, and that application of the recommendations from the *Highway Safety Manual* should be critically assessed because of their basis in such studies, I discuss the theoretical backings for the relevance of contextual variables, the omission of which causes specification error. Finally, I review several empirical safety studies that appear to substantiate this theoretical backing for the relevance of contextual variables by examining their correlation with crash frequency and end with a review of methods used in crash frequency analysis.

The *Highway Safety Manual* and the emerging expert paradigm

The American Association of State Highway and Transportation Officials (AASHTO) intended the *Highway Safety Manual* (HSM) as an optional guide for use by transportation agencies at all geographies, in making and implementing decisions towards the goal of increased highway safety. The *Highway Safety Manual* started out as a Transportation Research Board initiative, as several researchers identified a lack of emphasis on safety in transportation decision-making (Babar, Parkhill 2006). They saw this lack of safety emphasis as caused by the absence of a single authoritative document that could assist decision makers by filling a knowledge gap. In other words, the problem that the researchers saw was that safety related decision-making in transportation was being forgone in favor of other transportation needs such as maintaining optimal road capacity. They believed that this problem was a result of the absence of a widely accepted source of expert knowledge that transportation officials could refer to, for making safety related decisions. Thus, the *Highway Safety Manual*, a collection of such expert knowledge, was published and a paradigm in which transportation officials depend heavily on this expert knowledge is potentially emerging. Next, I discuss the theoretical backing for the notion that the *Highway Safety Manual* is being used as expert knowledge in highway safety decision-making or at least, has the potential and likelihood of being so.

The *Highway Safety Manual as Expert Knowledge*

I propose that expert knowledge is the body of what is known objectively or as fact, to the highest extent possible, and is both technical and authoritative. Its attribute of technicality implies the use of fact-based rationality to solve problems, while its attribute of authority implies its power to out-compete other kinds of knowledge with the potential to solve the same problem, by becoming the chosen alternative. In the case of expert knowledge in transportation analysis and decision-making, the technical attribute implies knowledge about how to determine capacity, safety, scheduling and how to design transportation systems for greater efficiency. The attribute of authority on the other hand, implies the potential of the knowledge or solution in question to be chosen from among other possible alternatives such as the discretionary knowledge of individual professionals. Mitchell (2002) discusses expert knowledge by drawing attention to the technical nature of such knowledge as the basis for its capability to solve problems, and its backing by the law as the basis for its influence over other alternatives.

These attributes of technicality and authority are evident in the relationship seen between technical knowledge and legal backing in many fields. In most developed countries, everyday problem solving is regulated such that only experts- those who are licensed professionals and have proven their technical knowledge and ability or expertise, can confer the solutions. A person who needs a home may not simply build a structure and live in it. There must be environmental and traffic impact analyses, approvals by public planning professionals and home inspections by licensed inspectors. Yet, even in the absence of a defined legal backing for any alternative solution to an existing problem, a single one of those alternatives in question may outcompete the others based on the influence of the person or people backing it.

In a case study of Aarlborg, Denmark, Flyvbjerg (1988) examines a multi-modal transportation plan that city planners and engineers had put together through a process of rational planning. The plan faced strong opposition from commercial interest groups and was in danger of not being implemented. In opposition, these groups proposed an alternative which was based on far less rationality or technical knowledge but was backed by persons who were very influential in local politics and in the local economy. Flyvbjerg discusses this phenomenon as more common in reality than researchers and others who grapple with the issue of the power behind rationality tend to believe.

The procedures put forward in the *Highway Safety Manual* are not legally required as the only way to assess highway safety or to make safety decisions. I argue however, that the dynamics between its capability to solve problems, based on technical knowledge, and the influence and authoritativeness of AASHTO, the publishing agency of the manual, will nevertheless position the *Highway Safety Manual* as the source of authoritative knowledge that transportation professionals and decision makers will increasingly rely on. There is precedent to support this assertion. A similar manual also published by AASHTO- the *Highway Capacity Manual* (AASHTO 2011) is already established as an authoritative manual in conducting highway capacity related analyses and in making capacity related decisions. This is also true for AASHTO's *A Policy on the Geometric Design of Streets and Highways (Green Book) (AASHTO 2011)*. In the next section, I discuss certain important problems arising from the use of authoritative manuals as expert knowledge.

The problem with an expert paradigm in highway safety

There are certain problems with this expert paradigm created by the emergence of the *Highway Safety Manual*. The problems arise partly from the general nature of expert knowledge and partly from complications inherent to the procedures by which the *Highway Safety Manual* recommendations are derived.

The positivist problem

One problem with the general nature of expert knowledge is that it assumes that all problems can be solved through a positivist approach. It ignores the human or other factors that can complicate problems and make them untenable to one-size-fits-all solutions. Not much question about the applicability of the solution is raised. The fact that there may be several alternative forms of knowledge that might apply to a problem, and yet a single alternative is considered expert knowledge is reflective of the value of positivism, a value which can be true of certain solutions in some areas of study but certainly not in all. In *Making Social Science Matter,* Flyvbjerg (2001) refutes the view of expert knowledge as somehow superior to other forms of knowledge about issues and the implications of issues that affect people and their societies. He argues that expert knowledge, which usually consists of facts about objects, is suited to solving problems about objects, since objects have a relatively static nature from context to context and are not influenced by the subjectivity of the observer. It is possible for the same observations to be made about such objects from context to context and observer to observer such that the positivist approach is possible and useful. It is questionable however, whether the field of highway safety is helped by the attempt to solve its problems by applying static technical knowledge to varying contexts that also involve the unpredictable element of human behavior.

As a result of the factors of varying contexts and human behavior, highway safety certainly seems to be an area that is also suited to practical rationality rather than only technical rationality or expert knowledge. Practical rationality refers to knowledge based on experience and is more suitable for inquiry about problems faced by people and societies which are by nature, context sensitive and subject to observer interpretation. For this reason, positivism is not always important or possible. I argue that in transportation planning, technical knowledge is not sufficient to understand problems because the subject matter of transportation problems is not limited to infrastructure such as roads and rail, or operations such as mass transit, but necessarily includes people and society. As such practical rationality is also needed.

Internal validity problems

A more specific problem with the *Highway Safety Manual* as expert knowledge in highway safety decision-making has to do with internal validity issues in the derivation of its recommendations. In the next few sections, I discuss some common internal validity problems in other transportation manuals, and then the problems of indeterminacy and transferability in the HSM.

The *Highway Safety Manual* is not the first authoritative manual with internal validity issues in the fields of transportation planning or engineering. One of the most widely known U.S. based studies of this problem was conducted by Donald Shoup where he examined the parking requirement standards recommended by the Institute of Transportation Engineers (ITE) in their *Parking Generation Manual*. The central finding of Shoup's study is that conventionally, parking requirements are decided based on factors that have little to do with transportation systems and modal share (Shoup 2005). They are instead based on factors such as square footage of buildings or employee counts which while being easier to measure, are not direct determinants of parking demand. Another problem he found was that while parking needs are contextually sensitive, the ITE recommendations are based on a very small sample of suburban contexts where driving is the main and sometimes virtually the only mode of transportation. The result is that using the ITE recommendations for more urban contexts means that parking will be oversupplied.

This internal validity problem is compounded by the fact that many of these guidelines are often codified into the municipal ordinances of many municipalities. This means that they become the standard, and the problems that they cause then become commonplace by proliferation from project to project. Shoup discusses how this occurs. Developers use parking generation rates codified by planners, who obtain these rates from other municipalities or from the *Parking Generation Manual.* The guidelines in the manual which were derived from suburban contexts during peak parking demand are then codified into municipal ordinances as minimum parking requirements (Shoup 2005). The internal validity problems of the HSM can also be compounded, as in the case of the *Parking Generation Manual* if through certain processes, the use of the HSM becomes a requirement for highway safety projects at all levels.

Liability and other reasons for codification of problematic standards

There are several reasons why guidelines with internal validity issues might be codified into municipal ordinances or required for use in other ways. One of the main reasons is that research and analysis to come up with jurisdiction specific solutions is labor-intensive and costly. The simplest alternative to incurring such costs is to adopt standards like the *Parking Generation Manual* or the *Highway Safety Manual* (Urgo, Wilensky et al. 2010). Another important reason is the fact that apart from standards developed by authoritative institutions such as ITE or AASHTO, there are no theories that are known in the planning discipline for the generation of solutions to problems like parking that professionals can discretionarily apply to problems in their locale. Shoup explains that

> "…zoning codes throughout the country contain thousands of different parking requirements- the Ten Thousand Commandments of off-street parking. Planners set parking requirements almost as if they were physicians prescribing drugs, but they have no theory, no training, and often no data to help them. No textbook explains the theory of parking requirements because there is none." (Shoup 2005, pg. 26)

The absence of such a theory means that planners must depend on standards like the *Parking Generation Manual.*

Liability is another very important reason why practitioners resort to standards. The legal system, which executes and creates government policy in court plays a role in setting precedent where the use of standards in transportation planning become unofficially mandated. A very good example of a standard that has historically protected localities from liability is AASHTO's *A Policy on Geometric Design of Highways and Streets (Green Book).* Like the *Green Book,* the *Highway Safety Manual,* is an AASHTO publication. If the adoption of *A Policy on Geometric Design of Highways and Streets* as a standard in highway design was influenced by AASHTO's status as an authoritative institution, then it is not unlikely that the *Highway Safety Manual* will benefit from the same influence. The *Green Book* is currently a mandatory standard in practice, even though just like the *Highway Safety Manual,* AASHTO describes it as a set of optional guidelines. It is practically a mandatory standard because it has been adopted by the Federal Highway Administration (FHWA) as one of the guidelines to be used for the design of projects that are to be part of the National Highway System. Such highways must conform to the standard therein (Urgo et al. 2010). In the next section, I discuss a few cases on the use of the *Green Book* from the state of California to illustrate this issue of liability.

California municipalities are not bound by the use of design standards from the *Green Book* since they are not often concerned with the design of roads that are part of the National Highway System. This notwithstanding, they may demonstrate compliance with standards in the *Green Book* as a way of avoiding liability in the event of being named in lawsuits where the plaintiff is alleging injury due to a dangerous roadway condition, under the California Government Code Section 835. The California Government Code allows the use of design immunity for localities that have been sued for injuries due to dangerous road conditions under section 835. This means that if the municipality can show that the injury in the lawsuit is caused by a design that was based on a prescription in the standard followed (such as standards in AASHTO's *Green Book)*, then they can be eligible for protection from liability (Urgo et al. 2010).

In a similar way, an ITE standard speed limit, set based on studies conducted by ITE was used in *James v. New York State Bridge Authority* to determine a judgement for the city due to immunity from having used this standard. The use of established standards in highway tort liability is not limited to municipalities. Typically, plaintiffs use such standards as the *Green Book* in attempting to prove negligence on the part of the municipalities that an action is being brought against (Blaschke, Mason Jr. 1986).

Tort liability is an important issue that highway authorities must take into consideration and even budget for in advance. As discussed above, highway authorities take it into consideration by designing highways according to established standards in such authoritative manuals as the *Green Book* or the *Highway Capacity Manual,* to be able to use design immunity since precedent shows that this is advantageous. In this way, just as was the case with the *Green Book* in California, an unofficial codification of the *Highway Safety Manual* can occur through precedent set by case law, as various locales proactively seek to protect themselves from liability, using design immunity.

Indeterminacy and Transferability

In this section, I discuss the first internal validity problem specific to the *Highway Safety Manual.* While it is an improvement from the conventional means of safety analysis, some researchers have found that the application of the procedure for safety analysis recommended in the *Highway Safety Manual* can introduce error due to its dependence on certain questionable assumptions. One of the problem areas that researchers have identified is the indeterminacy of crash modification factors (CMF).

The *Highway Safety Manual* contains multipliers that users can apply to baseline crash frequency figures to estimate how specific safety treatments will increase or decrease the baseline crash frequency. For example, the baseline crash frequency of an intersection has been found to be about 12.37 crashes per year. The agency conducting the assessment is interested in applying several treatments to the highway to increase safety. These safety treatments include adding a median, reducing the number of lanes and adding signals to the intersection. This agency will use CMFs from the *Highway Safety Manual* for those variables and apply it to the baseline crash frequency of 12.37. The result, say 8.54 crashes/year, is the new crash frequency figure that can be expected if these treatments are implemented. The *Highway Safety Manual* contains hundreds of these multipliers or crash modification factors for a wide range of safety treatments including lane additions and reductions, medians, lane or road width increases by crash type, and for urban and rural settings (AASHTO 2010b).

Hauer et al (2012) argue that a foundational flaw in the way that CMFs are used in the *Highway Safety Manual* and other authoritative guides lies in the assumption that they are constants (determinate) that can be estimated as the weighted average of available research results, as opposed to random variables whose values can vary widely from context to context based on a range of circumstances that affect them. If they are the former, then their standard error, which is reported in the manual is important, but if they are the latter, the standard deviation of their distribution becomes important.

For CMFs of treatment effects that vary widely across projects and contexts, using a weighted average CMF raises the question of transferability. Using the weighted average CMF or any one CMF chosen from such a case could result in erroneous safety predictions and inefficient decision-making. This means that while the idea of CMFs that can be transferred from past research to future scenarios may work for certain safety treatments, there is no reason to suppose that this is a general case. An illustration of an inference that can emerge from the application of erroneous CMFs from the HSM might be getting a crash frequency change from 12.37 crashes/year from the previous example to 5 crashes/year upon application of the safety treatments erroneously, when the actual change will only be a reduction to 9 crashes/year. Sacchi, Sayed et al (2014) also argue that point estimates for CMF are inadequate because safety treatment associations are continuous over time and should therefore reflect changes over time for countermeasure associations.

I posit that a factor that further limits transferability of these *Highway Safety Manual* CMFs is the presence of specification error due to omitted variables in the models that they are based on. The *Highway Safety Manual* CMFs are derived through an averaging of CMFs found in published and other research, provided that the CMFs have a small enough standard error (Babar, Parkhill 2006). This means that whatever statistical biases affect the original research will also likely transfer over to the *Highway Safety Manual* CMFs. If CMFs are based on models that are affected by the indeterminacy problem I discussed above or omitted variable bias so that the associations of certain safety treatments are overestimated or wrongly estimated, then it is reasonable to expect that the CMFs may also overestimate the change from baseline crash frequency that occurs from the installation of a treatment.

Most published crash frequency research estimate crash frequency parameters using models that exclude contextual variables. Some of the reasons for this include convention in the transportation engineering field, the existence of models that can control for unobserved factors, data availability and quality issues and the methodological complexities of accounting for such contextual variables (Mitra, Washington 2012). This omission of contextual variables has been found to limit the transferability of the crash prediction procedure outlined in the *Highway Safety Manual*.

The *Highway Safety Manual* acknowledges that factors that are unique to certain jurisdictions may affect crash frequency including such factors as climate, driver populations, animal populations and crash reporting procedures. It therefore recommends the use of a calibration factor, to calibrate the safety performance function, to ensure results that are accurate to that jurisdiction (AASHTO 2010, Vol 2, C.8). It also acknowledges that these calibration factors are only adequate to making the models suitable at the state-level but limited in making models accurate to the site level. The calibration factor can be obtained by taking the ratio of observed crashes from the locality in question, to predicted crashes from the HSM safety performance function. When the calibration factor is less than one, it is indicative of a lower crash frequency in the jurisdiction in question, as compared with the crash frequency of the sites used to generate the HSM safety performance function used. When it is greater than one, the opposite is true. In this way, it is expected that the jurisdictional factors that affect crash frequency are accounted for, and the resulting estimation is true to the jurisdiction. In the following section, I discuss some reasons why the calibration factors may or may not actually account for the association of contextual variables adequately, even at state-level-jurisdiction.

How HSM calibration factors might account for contextual variables.

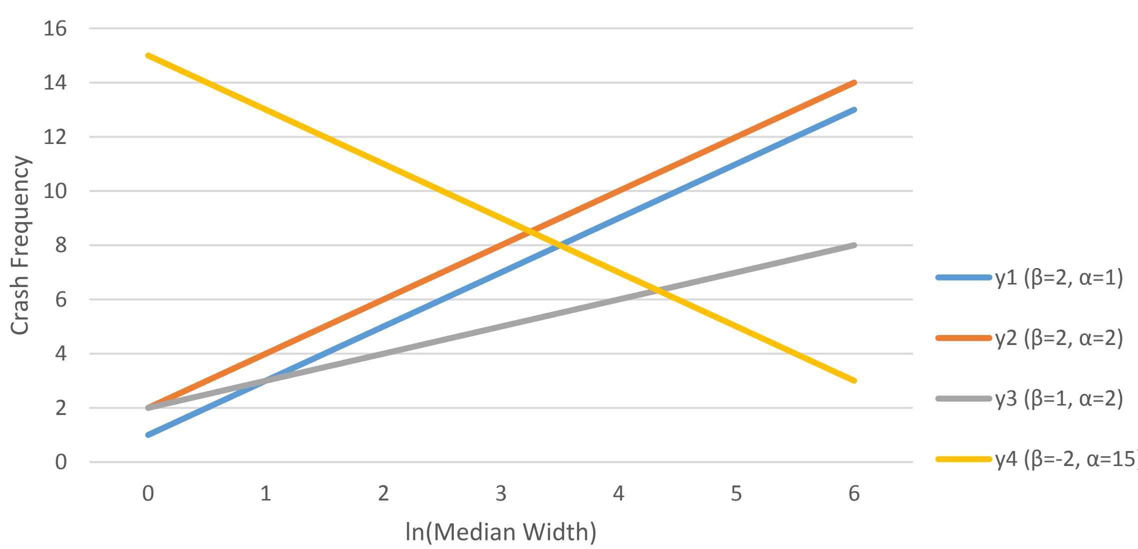
The *Highway Safety Manual* discusses the inability of its predictive method to account for the ways in which contextual factors affect crash frequency as an important limitation. It says:

> "Driver Populations vary substantially from site to site in age distribution, years of driving experience, seat belt usage, alcohol usage, and other behavioral factors. The predictive method accounts for the statewide or community wide influence of these factors on crash frequencies through calibration, but not site-specific variations in these factors, which may be substantial" (AASHTO, Vol 2, pg C-19).

I discuss driver behavior theories and the ways in which contextual variables act as proxies which capture the associations of driver attribute and behavior on crash frequency in a later section. For now, it is important to note that the *Highway Safety Manual* implies that calibration factors cannot be expected to account for contextual variables, which may differ in value quite substantially from site to site. While the HSM mentions that calibration factors account for behavioral effects on a statewide level, it is the case that countermeasures are applied at a site-specific level, making it very important for crash frequency modeling to function adequately and yield results that are as accurate as possible at the site level. I now go on to discuss some of the theory behind the HSM calibration procedure, as well as some research work assessing its adequacy in application.

Researchers examining the HSM calibration procedure have interpreted how the calibration factor works in theory as a re-estimation of the HSM safety performance function (SPF) intercept to what it would be if the SPF had been made based on the local jurisdiction's data (Farid, Abdel-Aty et al. 2016, Sawalha, Sayed 2006). The application of such a calibration factor scales the base model to give estimations that are closer to those of crashes observed in the local jurisdiction. This works, if the contextual variables can be expected to affect variability in crash frequency alone but not the way that crash frequency changes with a geometric variable, for example, median width. Figure 1 below illustrates this kind of change.

Figure 1: Change in Crash Frequency with Change in Median Width



Assume that scenario y1 is the baseline SPF from the *Highway Safety Manual*, and Scenario y2 is this SPF after calibration to a specific jurisdiction. As shown in figure 1, the change in crash frequency as a result of a change in median width is the same in both scenarios. While crashes are higher in scenario y2, a unit change in median width still causes a change by two units in crash frequency. This diagram shows how the HSM calibration procedure scales the SPF to the local jurisdiction in question.

This however is not sufficient, if the contextual variables that the calibration factors are supposed to account for can be expected to also affect the change in crash frequency as a result of a change in a geometric variable. This is illustrated in scenarios y3 and y4. In y3, the intercept of the model is increased from 1 to 2, as in y2, but the coefficient of the median width variable is now 1 from its initial value of 2. This yields a one-unit change in crash frequency as a result of a one-unit change in median width instead of a two unit change in crash frequency as a result of a one unit change in median width from the y1 and y2 scenarios. A more drastic example is shown by scenario y4, where the relationship becomes negative. A one-unit change in median width still yields a two-unit change in crash frequency but in the opposite direction. In other words, median width is now seen to have a negative association with crash frequency. The underlying assumption whereby recalibration to the y2 scenario from the y1 scenario occurs, is that the geometric variables have the correct relationship with crash frequency with respect to magnitude and especially direction of association in the first place. As I have previously hypothesized, this may not be a correct assumption if omitted variable bias is a factor.

Several published studies have attempted to address the question of whether a calibration of the HSM SPF is enough to make it transferable. Sawalha and Sayed (2006) propose that recalibrating the model constant accounts for most factors not included as independent variables in the model. While this may be true, it most likely does not include factors that can cause omitted variable bias, since their association is not included in the constant but in the associations of the independent variables, causing them to be biased.

On the other hand, the study cautions that recalibrating the model constant cannot be sufficient, because the model shape or overdispersion parameter should also be considered (Sawalha and Sayed, 2006). This is based on the fact that different datasets have different shape or overdispersion parameters. The assumption that the overdispersion parameter of a model will be the same as that of the local jurisdiction it is transferred to therefore questionable. The shape or overdispersion parameter is important because it affects precision. The higher the parameter or the greater the dispersion in the data, the lower the precision of the SPF and the lower the overdispersion parameter, the higher the precision and reliability of the SPF (AASHTO 2010). Cunto et al (2015) suggest that the use of calibration alone is insufficient and using other functional forms or explanatory variables may also be needed to improve the performance of the HSM predictive methods.

From this discussion, it can be seen that calibration of the model constant alone may not be sufficient for adequate transferability. Several researchers have pointed out other components of SPFs that need to be considered when being transferred to other jurisdictions. I posit in my research, that omitted variable bias caused by the omission of contextual variables might be another factor that should be taken into consideration for transferability.

Support for the transferability of the HSM calibration factors

One way to find out if scaling the HSM SPFs using a calibration factor is adequate enough for transferability is to apply said SPF and calibration factor to a local jurisdiction. Cunto et al (2015) carried out such a validation procedure to assess the transferability of the HSM SPFs to urban roads in Fortaleza City, Brazil. They obtained a calibration factor by dividing observed crashes by crashes predicted using the HSM SPFs, for a sample of signalized and unsignalized intersections. They then applied this derived calibration factor to predicted crashes for separate samples of signalized and unsignalized intersections. The results show predicted crashes to be very close in value to observed crashes for both the calibration and the validation samples for signalized intersections with very similar results for the mean absolute deviance and mean absolute percentage error for each sample's predictive performance. The results for the unsignalized intersections were similar but showed that prediction did not perform as well as it did for the signalized intersection samples.

An assessment of the cumulative residual plot, which is a common way to check the fit of a model, showed that the SPF, even with the calibration factor applied did not fit the data in the validation samples well for both the signalized and unsignalized intersection groups. The same finding was made when scatter plots showing crashes observed and predicted for the validation samples were assessed. Several other studies have been carried out for various US states, to assess the transferability of the HSM SPFs, with many seeing mixed results and recommending the use of local SPFs instead of the calibrated HSM SPFs including Maryland (Shin, Dadvar et al. 2015), Utah (Brimley, Saito et al. 2012), Louisiana (Sun, Li et al. 2006), and Alabama (Mehta, Lou 2013). The HSM safety performance function was found to overestimate crashes in Maryland, with the recommendation for locally-developed calibration factors where possible. The Louisiana study found better transferability from the HSM with negligible differences between observed crashes and crashes predicted using the HSM predictive method. Abdel-Aty et al (2016) also explore the transferability of SPFs using pooled data from a number of states and found that SPFs using pooled data from two to three states such as the HSM SPFs have better transferability than SPFs that do not.

These studies at the very least, show that the use of calibration factors does not necessarily address the issue of the applicability of the HSM models to local contexts. In the next section, I briefly review existing mandates in certain states, for the use of the *Highway Safety Manual* in conducting safety analysis, after which I review a number of crash frequency studies in order to illustrate the commonality of omitting contextual variables, as well as the wide range of variability of safety treatment associations.

The expert paradigm in practice; State mandates for the use of the *Highway Safety Manual*

I discussed above, the various reasons including labor and research cost, lack of theory, and liability as reasons why standards like the *Highway Safety Manual* may become codified or required. This codification or requirement is an issue because it causes the standards to be used quite extensively even though they can be quite problematic. In the case of the *Highway Safety Manual,* the problems I discussed above include indeterminacy, transferability, and the fact that practical rationality is ignored. In this sub-section, I discuss several examples of states that have made the use of the HSM a legal requirement, in the last decade since its publication.

Most of these states have created this mandate for the funding of highway safety projects through the Highway Safety Improvement Program (HSIP) under the FAST Act. The Fixing America's Surface Transportation Act was signed into law by President Obama in 2015 to secure long-term funding for surface transportation. The HSIP program under the FAST Act is a continuation from previous transportation legislation and sets aside about $2.2 billion per year over the 5 fiscal years from 2016 to 2020 for spending on the improvement of highway safety (LDOTD 2012a). The FHWA has made it a requirement that state highway safety projects that will be funded by this program are to be aligned with the strategic highway safety plan of the various states and must address a specific highway safety problem. Since these state Strategic Highway Safety Plans are required to be data-driven and involve countermeasure analysis, the *Highway Safety Manual* becomes very useful since it prescribes and provides detailed guidelines, data and other resources that facilitate just this kind of analysis. In addition to this, there are specific questions about whether a state has used the *Highway Safety Manual* and how it has used it, in the state annual reports that FHWA requires states to submit for the HSIP program. In addition to this, the FHWA lists the HSM as one of the tools available for safety analysis under the HSIP, along with a recommendation for its use:

> “The Highway Safety Manual (HSM) provides practitioners with the best factual information and tools to facilitate roadway design and operational decisions based on explicit consideration of the safety consequences. The HSM serves as a resource for information related to the fundamentals of road safety, road safety management processes, predictive methods, and CMFs. The road safety management process outlined in the HSM aligns very closely with the HSIP process. Related to the HSIP, the HSM guides safety practitioners in several applications, including: identifying sites with potential for safety improvement, identification of contributing factors and potential countermeasures; economic appraisals and prioritization of projects; and evaluation of implemented improvements.” (LDOTD 2012a)

While this does not constitute an explicitly stated requirement for the use of the *Highway Safety Manual,* this requirement for a specific kind of data-driven analysis, as well as the recommendation of, and questions about the HSM use make its requirement strongly implicit.

The state of New Jersey has been apportioned $57 million annually in HSIP funding. At least one state transportation agency in New Jersey has already mandated the use of the *Highway Safety Manual*- the North Jersey Transportation Planning Authority (NJTPA). NJTPA currently has two programs, under which county roads, local roads, and high-risk rural roadways are funded towards the implementation of small-scale safety treatments. These programs are the Local Safety Program, and the High-Risk Rural Roads Program. Under these programs, project sponsors are required to complete HSM calculations (FHWA 2017). It is stipulated in the New Jersey HSIP manual that the HSM must be used in performing safety analysis, after which the calculations performed according to its guidelines are to be reviewed by FHWA, NJDOT and NJTPA representatives for accuracy before successful submission of applications for funding.

The state of North Carolina dedicated $50 million of state funds reimbursable through HSIP funding, for highway safety improvement projects in 2015. For the same period, $89.7 million in HSIP funding were programmed and $77.8 million were obligated (LDOTD 2012a). In North Carolina, HSIP funds are centrally administered, in partnership with Metropolitan Planning Organizations which may be sponsors of highway safety projects. Although the HSM is used in safety analysis in the state of North Carolina, there is no explicitly stated requirement for its use that can be found in any state safety analysis manuals.

The Virginia Department of Transportation (VDOT) administers the state Highway Safety Improvement Program, through various government agencies, municipalities, organizations, citizen groups or private individuals that may act as sponsors for safety improvement projects (LDOTD 2012a). Virginia currently has an apportionment of approximately $54 million for highway and non-motorized safety improvements. While there is no explicitly stated requirement for the use of the HSM in safety analysis for the state of Virginia, the state provides its own safety analysis tool which includes state-specific highway safety performance functions based on methodology outlined in the *Highway Safety Manual.* It allows agencies to perform network screening and prioritization of crash locations for safety treatment implementation and to select appropriate countermeasures, all heavily informed by standards and procedures from the HSM. This means that all state transportation organizations or municipalities proposing safety projects for HSIP funding in the state of Virginia will invariably use this HSM-based method.

The state of Louisiana has an apportionment of about $65 million per year in HSIP funding. HSIP funds and projects are administered under the Louisiana Department of Transportation and Development (DOTD) through various metropolitan planning organizations and other transportation agencies that may sponsor safety projects. The

DOTD maintains a database of Louisiana-specific data and calibration factors needed for carrying out safety analysis according to the methods prescribed in the *Highway Safety Manual.* Projects that are to be eligible for Louisiana HSIP funding are encouraged to quantify the safety benefit of their proposed safety treatment in dollars, using steps outlined in the *Highway Safety Manual.* In addition to prescriptions from the Louisiana DOTD for the use of guidelines in the *Highway Safety Manual,* the DOTD also has a *Louisiana Highway Safety Manual Implementation Plan*. Under this HSM implementation plan, the DOTD has identified the goal of conducting training courses to "ensure the integration of the HSM into daily project planning, programming and engineering activities" (LDOTD 2012b). Also, under this plan, the DOTD has identified the goal of adopting the HSM as the guideline for DOTD project safety analysis. As with several other states discussed above, while there is no explicit requirement for HSM use, the stated expectations outlined by the DOTD regarding eligibility for HSIP funding for safety projects strongly imply that use of the HSM is in fact a requirement. In addition to this, the DOTD even provides a quick-lookup sheet that makes it possible to check what HSM chapters and applications pertain to the various stages in safety project implementation for HSIP funds in the state of Louisiana (LDOTD 2012a).

Under the Washington Highway Safety Improvement Program, approximately \$78.9 million has been programmed for HSIP project funding, while approximately \$44.8 million has been obligated. Like the state of Louisiana, the Washington State Department of Transportation (WSDOT) also implements highway design through its own state specific design manual. Recently, WSDOT incorporated the *Highway Safety Manual* methods into their WSDOT *Design Manual* (WSDOT 2012).

So far in this literature review, I have made a case against the uncritical use of the Highway Safety Manual due to it positivist and internal validity problems and discussed the exacerbation of these problems through the widespread use of the HSM, as such use becomes a requirement through an expert paradigm. I review theory and empirical studies surrounding the specific internal validity issues of specification and measurement error, in the next section.

A review of crash frequency studies with potential specification and measurement error

The studies I discuss in this section are examples of studies that model crash frequency using only road geometry variables. I highlight them to illustrate the fact that this kind of modeling is conventional in crash frequency research. This convention needs re-examination because the omission of contextual variables contributes to the problem of indeterminacy of crash modification factors, which in turn, limits their transferability.

I examined several studies that looked at the associations of lane and road widening on crash frequency. All the studies I examined that included lane and pavement width variables found negative associations, with values ranging from -0.42 to -0.09. Abdel-Aty and Radwan (2000) examined the associations of traffic volume, road segment length, horizontal curvature, shoulder width, median width, lane width and number of lanes, and urban or rural location on crash frequency. Labi (2011) examined the associations of lane width, shoulder width, pavement surface friction, vertical and horizontal alignment on crash frequency on rural two-lane roads in Indiana. Council and Stewart (1999) examined the association of rural road conversion from two to four lanes. They also studied the associations of the additional variables of traffic volume, shoulder width and road width. Garnowski and Manner (2011) examine the associations of segment length, horizontal curvature, vertical curvature, lane width, traffic volume, and quantity of truck traffic on crash frequency for German autobahn connectors. None of these studies examined any specific contextual variables although some of the associations of some contextual variables such as population density might be captured in a dummy variable like urban or rural location, which Abdel-Aty and Radwan included. The range of associations from -0.42 to -0.09 is also quite large and gives substance to the indeterminacy arguments discussed above.

For the treatment variable of lane count however, I found the range of associations to be smaller than that for lane and road width. A study that examined urban arterials in Vancouver, BC found a positive association with crash frequency, with a coefficient of 0.085 (Sawalha, Sayed 2001). Sawalha and Sayed examined segment length, traffic volume, number of lanes, number of bus stops per kilometer, median type, land use type and percentage of arterial roadway along which parking is allowed as control variables. Zeng and Huang (2014) examined urban roadways in Florida and also found and positive association with a coefficient of 0.167. They examined the control variables of segment length, traffic volume, and number of lanes.

Traffic volume had a wide range of coefficient values in the studies I examined, with the lowest coefficient being 0.24 (Zeng & Huang, 2014) and the highest being 1.18 (Council & Stewart, 1999). Both the Zeng and Huang and the Council and Stewart studies examined only geometric variables in addition to traffic volume.

The range of shoulder width coefficients was also relatively low compared with the range for some of the other variables I discussed previously. The Sawalha and Sayed study found a negative association of magnitude 0.15 (Sawalha & Sayed, 2001) while another study which examined Washington state principal arterials found a negative association of magnitude 0.30 (Milton, Mannering 1998). In addition to shoulder width, this study examined the variables of segment length, horizontal curvature, and vertical curvature.

I examined two studies that looked at the association of median width. One was a study by Malyshkina and Mannering (2010), and the other was the Abdel-Aty and Radwan study I previously discussed. The Abdel-Aty study found a negative association of magnitude 0.024 which was not statistically significant and the Malyshkina and Mannering study found a positive association of 0.905. This was the highest range for the coefficients of any variable with the signs showing opposing directions of association, although one of the coefficients was not significant. The Malyshkina and Mannering study examined the variables for segment length, the curvature of the sharpest horizontal curve found on the roadway segment, interior highway shoulder indicator, and median width indicator.

Crash frequency in the above studies was modeled using various combinations of geometric variables based on the analyst's discretion. This of course, affects the coefficients of the variables specified, and in turn, the inferences made from these models. Along with the nature of the coefficients, and the problem of contextual variable omission, this also contributes to the problem of indeterminacy. In the Highway Safety Manual, crash modification factors are derived from a combination of studies that use different variable specifications for the crash frequency models.

In the next section, I discuss theory and empirical studies that point to the importance of contextual variables. The focus on geometric variables and the corresponding unexplored effects of omitting contextual variables seen in studies like the ones I discussed above, is inconsistent with what driver behavior models indicate and what empirical research on the associations of contextual variables with crash frequency shows. In general, these driver behavior models and empirical research show that contextual variables are important determinants of crash occurrence.

A review of theories of driver behavior

In examining safety research, it is important to consider the human factor for several reasons. The most obvious reason is that driver behavior is known to be a causal factor for crash occurrence. Another reason is that many researchers believe that behavioral changes in reaction to road safety modifications can have such an impact as to reverse the improved safety effects of the modification although there is no agreement on the extent of this reversal (Smeed 1949). This means that knowing how drivers behave ordinarily and more specifically, in reaction to road design modifications is important to assessing the effects of any modification. Many theories about driver behavior exist.

One of the earliest studies that examined the ways in which driver behavior changes was published in 1964 and found that drivers respond to riskier situations and adjust factors such as speed to account for such a perceived risk (Taylor 1964). Taylor theorized that drivers act to maintain a certain level of arousal, as indicated by the driver's galvanic skin response (GSR) which is inversely related to the response reaction such as speed adjustment. Taylor's study formed the basis upon which Wilde (1982) developed his risk homeostasis theory. Wilde argued that drivers arrive at a level of acceptable risk of being involved in a crash that is the result of weighing that risk against benefits such as covering a certain distance in less time through adjusting speed. To illustrate, this means that a driver that is perceiving relatively low risk might respond by increasing driving speed to achieve more distance or save time and conversely when perceiving relatively higher risk, might respond by decreasing driving speed, thus covering less distance in the same amount of time. The implication is that safety treatments to roads might not be as effective as transportation agencies might expect if drivers increase risk taking behavior to garner the benefit of increased mobility, in response to the perceived reduced risk brought about by the safety treatments. Therefore, if drivers act to maintain the same target risk and adjust risk taking behavior up or down to achieve the target, more effective policies may be to change the target risk, as opposed to adjusting road conditions. Methods to change drivers' target risk might be through education and safety campaigns.

Another theory that continues to be debated by researchers is the zero-risk theory posited by Näätänen and Summala (1974). Näätänen and Summala argue that perception of risk is not a direct factor in drivers' decision-making and that what is instead key to driver behavior is motivation. Motivation can be described as the prioritization of such goals as are secondary to the primary goal of arriving at a destination, for instance, the goal of expending the minimum energy possible in taking a trip or doing so in as little time as possible or doing so in such a way as to create an exciting experience. Summala (1988) discusses the implications of this theory for road safety as indicating the need to control drivers' ability to carry out their motivations for instance, using speed control.

There is also the rule-based model proposed by Michon (1985). Michon saw cognitive driver behavior as consisting of a hierarchy of three levels, including the strategic level, the tactical level and the operational level, between which the driver actively makes decisions for execution. He saw these categories as occurring concurrently as the driver switches from one level to another in response to the driving conditions that arise. The strategic level might involve all planning towards certain goals for instance avoiding congested roads, the tactical might involve maneuvers towards negotiating driving conditions such as driving through curves or passing and the operational involves automated actions including braking and steering. Michon sees this rule based model as more capable of practical relevance for improving safety policy based on its suitability for analysis through computational frameworks (Michon 1985, Noland 2013).

Task difficulty homeostasis, a theory put forward by Fuller (2005) takes a departure from the theories I discussed above by proposing that it is not a given level of risk or motivation that drivers aim to remain at, but a certain level of task difficulty (Fuller, 2005). Therefore, a driver might adjust speed by driving more slowly when carrying out tasks that have a higher level of difficulty such as attempting to merge from an entry ramp into fast moving traffic or turning through a sharp bend on the road. Fuller argues that task difficulty may differ between drivers since factors such as their differential capabilities based on experience, familiarity, age, impairment and other attributes not limited to these will vary from person to person. Fuller later included the association of perceived risk to this theory, as the risk allostasis theory. He proposed that a driver will seek to maintain a certain level of risk, in addition to seeking to maintain a certain level of task difficulty based on the driving circumstances and the driver's capability (Fuller, 2008).

All the above theories represent ideas on how the psychological makeup of drivers affect their driving behavior and are more comprehensively examined by Ranney (1994) in his review of driver behavior models. Wilde proposes that risk is the major mental preoccupation of drivers while the other theories discuss motivation or drivers' responses to their capability to carry out certain driving tasks as the main factor. These theories are collectively called cognitive or psychological models of driver behavior and differ from economic models. In the latter, the driver is an economic being that rationally weighs the costs against the benefits of certain actions in order to come to the most favorable decision, for instance, a driver that weighs the gains of reducing trip time against the cost of the increased potential for crash frequency (Hedlund 2000).

Risk compensation, put forward by Peltzman (1975) is one such theory. Peltzman argued that in response to safety regulation to mitigate risk, drivers will respond in such a way as to reduce or negate the mitigating effect of the safety regulation. The logic was that drivers would respond to increases in vehicle safety such as regulation, by trading off the increased safety for time savings through faster driving, which Peltzman termed "driving intensity". The results of econometric analysis that Peltzman carried out led him to conclude that such government regulation was counterproductive as there was no effect on overall traffic fatalities, largely due to increases in pedestrian fatalities.

Other researchers have approached this idea of risk compensation by basing it in the microeconomic theory of utility. Utility is the total amount of satisfaction that can be derived from a good or a service. Since it cannot be measured directly, the preference for a good or service over another as indicated by perhaps the willingness to pay is used to measure utility. Both O'Neill (1977) and Blomquist (1986) estimated utility maximization equations to describe how drivers maximize travel utility by making trade-offs between safety and travel time through speed adjustment. Noland (2013) also proposed a utility maximization function, but based the utility of travel on the price of travel, travel time, the driver capability, in-vehicle activities that may cause distractions, and risk. Noland argues that many of the proposals in the other theories can be brought together such as in his utility maximization function, to reflect the many trade-offs that drivers might make based on varying perception and capabilities.

The state of theories in the safety literature on driver behavior is such that there is both convergence and differentiation. They converge upon the relevance of driver attributes in determining driver behavior that may contribute to crash occurrence. For instance, Wilde's risk homeostasis theory indicates the importance of target risk, the zero-risk theory indicates the importance of motivation and task difficulty homeostasis highlights the importance of willingness and capability to undertake difficult tasks. Target risk, motivation and capability are all dependent on driver attributes which may include gender, age, income and other demographic or socioeconomic attributes.

In the next sub-section, I discuss some research that links some of these attributes with the likelihood of crash occurrence. While these theories differ on what specific factor is important, they exist side by side largely without consensus on which is "correct" and should thereby be the basis for the formation of safety policy. This might appear to be a problematic state of affairs since progress is often assumed to be largely dependent on a known "way forward". To an extent, some progress has already been made, for example, in the area of risk compensation. While the debate sparked by Peltzman's initial proposal of risk compensation did not yield any changes to the development and implementation of Federal Motor Vehicle Safety Standards (FMVSS)[1], it introduced the issue of risk compensation as a reality that implies the need for a more comprehensive assessment of the potential externalities of safety policy (Hedlund, 2000).

This condition is also quite normal in fields such as economics, urban planning and other social sciences. In *Making Social Science Matter,* Flyvbjerg (2001) discusses this ability of social science theories to exist side by side as a major differentiating factor from theories in the natural science which are incrementally developed, such that each subsequent theory must either build upon or refute a prior theory.

[1] FMVSS were a series of regulations to implement safety standards in newly manufacture motor vehicles, mandated by the National Highway Traffic Safety Administration in 1968.

A review of empirical safety research with contextual variables

These theories and their convergence upon the important role played by driver attributes, support the argument that researchers should account for contextual variables in safety research and in understanding the associations of safety policy on crash frequency, since contextual variables can serve as indicators of driver attributes. The underlying assumption that researchers who have incorporated contextual variables into their analyses have based their studies on is that certain driver behavior patterns can be accounted for by the use of variables that represent such behaviors. For example, with respect to the attribute of age, one concern is that younger drivers may not have the experience required to navigate difficult situations, especially when under stress, or may be prone to distracted driving, or excessive driving speed (Chen, Baker et al. 2006, Klauer, Dingus et al. 2006, Deery 1999, Clarke, Ward et al. 2006). This driver behavior pattern can be accounted for, by including age variables in statistical models. If the underlying assumption is correct, roadways in localities that have a relatively larger proportion of younger drivers should show a statistically significant difference in crash frequency from those with smaller proportions.

This same theory can be applied to the effects of other contextually differentiated attributes. Noland (2013), presents some other logical hypothetical behavioral scenarios exploring driver responses to road engineering and design policies and to changes in driving conditions as a result of weather, variations in driver behavior by age and experience of driver, and driver behavior on congested roads that help show the relevance of contextual variables. With respect to the attribute of congestion, some studies show that congestion and traffic calming designs common to urban areas can help to decrease crash severity and that sprawl is associated with higher risk of crash frequency (Ewing, Schieber et al. 2003, Ewing, Dumbaugh 2009).

One study found that trip time, trip cost, and driver characteristics such as gender, family status, driving experience and annual family income were important factors in determining driver choices between various transportation alternatives with the goal of crash risk reduction. The researchers found that married males with more driving experience and higher annual family income showed greater resistance to reducing their crash risk by choosing a safer transport mode alternative if the option in question increased trip time or trip cost than married females (Yannis, Kanellopoulou et al. 2005). Attributes like marriage or family status, and annual income are contextually differentiated attributes in that they differ by location. Data for these attributes are also readily available and easily used in statistical models.

Another study examined risk taking behavior including the tendency to exceed speed limits, to follow the rules for passing and keeping a safe driving distance, the tendency to drive distractedly and conditional capability of the driver, defined as the personal state of the driver at the time of driving including such conditions as energetic or fatigued. The researchers report that subjective risk perception, which affects driver behavior and the likelihood of crash occurrence, differs by various demographic or socioeconomic factors including gender, income, household size and driving experience (Machado-León, de Oña et al. 2016).

Several studies have examined the effects of various contextual variables on crash frequency using statistical models. One study modeled the associations of population density, employment density, average annual daily traffic density, % Hispanic population, and land use variables such as % vacant land, and % commercial, residential and industrial land uses on pedestrian collision density as the dependent variable. The findings were that population density, and certain land uses have significant effects on pedestrian collisions (Loukaitou-Sideris, Sung et al. 2007). Another study found that socioeconomic status was an important variable affecting crashes involving pedestrians, with pedestrian crash likelihood being about 2.7 times greater in poorer neighborhoods (Graham, Glaister 2003). A more recent study, also focusing on pedestrian casualties found that the median household income in an area was a more important variable affecting pedestrian casualties than motor-vehicle casualties (Noland, Klein et al. 2013, Noland 2013). This study supports the findings of an earlier study that also examined various localities with different income levels among other poverty measures, as indicated by the index of multiple deprivation (Noland, Quddus 2004).

In addition to the existing theory that supports the argument that contextual variables have relevance in safety research, these empirical studies show not just the presence of the effect but the magnitude of contextual variable associations with crash frequency. Together, they provide a strong argument for including such variables in safety research and yet, only one study to date has examined the effect of omitting contextual variables on parameter estimates. This study examined the associations of traffic volume, along with proximity to drinking establishments and proximity to schools, on crashes occurring at road intersections (Mitra, Washington 2012). The researchers found that the parameters for the model that omitted the contextual variables were overestimated by up to 40% of their value in the model that included contextual variables. To my knowledge, no study has examined the effect of the omission of contextual variables in models that estimate crash frequency for road segments.

There is some rationale for the focus of crash frequency studies on road geometry despite strong theoretical and empirical support for their relevance. In the next sub-sections, I examine this rationale and review the methods used in crash frequency research.

Policy measures to alter driver behavior

One reason for the focus of safety improvement efforts on roadway design, despite the established importance of driver attributes and contextual factors, might be that the results of driver targeted policies have not been completely positive. A number of studies have demonstrated the effect of safety reversal from risk compensation (Peltzman 1975) in the area of policies to increase safety through the incorporation of airbags in vehicles and mandatory use of seat belts (Peterson et al. 1995, Sen 2001). Certain studies have found significant impacts of policies that mandate seat belt use on safety (Loeb 2001, Coehn and Einav 2003) while others found little evidence of this (Derrig et al. 2002, Garbacz 1991, Harvey and Durbin 1986). Policies to reduce the legal drinking age have been found to be somewhat effective (Asch and Levy 1990, Carpenter and Stehr 2008). Despite the mixed results, studies that investigate the associations between contextual variables and crash frequency are still relevant, because driver attributes, which contextual variable serve as proxies for, are thought to be the main factor causing around 90% of crashes.

A review of estimation methods used in safety analysis

The model type most frequently used by researchers in analyzing crash frequency data is the negative binomial or poisson-gamma model. One reason is that crash frequency data is count data, and cannot be adequately estimated using ordinary least squares regression, since the assumption of normality is not met. Another reason is that the negative binomial model is one of the simpler models to estimate and can appropriately account for over-dispersion, which is a condition that describes a data distribution where the variance exceeds the mean. Using models that cannot deal with over-dispersion results in biased parameter estimates (Lord & Mannering, 2010). In Lord and Mannering's 2010 review of methodological alternatives used in the analysis of crash-frequency data, they examined about 103 studies published between 1984 and 2010. Of the 16 statistical methods used in the studies, the negative binomial model was the most used with 30 studies. The next most frequently used were bivariate/multivariate models which were used in 16 studies, and random effects model which was used in 13 studies. 10 studies used both zero-inflated poisson models and negative binomial models, bringing the total number of studies in which the negative binomial model was used to 40. All other models used such as the poisson, poisson lognormal, negative multinomial, random parameters and hierarchical/multilevel models were used in only 2 to 5 of the 103 studies that Lord and Mannering reviewed.

The negative binomial model is however not without its own short comings. One problem is that it is not capable of accounting for under-dispersion, but this is not an important issue for crash frequency data since the data has a far greater tendency to be over-dispersed (Lord & Mannering, 2010). Another problem is the data requirements. Many model types, including the negative binomial model are prone to error when the sample size is small (Lord, Mannering 2010) and many states simply do not collect the quantity and quality of data that is required for this kind of analysis (Mitra & Washington, 2012).

There are also problems to do with modeling spatial data. It is reasonable to expect that spatial correlation is an issue with studies that analyze contextual variables since spatial entities such as blocks, or block groups may be affected by the same unobserved effects (Lord & Mannering, 2010). Some researchers that have examined the association of contextual variables on crash frequency have addressed this by re-estimating models in their studies using models that account for spatial correlation and comparing them with the initial estimation where spatial correlation was not accounted for (Noland et al., 2013).

Finally, another methodological issue that potentially complicates and limits research on the associations of contextual variables on crash frequency is the issue of geographic units in determining values for the contextual variables analyzed. In the body of research work where spatial aggregation of spatial variables occurs, there is no general theory to base the scale of aggregation on. In other words, there is no general theory that can guide how a variable such as median income may be assigned to a road segment so that the association of median income on crash frequency can be determined. One hypothetical method can be to assign the average median income for instance, of all the block groups, if this is the chosen geographic unit, within a distance of a quarter mile, or a half mile to the road segment in question. The question that then arises is whether the scale of the influence of median income on the road segment in question is in fact a quarter mile or a half mile, or something else. There is also the question of whether this assumed scale varies by the specific spatial variable. This problem is known as the modifiable areal unit problem (MAUP). It points to the modifiable nature of geographical units of analysis based on analyst discretion and is often done in a relatively arbitrary fashion (Pietrzak 2014). This problem can introduce measurement error because there is no way to ascertain whether or not the chosen scale of aggregation accounts for the entire variable effect being studied or accounts for other variable effects not being studied. This is an important issue because parameter estimates will change based on the scale chosen and the wrong scale will mean that these estimates are biased (Openshaw 1984).

A number of researchers have examined the issue of modifiable geographical units in the analysis of transportation data. Abdel-Aty et al gives a review of some of the research work done to examine how modifiable geographic units affect parameter estimate variability (Abdel-Aty, Lee et al. 2013). In the absence of a generally accepted theory or set of guidelines for choosing an appropriate geographic scale of analysis, many researchers routinely use various units in their analysis, including traffic analysis zones (TAZ), census tracts, census block groups and census blocks depending on data availability. This will affect the inferences made, and can lead to internal validity problems if the chosen scale is not appropriate for the variable of interest. There is some research that has compared parameter estimate variability and model goodness of fit from one geographic scale of analysis to another. One study concluded that the effects of variability in scale is unpredictable for multivariate analysis but may be predictable for bivariate analysis (Fotheringham, Wong 1991). Another study examined variability across traffic analysis zones, census block groups and census tracts for the dependent variables of total crashes, severe crashes and pedestrian crashes and found that there was variability in the significance of the parameter estimates across the various geographic zones but not in the direction of association (Abdel-Aty et al., 2013). The direction of association remained unchanged as long as the variables maintained significance across the models with varying geographical units.

An important problem from the use of aggregate data is ecological fallacy. Ecological models use aggregate data, and ecological fallacy is the misinterpretation of relationships seen at aggregate level as affecting individual cases (Freedman 1997). In crash studies, an ecological fallacy would be interpreting a positive association between lane count for instance, and crash frequency as being true of individual sites. Depending on the dataset, this may or may not be true. Despite this important issue, ecological models are still widely used in crash studies because they can be a powerful tool for making inferences, as long as the correct interpretation is made. The value of using aggregate data in ecological models is that when both the inferences and their interpretation are correct, generalization is possible, and much can be understood about many cases as opposed to a single case. In the field of transportation, ecological models are used because aggregate data is easier to obtain than data on individuals (an individual level model would also need individual level data on those not involved in crashes). Making correct inferences from ecological models is dependent on having prior and detailed knowledge of the dataset and the way that it was aggregated (Davis, G.A. 2004).

A note on variable choice

Also important to the question of methods used in assessing variable associations is the issue of model specification. Some of the crash frequency studies I reviewed specified roadway width, shoulder and median width, segment length, lane width and lane count, as explanatory variables. Most specified a combination of a smaller subset of these variables. Roadway width, shoulder and median width, segment length, lane width and lane count appear to be by far, the most commonly specified variables, perhaps because it is easier to obtain data for these variables than other variables. Horizontal curvature on the other hand, is reported in many studies as being a limiting factor on model specification because curvature data is difficult to obtain or tedious to measure. In addition to the more commonly specified variables of roadway width, shoulder and median width, segment length, lane width and lane count, some studies specify a variety of traffic composition variables. They include percent small-vehicle composition (Chiou, Fu 2013) and percent truck composition (Garnowski, Manner 2011, Milton, Mannering 1998). A few include weather related variables (Shankar, Mannering et al. 1995), and indicators for the presence of and variables for roadway lighting, quantity of driveways, pedestrian crossings and bus stops, parking, and pavement quality (Sawalha, Sayed 2001, Labi 2011, Malyshkina, Mannering 2010).

A number of important factors determine the combination of variables used in estimating crash frequency. Data is one of the most important limiting factors. Studies are limited to the assessment of variables for which data is available. This can in turn, create methodological limitations, such as the introduction of omitted variable bias among others, into the study. Irrespective of this, data availability is one reason why a study may limit the variables included to AADT and a few other variables even if theoretically, it makes sense to include other variables.

Having a theoretical basis is also an important determining factor for model composition. A study might reasonably exclude a certain variable because no theoretical basis can be found to include it, even when other studies have included such a variable. Again, this is dependent on analyst discretion and can cause models specifics to vary widely across studies.

Methodological reasons include the need to avoid multicollinearity, which can create large variances, rendering variable parameters meaningless. One of the affected variables will usually be dropped from the model as a solution to the problem of multicollinearity. Most studies will have to consider these issues.

The reason for including contextual variables in this study is because theory suggests their importance to crash frequency, and the likelihood of creating statistical errors when such variables are left out. Data availability plays an important role in the choice of contextual variables included. Data is more easily available for such contextual variables as percent residential, commercial, industrial and other types of land use variables, population and population density, employment, income, proximity to various destinations of interest such as schools, bars or transit access points. In both studies included in this dissertation, the presence of multicollinearity with the inclusion of certain variables limited the contextual variables included to population density, employment density and median income, since population density tended to be multicollinear with population, and employment density tended to be multicollinear with percent commercial land use.

Another factor that determines the choice of contextual variables are the causal theories that the operationalization of these variables are based on. Driver motivations and attributes such as proficiency, risk-taking tendency, the need to maintain a certain level of mental engagement or task difficulty can be theoretically operationalized as variables that can be specified in crash frequency models. While a number of studies have examined the associations of such contextual variables in crash models, there are no studies to my knowledge that assess the measure to which these variables have correctly operationalized driver attributes. One of the most commonly explored relationships is between the driver attribute of proficiency and age. Chen, Baker et al. (2006) Klauer, Dingus et al. (2006), Deery (1999), Clarke, Ward et al. (2006) are some examples of studies that use age with the assumption that it is representative of driver experience or proficiency. Place attributes can be more easily operationalized because many are measured quantitatively. For example, sprawled neighborhoods can be expected to have lower population density while dense urban cores have higher population density. Some studies exploring such place-based contextual variables include Ewing, Schieber et al. (2003), and Ewing, Dumbaugh (2009). Another study, speculating that married individuals may be more risk averse because of their family responsibilities found a negative association between being married and crashes (Yannis, Kanellopoulou et al. 2005). Without an established method of making the link between driver and place-based attributes and the contextual variables explored, analysts must base their choice of contextual variables on theory and this choice will in turn affect results and contribute to indeterminacy.

In the above sections, I have discussed the problem that my research is focused on, and the questions that this problem poses. I have examined pertinent literature that help de-mystify these questions and their implications and discussed methods typically used in similar research. In the next and final section before I discuss the research carried out, I briefly touch on the significance of my research and expected findings.

Expected Findings and Research Significance

The following is a discussion of my expected findings and the significance of this dissertation study.

Specification error and spatial autocorrelation

I expect that my research results will show coefficient value change of considerable magnitude for some road related variables, and possibly also show change of direction of association as a result of including formerly omitted variables, thereby indicating the presence of specification error in the models where such variables are excluded. I also expect that the test for the presence of spatial autocorrelation will show that it is a factor causing bias in my dataset for the simple reason that the road segments and block groups I analyze, are all located within close proximity to one other.

I expect also to find that certain geometric and traffic volume variables are more substantially affected by the omission of certain contextual variables than others. This has some significance for understanding the relationship between certain geometric and traffic volume variables and certain contextual variables. Having such an understanding might enable decision makers and transportation professionals to exercise more control over road design in diverse environments towards better road safety performance.

One other finding that may emerge is the relative importance of specification error compared to other conditions that cause bias, in modeling. One such condition I discussed is spatial correlation.

Data issues:

I expect to find that data availability and quality are significant limitations affecting the validity of results. Road safety research depends to a large extent on data collected by public agencies from first responders and such datasets are known to have significant amounts of missing data. One specific case of this is seen in police reports from crashes involving only property damage. This limits the number of observations that can be analyzed as well as introduces systematic error if the observations that are excluded from the analysis are not random.

Research Significance:

The findings that result from my analyses will be significant in certain important ways. My research will show the importance of the problem of indeterminacy arising from specification error and analyst research methods, to sound research practices and to safe-guarding lives and property on our roads. My research will also show how variables might be differentially affected by specification error. This could lead to better road safety decision results if variables known to be more impacted by specification error are treated with caution.

A more practical interpretation of the amount of bias is what it means in terms of dollars spent or spent inefficiently. It could mean errors made in capital project budgeting for localities. For instance, a municipality might have expected to spend approximately $24,000 in increasing the number of lanes of a road segment from 2 to 3 lanes, expecting that it would result in 1.56 less crashes, when it would result in only 0.352 less crashes. This means that while the expectation was that crashes would diminish by 156%, they would in reality, diminish only by 35.2%, a substantially smaller reduction than was expected. This municipality under-budgeted for the amount of funding that was actually needed to achieve the goal of crash reduction that was initially set. This kind of error can give the perception of inefficient spending since the measures taken did not yield the expected results. The outcomes could be worse in cases where specification error from omitting contextual variables causes the direction of association of the variables in question to be wrongly indicated.

Another important contribution of my research might be to further the ongoing conversations on the adequacy of current procedures in the research and practice of crash frequency reduction. I mentioned earlier that only one other research work has examined the adequacy of specifying crash reduction models with only road geometry and traffic volume related variables. By all indications, this kind of model specification is virtually the standard in research today and may also be standard in practice. Furthering the conversation about its adequacy can lead to methodological improvements in the future.

Crash Frequency Analysis for Pennsylvania State Roads

In this chapter I examine the hypothesis that there is specification error in crash frequency models that exclude contextual variables, using a dataset of Pennsylvania roadways. This section addresses a part of my first research question: “is there evidence to point to specification and measurement error in crash frequency modeling?”

I discussed previously that one of the issues with the *Highway Safety Manual* as expert knowledge is that its recommendations are based on statistical modeling processes known to introduce errors that limit its use across contexts or its transferability. This well documented transferability issue points to the importance of contextual factors affecting crash frequency, even though the specific issue of how contextual factors are important to model specification has largely been unexplored. This issue of model specification with contextual factors is important for gaining an understanding of the nature of the errors that affect inferences from crash models, especially those inferences that will guide road safety decisions. In the next few paragraphs, I discuss data for crash frequency analysis in general as well as a more specific discussion of my Pennsylvania dataset.

Data

I chose Pennsylvania for my analysis because data was readily available from the Pennsylvania Department of Transportation (PENNDOT 2017). I carried out my analysis using three main data sub-sets including crash frequency data, road attribute data, and place attribute data (contextual variables). Road attribute data is typically immediately usable for regression analysis since each crash observation has only one possible value for the corresponding road attribute in question. For instance, the road attributes corresponding to crash X will be the attributes of the road segment where crash X occurred. This is in line with the assumption that the road attributes at point X are factors that are associated with crash X. I obtained a total of 112,502 road segments that consisted of both rural and urban interstates, principal arterials, minor arterials, collectors, local access roads and interchange ramps. Table 1 below shows a summary of the dataset by road type:

Table 1: Summary of Pennsylvania road network:

	Number of segments	% of Total	Total length (miles)
Interstates	7,580	6.74%	3490.20
Principal Arterials	17,859	15.87%	7384.22
Minor Arterials	21,551	19.16%	9233.49
Collectors	45,843	40.75%	20203.72
Local Access Roads	18,965	16.86%	8598.36
Interchange Ramps	704	0.63%	229.30
Total	112,502	100%	49139.28

PennDOT Pennsylvania State Roads 2015 data

There were approximately 50,000 miles of roadway in the Pennsylvania dataset, with most of the roads being collectors, principal arterials, minor arterials and local roads. I obtained the dataset as a shapefile from PennDOT, with the roads already divided up as segments.

The geometric attributes analyzed were already appended to the shapefile. They included pavement width, lane count, median width, segment length and average annual daily traffic. Certain variables used in my dataset were derived from these original variables. The first was vehicle miles traveled which I derived by multiplying segment length and annual average daily traffic (AADT). The other was sinuosity, a measure of curvature, which I obtained by running a function in ArcGIS that was a ratio of the straight-line distance between the start and end points of a road segment to the actual length of the segment (ESRI 2011). A sinuosity value of close to one means that the segment was relatively straight, while a value closer to zero indicates a curvy road segment. Table 2 shows the distribution of these variables along with the total crashes dependent variable.

Table 2: Distribution of Total Crashes & Geometric Variable

Variable	Mean	Standard Deviation
Total Crashes	7.89	13.63
Total Width	23.43	7.01
Lane Count	1.99	.36
Median Width	10.77	70.01
Annual Vehicle Miles Traveled	2172.50	4064.38
Sinuosity	.96	.09

(PENNDOT 2017)

For each road segment I analyzed, there was an average of 7.9 crashes, with a standard deviation of 13.6. Most segments were two lane roads with an average total width of 23 feet. Just as in the typical count model, my main dependent variable, crashes, is overdispersed, making the use of the negative binomial model appropriate.

I obtained crash frequency data from PennDOT for the years 2009 to 2013. Five-year data blocks are typically used in order to have a large enough dataset for analysis, since crashes are rare in general. As can be seen from Table 3, the vast majority of the crashes that occurred in this period were either property damage or injury crashes. Crashes with fatalities were relatively rare.

Table 3: Summary of Crash Types, 2009 - 2013

Crash Type	Frequency
Crashes with only property damage	418,490
Crashes with injuries	464,052
Crashes with fatalities	7,964
Other Crashes	2,507
Total crashes	887,999

(PENNDOT 2017)

The final subset of data that I needed to conduct my crash frequency analysis for this dataset was contextual data. I discussed theory as the basis for specifying contextual variables in crash models in my literature review. Place-based variables which are often quantitatively measured are easily included in crash studies. Population density and employment density, which both measure the extent of sprawl or lack thereof have been used in past studies including Ewing, Schieber et al. (2003), and Ewing, Dumbaugh (2009), and are also specified in my crash models for both my Pennsylvania and North Carolina datasets. Median income, has been found to have a strong association with crash frequency with one study finding that males with higher incomes were more likely to be involved in crashes (Yannis, Kanellopoulou et al. 2005). I also specify median income in my models.

Contextual data requires more processing than road geometry data before it can be used in regression analysis along with geometric variables. For road geometry variables, the value of the variable in question, on the specific road segment on which the crash occurred is important. For contextual variables on the other hand, conditions beyond the immediate vicinity of the crash are important because such attributes have a wide area of influence. Take population density for example, which is measured as the population per geographic unit (people per square mile in a block group, city or state). Data available for population density spans more than the immediate vicinity of a crash frequency since the segment is located in a block group, city or state. It is also reasonable to assume that in addition to the block group in which the crash occurred, other block groups located in close proximity to the crash location are important factors since driver behavior is not localized to block groups.

In order to account for the associations of several block groups on a specific road segment, an aggregation of the attribute by the block groups in question is necessary. Mitra and Washington (2012) used a summation of totals from all block groups within a certain distance of crash locations for their chosen contextual variables including number of schools and number of drinking establishments as their method of aggregation. For my variables, I used average of block group data.

I obtained my contextual data using block groups as my unit of observation. I then attributed my contextual variables to road segments by assigning the average value of the variable in question, for each block group found within a quarter mile of a road segment to that road segment. For example, if there are three block groups within a quarter mile radius of a road segment, each with a population density value of 200, 350 and 400 people per square mile, then the population density value assigned to that road segment would be 316.7 people per square mile. In this way, road segments can have not just geometric attributes, but contextual as well. This is where the MAUP issue discussed previously comes in. There is currently no recommended method of aggregation and researchers must therefore address this issue based on their individual judgement.

I obtained contextual data from the US Census Bureau, for median income, population, and employment at the block group level and for the years between 2009 and 2013. The employment data consists of employment totals by block group in terms of work destination. I used employment at work destination rather than just the number of those employed by block group because the former has a greater significance for transportation, since it entails commutes to work. At the time of analysis, Pennsylvania had a total of 9740 block groups. The actual variables accounted for in the models were average median income, population density by square mile and employment density by square mile.

Table 4: Distribution of Contextual Variables

Variable	Mean	Standard Deviation
Median Income	54,468.03	27,536.76
Population	131.58	95.25
Employment	569.42	1459.70

(US Census Bureau)

The choice of variables to specify in my crash frequency models is an important issue I discussed in my literature review. In order to explore the adequacy of the variable combination that I use in my models, I checked their correlations. The results are presented in Table 5 below. None of the variables have particularly high correlations, except population density and employment density, which have a correlation of 0.85. This high positive correlation between population density and employment density is most likely indicative of the tendency of businesses to locate where there are concentrations of residential land uses, in order to easily access human capital.

Table 5: Variable Correlations

	Pavement Width	Lane Count	Median Width	Vehicle Miles Traveled	Sinuosity	Population Density	Employment Density	Median Income
Pavement Width	1							
Lane Count	0.4896	1						
Median Width	0.2138	0.1137	1					
Vehicle Miles Traveled	0.4295	0.1869	0.4215	1				
Sinuosity	0.1803	0.246	0.0851	0.093	1			
Population Density	0.4433	-0.0088	0.0622	0.3256	0.0183	1		
Employment Density	0.4464	-0.0393	0.1441	0.4247	-0.0044	0.8524	1	
Median Income	-0.074	-0.0162	0.0218	0.1532	-0.0081	0.0685	0.0694	1

Since population density and employment density have a high correlation, there are concerns about multicollinearity. To further explore this, I calculated the variance inflation factors of my variables (VIF) (Table 6). While this is more appropriate for ordinary least squares regression, it can still be used to determine the degree of multicollinearity between variables intended for specification in non-linear models. The VIF for population density and employment density are higher than those for the other variables, but still quite far from raising concerns of multicollinearity affecting the results. As such, I included both variables for all the models in this dataset.

Table 6: Variable Inflation Factors

Variable	VIF	1/VIF
Pavement Width	2.05	0.48781
Lane Count	1.53	0.65416
Median Width	1.24	0.80682
Vehicle Miles Traveled	1.63	0.61535
Sinuosity	1.08	0.92989
Population Density	3.84	0.26015
Employment Density	4.19	0.2389
Median Income	1.06	0.94219
Mean VIF	2.08	

Methods

The topic of methods used for detecting omitted variable bias has been given some attention by a number of researchers. Washington et al (2011) discuss some of the complexities of detecting omitted variable bias in non-linear count models. One complexity arises from the fact that the indicators that can be used to detect omitted variable bias, are also indicative of other conditions that can lead to erroneous inferences. Bias in parameter estimates is one such indicator since it is not only the result of omitted variable bias, but also of over-dispersion, among other problems (Lord, Mannering 2010). Over-dispersion is commonly dealt with by using negative binomial regression since it accounts for this condition, unlike Poisson models. This means that if a negative binomial model is used in analyzing crash frequency, it is reasonable to assume that over-dispersion is not a factor causing bias in the independent variable parameters and any bias seen can be attributed to other reasons such as omitted variables, if there is a strong enough theory to indicate that omitted variables might be the issue.

Another condition that can affect the accuracy of coefficients is spatial correlation. It should also be accounted for in order to adequately isolate the bias of parameter estimates due to specification error arising from omitted variables. This is because the presence of spatial correlation can cause imprecise parameter estimates (Lord, Mannering 2010). This means that the best way to detect omitted variable bias is to eliminate other possible explanations for biased parameter estimates. I am therefore interested in testing whether there is a change of considerable magnitude in the coefficients of the variables of interest, or a change in the direction of association, when compared to their associations in more fully specified models. I am also interested in seeing the degree to which the change persists when spatial correlation is controlled for.

I carried out two main procedures for detecting omitted variable bias. In the first procedure, I compared the associations of geometric variables on crash frequency in models where contextual variables are included (combined models) to their associations in models where contextual variables are excluded (link-based models). Mitra and Washington (2012) use this method in detecting the associations of omitting the contextual variables of proximity to drinking establishments, and proximity to schools, among other contextual variables, on two traffic volume variables. Based on the theory that the associations of omitted variables are captured in the associations of specified variables, associations of greater magnitude seen in the link-based models when compared with the combined models may be indicative of omitted variable bias.

In the second procedure, my goal is to be able to confidently conclude that any bias seen is from the omission of contextual variables. I use the negative binomial regression, which uses maximum likelihood estimation (MLE) to specify these models initially, and then re-estimate with negative binomial conditional autoregressive models using *CrimeStat* software (Levine, Lord et al. 2010) so as to be able to detect and minimize the effect of spatial correlation. It is important to account for spatial correlation because spatial entities such as road segments, blocks or block groups can be affected by the same unobserved effects (Lord & Mannering, 2010). These models use Markov Chain Monte Carlo estimation (MCMC), a Bayesian estimation technique. This has the added advantage of producing a credible interval that shows with a specified confidence level (e.g. 95%) the range of values that the correct estimate lies within. I then compared the coefficients of the variables specified with the omission of contextual variables against the coefficients in the models where contextual variables were included in the re-estimated models to see if there is still a significant difference in the magnitude or direction of associations. This step helps narrow down the source of any bias found to the omission of contextual variables. Since using MCMC models takes care of more conditions that may confound inferences made from my results, I present and discuss the MCMC models here, while also touching briefly on the MLE models, and displaying them in appendix B.

Results and Discussion

I carried out the analysis in three stages. First, I modeled the associations of the contextual variables, along with 5 variables for road density by functional classification. The purpose of this part of the analysis is to be able to compare these baseline associations of contextual variables with the associations of the same variables when specified in models where road geometry variables are included. In the second stage, I modeled the associations of just the road variables on crash frequency (link-based models), so as to be able to compare their associations in these link-based models against their associations in combined models where they are specified alongside contextual variables (combined models). This final stage of modeling, where both the road geometry and contextual variables are specified was the third stage of the analysis. There were three main dependent variables specified for all the models in the three stages. They were total crashes, fatal and major injury crashes, and fatal and all injury crashes. Fatal and major injury crashes are those for which a fatality or an injury which caused temporary or permanent incapacitation were recorded. Fatal and all injury crashes are those for which a fatality or any severity of injury was recorded.

All the model results displayed are numbered in sequence. When discussed in the text, they appear with the same labels used when displayed as tables, as well as another label in parenthesis that shows whether they are spatial (S), link-based (L) or combined models (C), MCMC or MLE models, and the dependent variable of the model represented by a code of 1, 2 or 3. The dependent variables are represented by 1 for total crashes, 2 for fatal and major injury crashes, and 3 for fatal and injury crashes. To illustrate, a model label of L-MCMC-1 means that the model is a link-based MCMC model, with total crashes as the dependent variable. A model label of C-MLE-2 means that the model is a combined MLE model with fatal and major injury crashes as the dependent variable.

Crash frequency analysis for entire Pennsylvania road network

Previously, I made the case that using negative binomial conditional autoregressive models is a better way to model this kind of data, rather than simply using negative binomial models. I therefore present and analyze the results of the negative binomial conditional autoregressive models (MCMC) in this section, while showing the results for the negative binomial models (MLE) in the appendix.

Table 7: Spatial MCMC Models (Entire network)

	Model **1** (S-MCMC-1)				Model **2** *(S-MCMC-2)*			
	Crashes	**t-stat**	**2.5th Percent.**	**97.5th Percent.**	**Fatal & Major Injury Crashes**	**t-stat**	**2.5th Percent.**	**97.5th Percent.**
Population density (ln)	-0.457	-58.468	-0.532	-0.381	-0.536	-45.289	-0.651	-0.422
Employment density (ln)	0.217	48.318	0.173	0.260	0.124	17.853	0.058	0.191
Median income (ln)	0.081	7.870	-0.024	0.177	0.065	4.742	-0.073	0.193
Interstates density (ln)	0.366	9.537	0.016	0.761	0.274	5.502	-0.206	0.760
Principal density (ln)	0.096	7.593	-0.026	0.218	0.084	4.363	-0.102	0.269
Minor arterials density (ln)	0.040	3.287	-0.078	0.160	0.003	0.151	-0.190	0.195
Collectors density (ln)	-0.004	-0.328	-0.126	0.118	-0.045	-2.210	-0.241	0.151
Local roads density (ln)	-0.004	-0.100	-0.393	0.423	0.028	0.486	-0.535	0.582
Constant	4.330	36.375	3.233	5.533	1.934	12.636	0.482	3.484
Spatial Correlation (phi)	-0.011	-5.306	-0.032	0.009	-0.041	-7.992	-0.096	0.000
Observations	9740				9740			
Df	9729				9729			
Log likelihood	-49289.6				-18363.4			

Table 8: Spatial MCMC Models (Fatal Injury Crashes, Entire network)

	Model **3** (S-MCMC-3)			
FATAL & INJURY CRASHES	**Coefficient**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Population Density (ln)	-0.419	-51.341	-0.498	-0.340
Employment Density (ln)	0.205	45.810	0.162	0.248
Median Income (ln)	0.032	2.935	-0.080	0.134
Interstates density (ln)	0.464	12.391	0.123	0.848
Principal arterials density (ln)	0.165	12.810	0.041	0.291
Minor arterials density (ln)	0.097	7.692	-0.025	0.219
Collectors density (ln)	0.048	3.764	-0.076	0.173
Local roads density (ln)	0.042	1.008	-0.344	0.462
Constant	4.023	31.605	2.855	5.314
Spatial Correlation (phi)	-0.010	-4.478	-0.030	0.011
Observations	9740			
Df	9729			
Log likelihood	-43358.8			

Table 7 and Table 8 show the effects of the contextual variables for population density, employment density and median income, along with the variables for road density for each functional classification, on total crashes, fatal and major injury crashes and fatal and injury crashes. In all three models in Table 7 and Table 8, population density has negative associations, in line with the Ewing, et al. (2003), and Ewing and Dumbaugh (2009) studies that show that sprawl is associated with a higher risk of crash occurrence. Employment density has positive association in all three models. This might be indicative of office parks or campuses and other such land uses which tend to be located along fast-moving highways, and feature less density and pedestrian activity. Median income shows positive associations through all three models, but with relatively low significance levels.

Table 9 shows the results for the total crashes, fatal and major injury crashes and fatal and injury crashes models specified with only geometric variables, and then with both geometric and contextual variables. While all the other models displayed in this section are MCMC models, the Table 9 models below are MLE models. The reason I display these results is that I ran into some difficulties using *CrimeStat* to estimate the MCMC models potentially because of their complexity. In this case, each model had between five and eight variables with 112,502 observations and the models did not converge after 120 hours of running.

Table 9: Negative Binomial Models with Geometric Variables and then with all Variables

	Model **4** (L-MLE-1)	Model **5** (C-MLE-1)	Model **6** (L-MLE-2)	Model **7** (C-MLE-2)	Model **8** (L-MLE-3)	Model **9** (C-MLE-3)
VARIABLES	**Crashes**	**Crashes**	**Fatal & Major Injury Crashes**	**Fatal & Major Injury Crashes**	**Fatal & Injury Crashes**	**Fatal & Injury Crashes**
Total width (ln)	1.538***	0.465***	0.598***	0.189***	1.688***	0.485***
	(0.0156)	(0.0190)	(0.0319)	(0.0359)	(0.0177)	(0.0200)
Lane count (ln)	-1.597***	-0.352***	-0.400***	-0.00872	-1.632***	-0.345***
	(0.0297)	(0.0347)	(0.0629)	(0.0657)	(0.0338)	(0.0368)
Median width (ln)	-0.103***	-0.0947***	-0.123***	-0.118***	-0.116***	-0.118***
	(0.00264)	(0.00349)	(0.00626)	(0.00719)	(0.00299)	(0.00396)
Vmt (ln)	0.595***	0.516***	0.550***	0.535***	0.594***	0.517***
	(0.00255)	(0.00555)	(0.00658)	(0.00719)	(0.00304)	(0.00560)
Sinuosity (ln)	-0.344***	-0.115*	0.0356	0.169	-0.222***	0.0714
	(0.0479)	(0.0615)	(0.128)	(0.128)	(0.0576)	(0.0645)
Median income (ln)		0.0249***		-0.201***		-0.0938***
		(0.00926)		(0.0149)		(0.00886)
Population density (sq mi, ln)		0.125***		0.102***		0.162***
		(0.00478)		(0.00975)		(0.00522)
Employment density (sq mi, ln)		0.163***		0.0156**		0.153***
Constant	-5.555***	-4.451***	-7.210***	-4.506***	-6.878***	-4.233***
	(0.0494)	(0.116)	(0.117)	(0.201)	(0.0576)	(0.115)
	-0.266***	-0.583***	-0.530***	-0.629***	-0.110***	-0.456***
	(0.00573)	(0.00803)	(0.0347)	(0.0382)	(0.00655)	(0.00882)
Observations	112,502	112,502	112,502	112,502	112,502	112,502
Log likelihood	-295940	-284501	-62074	-61738	-236163	-226538
Ll Constant Only	-339565	-339565	-69493	-69493	-272339	-272339
LR Chi2	87251	132864	14837	14145	72353	103928
Pseudo_R2	0.128	0.162	0.107	0.112	0.133	0.168
Standard errors in parentheses						
*** p<0.01, ** p<0.05, * p<0.1						

Interestingly, the population density coefficients change direction of association compared to all three models in Table 7 and Table 8 where they are specified without geometric and traffic volume variables, to Model **5** (C-MLE-1), Model **7** (C-MLE-2) and Model **9** (C-MLE-3) in Table 9 where both groups of variables are specified. Median income also undergoes a similar change, having a negative association in both models

Model **7** (C-MLE-2) and Model **9** (C-MLE-3) in Table 9 where they are specified with geometric and traffic volume variables, but a positive association when specified without, in Model 1 (S-MCMC-1) , Model 2, and Model 3 of Table 7 and Table 8. These results may be indicative of omitted variable bias in the Table 7 and Table 8 models where geometric variables are excluded. One important caveat to note is that Model 1 (S-MCMC-1), Model 2, and Model 3 are not strictly comparable to Model **5** (C-MLE-1), Model **7** (C-MLE-2) and Model **9** (C-MLE-3) because the former models had 9740 observations, whereas the latter models had 112,502 observations. The contextual variables used in Model **5** (C-MLE-1), Model **7** (C-MLE-2) and Model **9** (C-MLE-3) were aggregated by distance and assigned to the 112,502 road segments, as described in the *Methods* section, and as such, are not entirely in the same form as the contextual variables in Model 1 (S-MCMC-1), Model 2, and Model 3. These results, especially the change in direction of the coefficients of population density and median income, show how various modeling decisions such as the decision to model only contextual variables, only linked based variables, or to combine both kinds of variables, can affect results and in turn support certain road safety decisions over others. The method an analyst chooses has a direct effect on the results obtained.

The corresponding MLE models to Model 1 (S-MCMC-1), Model 2, and Model 3 are displayed in Table 47, in Appendix B. In comparing Model 1 (S-MCMC-1), Model 2, and Model 3 with the Table 47 models, I found that the results largely remained the same across all the dependent variables, for all the independent variables.

In Table 9, all the coefficients for the six models displayed are statistically significant, except for sinuosity in Model **6** (L-MLE-2), both lane count and sinuosity in Model **7** (C-MLE-2), and sinuosity in Model **9** (C-MLE-3). Interestingly, while the coefficient for lane count was statistically significant at the 95% confidence level in Model **6** (L-MLE-2), it decreased considerably in magnitude and lost statistical significance with the addition of contextual variables in Model **7** (C-MLE-2). A similar change occurred with the sinuosity coefficient, in addition to a change in its direction of association from Model **8** (L-MLE-3) to Model **9** (C-MLE-3).

The first research question is whether there is specification error in models that omit contextual variables. At first glance, the modeling results show that the coefficients for the geometric variables decrease in magnitude when contextual variables are added, as can be seen in Table 9. These results would appear to indicate the presence of omitted variable bias in the geometric and traffic volume coefficients of Model **4** (L-MLE-1), Model **6** (L-MLE-2), and Model **8** (L-MLE-3) since omitted variable bias can be indicated by inflated coefficient magnitudes (such coefficients carry the effects of variables that are omitted). More indicative of this condition are the lane count coefficient in Model **7** (C-MLE-2), and sinuosity in Model **9** (C-MLE-3), which also both lose significance upon the addition of contextual variables to the models and change direction of association in the case of the sinuosity variable in Model **9** (C-MLE-3). Again, the absence of contextual variables may be responsible for the initial appearance of significance of lane count and sinuosity, and the initial direction of association of sinuosity. The results of these models have therefore not refuted my hypothesis.

Another factor that points to better specification in the combined models, compared with the link-based models in Table 9 is the Pseudo R^2 measure of fit. All three combined models show higher Pseudo R^2 values than their link-based model counterparts. While the results of the model comparisons generally indicate the presence of omitted variable bias, there is still the possibility that some of the bias seen is due to another factor such as spatial correlation. As I was unable to run corresponding MCMC models for the Table 9 models, I test this possibility in later sections of this chapter.

Crash frequency analysis for Pennsylvania principal arterials

The next stage of my analysis is to see if for my link-based and combined MLE models, there is a problem with spatial correlation by using conditional autoregressive models. Since I was unable to run conditional autoregressive models using the entire dataset, I chose to run the models based on subsets. I ran the models at this stage by the functional classifications of the roads being analyzed, so that for each dependent variable, there were multiple models for each functional classification. Not all functional classifications had enough crash occurrences per road segment for the MCMC models to run successfully. This is another documented issue with *CrimeStat* software (Levine 2013b, Levine 2013a). I was able to successfully run MCMC models for the two functional classifications of principal arterials (functional class B) and local roads (functional class E). The results are displayed below, while the MLE model results are displayed in the appendix.

Table 10: Total Crashes MCMC Models with only Principal Arterials (FC B)

	Model 10 (L-MCMC-1)				Model 11 (C-MCMC-1)			
CRASHES	**Coefficient**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**	**Coefficient**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Total width (ln)	1.209	34.56	0.754	1.67	0.426	13.865	0.021	0.831
Lane count (ln)	-0.488	-7.656	-1.324	0.344	0.211	3.876	-0.505	0.928
Median width (ln)	-0.113	-23.48	-0.175	-0.049	-0.11	-26.528	-0.163	-0.055
VMT (ln)	0.406	60.485	0.318	0.493	0.362	59.08	0.281	0.442
Sinuosity (ln)	0.078	0.715	-1.415	1.45	0.582	6.177	-0.689	1.791
Median income (ln)					0.034	2.417	-0.158	0.217
Population density (sq mi, ln)					0.225	25.861	0.111	0.34
Employment density (sq mi, ln)					0.142	25.181	0.068	0.216
Constant	-3.935	-37.589	-5.288	-2.548	-4.338	-23.424	-6.74	-1.849
Spatial Correlation (Phi)	-0.08	-11.3	-0.179	-0.009	-0.027	-8.376	-0.07	0.011
Observations	17859				17859			
Df	17851				17848			
Log likelihood	-63680				-60285			

Table 11: Fatal & Major Injury Crashes MCMC Models with only Principal Arterials (FC B)

	Model 12 (L-MCMC-2)				Model 13 (C-MCMC-2)			
FATAL & MAJOR INJURY CRASHES	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Total width (ln)	0.721	11.869	-0.076	1.475	0.341	5.49	-0.514	1.152
Lane count (ln)	-0.014	-0.119	-1.498	1.606	0.097	0.869	-1.349	1.599
Median width (ln)	-0.005	-0.095	-0.312	0.751	-0.149	-14.39	-0.288	-0.017
VMT (ln)	0.434	24.204	0.197	0.646	0.47	30.571	0.276	0.681
Sinuosity (ln)	0.878	2.919	-2.62	5.002	1.171	4.346	-2.073	4.933
Median income (ln)					-0.251	-8.422	-0.643	0.113
Population density (sq mi, ln)					0.161	8.925	-0.074	0.399
Employment density (sq mi, ln)					0.023	1.906	-0.133	0.18
Constant	-8.036	-21.67	-13.377	-4.493	-4.672	-12.11	-9.488	0.394
Spatial Correlation (Phi)	-8.029	-4.403	-37.24	-0.038	-0.188	-12	-0.395	-0.031
Observations	17859				17859			
Df	17851				17848			
Log likelihood	-73376				-14792			

Table 12: Fatal & Injury Crashes MCMC Models with only Principal Arterials (FC B)

	Model **14** (L-MCMC-3)				Model **15** (C-MCMC-3)			
FATAL & INJURY CRASHES	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Total width (ln)	1.42	37.965	0.937	1.919	0.407	11.694	-0.043	0.858
Lane count (ln)	-0.529	-7.843	-1.414	0.355	0.278	4.668	-0.498	1.048
Median width (ln)	-0.164	-29.652	-0.236	-0.091	-0.136	-28.909	-0.196	-0.074
VMT (ln)	0.422	53.267	0.318	0.525	0.367	50.691	0.273	0.461
Sinuosity (ln)	0.432	3.419	-1.265	2.07	1.107	9.646	-0.383	2.598
Median income (ln)					-0.013	-0.851	-0.221	0.19
Population density (sq mi, ln)					0.225	22.49	0.095	0.355
Employment density (sq mi, ln)					0.189	29.195	0.105	0.273
Constant	-5.538	-46.742	-7.08	-3.984	-5.11	-24.272	-7.868	-2.304
Observations	17859				17859			
Df	17851				17848			
Log likelihood	-53375				-50084			

In comparing the link-based models for total crashes, fatal and major injury crashes, and fatal and injury crashes, which are Model 10 (L-MCMC-1), Model 12 (L-MCMC-2), and Model 14 (L-MCMC-3) respectively, to their corresponding combined models which are Model 11 (C-MCMC-1), Model 13 (C-MCMC-2), and Model **15** (C-MCMC-3), a number of noteworthy changes are present. First is that compared with Table 9, the reduction in magnitude of geometric variables from the link-based models to the combined models is less consistent. For example, from Model 10 (L-MCMC-1) to Model 11 (C-MCMC-1), median width does not change in coefficient magnitude, and the sinuosity coefficient actually increases from 0.078 to 0.582. From Model 12 (L-MCMC-2) to Model 13 (C-MCMC-2), the lane count coefficient increases by a relatively large magnitude, from 0.014 to 0.097, while the sinuosity coefficient increases from 0.878 to 1.171. In addition to this, the expected reduction in magnitude from the link-based model coefficients to the combined model coefficients is not really seen when comparing Model 12 (L-MCMC-2) to Model 13 (C-MCMC-2), except in the total width variable. All the geometric variable coefficients unexpectedly show an increase in magnitude, with the exception of total width. For most of these geometric variables, the increases in magnitude are not large. From Model 14 (L-MCMC-3) to Model **15** (C-MCMC-3), the sinuosity coefficient increases. One reason why there is consistent increase in magnitude of geometric variable coefficients from Model 12 (L-MCMC-2) to Model 13 (C-MCMC-2) may be because of the zero-inflated nature of the fatal and major injury variable, which can cause problems in modeling. The fatal and major injury variable can be described as zero-inflated because the vast majority of observations have zero values for fatal and major injuries. This means that accidents involving fatalities or major injuries were rare. The problem with this is that there is not enough observations of fatal or major injuries to model results that are stable (Levine 2013b, Levine 2013a). Aside from these exceptions, there is generally a reduction in coefficient magnitude from the link-based to the combined models for the other two dependent variables. Increases in magnitude of the geometric variables may also be indicative of an interaction between a contextual variable and these variables. I explore possible variable interactions in the next chapter.

Another change of note is that the lane count coefficient consistently changes direction of association, from Model 10 (L-MCMC-1) to Model 11 (C-MCMC-1), Model 12 (L-MCMC-2) to Model 13 (C-MCMC-2) and Model 14 (L-MCMC-3) to Model 15 (C-MCMC-3). While only Model 12 (L-MCMC-2) to Model 13 (C-MCMC-2) shows an increase in magnitude, Model 10 (L-MCMC-1) to Model 11 (C-MCMC-1) and Model 14 (L-MCMC-3) to Model 15 (C-MCMC-3) show magnitude reductions and changes from a negative association with crash frequency to a positive association with crash frequency. The change in direction of association of lane count, from negative to positive from all three link-based models to their corresponding combined models agrees with Sawalha, Sayed (2001). This change in direction from the link-based to the combined models suggests that increasing the number of lanes on principal arterials in the Pennsylvania roadway system, is actually associated with more crash occurrences, contrary to what would be the assumption, when inferring from models that have excluded contextual variables and that this may not be a general association, applying specifically to principal arterials. This finding has implications for both methodology and theory in crash frequency studies. It appears to underscore the importance of specifying crash frequency models with contextual variables as well as highlights the error in the assumption of sound inferences from generalizing between differing contexts such as road types, urban versus rural and from one locality to another.

The spatial correlation coefficient is found to be significant in both Model 10 (L-MCMC-1) and Model 11 (C-MCMC-1), with a large reduction in magnitude in Model 11 (C-MCMC-1), indicating that spatial correlation has been accounted for in this model, to a notable degree, when compared with Model 10 (L-MCMC-1).

In examining the corresponding MLE models (

Table 48, appendix B), there are no notable changes in magnitude or direction of association of the geometric variables that were not seen in the MCMC models (Model 10 (L-MCMC-1) through Model 15 (C-MCMC-3)), suggesting that spatial correlation is not a significant issue and potentially ruling this out as the cause of any bias first seen in the MLE models. Comparing from Model 10 (L-MCMC-1) however, the sinuosity coefficient gains significance in Model 45 (L-MLE-1) in

Table 48. This could be an indication that in predicting total crashes on principal arterials for Pennsylvania state roads, sinuosity is not actually a significant variable, but

only appears to be so, due to the influence of spatial correlation. Again, my hypothesis that the omission of spatial variables is the cause of the bias seen is not refuted.

Comparing the models for fatal and major injury crashes in Table 11 (Model 12 (L-MCMC-2) and Model 13 (C-MCMC-2)), with their corresponding MLE models (Model 47 and Model 48), again there are not many important changes. For the link-based models, there are no notable changes from Model 47 (L-MLE-2) to Model 12 (L-MCMC-2). The exception is the median width coefficient which loses significance from Model 47 (L-MLE-2) to Model 12 (L-MCMC-2) and reduces notably in magnitude. In comparing Model 48 (C-MLE-2) to Model 13 (C-MCMC-2), the lane count coefficient loses significance. Finally, for the fatal and injury crashes, comparing Model 49 (L-MLE-3) and Model 50 (C-MLE-3) to Model 14 (L-MCMC-3) and Model 15 (C-MCMC-3) in Table 12 also shows no important changes from the MLE models to the MCMC models, except for the loss of significance of the median income coefficient from Model 50 (C-MLE-3) to Model 15 (C-MCMC-3).

While maximum likelihood methods show significance at a certain confidence level, Bayesian methods like the MCMC estimations I carry out using *CrimeStat* include a credible interval, which shows the range of coefficient estimates that the correct estimate is within, to a certain probability level. This means that for example, if the result shows a 95% credible interval, there is a 95% probability that the correct coefficient is in that range. This is very useful, because it makes it possible to see at a glance, what variables are not well suited to deriving CMFs, due to the wide range of their credible intervals. Several variables in both Model 10 (L-MCMC-1) and Model 11 (C-MCMC-1) have very wide credible intervals, some ranging from negative values on their lower limit to positive on their upper limit. This illustrates the point made earlier about the fact that point estimates such as those prescribed as part of the *Highway Safety Manual* can be misleading.

Crash frequency analysis for Pennsylvania local roads

The following are the results for local roads (functional class E). It is useful to observe changes between link-based and combined models for this road classification as well, to see if the results seen for principal arterials are similar and to explore the ways in which they might differ.

Table 13: Total Crashes MCMC Models [Local Roads (FCE)]

	Model 16 (L-MCMC-1)				Model 17 (C-MCMC-1)			
CRASHES	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Total width (ln)	1.746	23.626	0.772	2.796	1.401	19.906	0.455	2.275
Lane count (ln)	-2.224	-10.717	-4.925	0.397	-1.245	-5.776	-3.845	1.549
Median width (ln)	-0.287	-20.735	-0.472	-0.093	-0.224	-17.619	-0.392	-0.050
VMT (ln)	0.772	96.166	0.663	0.884	0.658	79.382	0.547	0.770
Sinuosity (ln)	-2.053	-8.466	-5.337	1.181	-1.070	-4.862	-3.969	1.751
Median income (ln)					0.260	9.029	-0.113	0.627
Population density (sq mi, ln)					0.055	3.247	-0.173	0.287
Employment density (sq mi, ln)					0.136	15.335	0.018	0.256
Constant	-5.394	-19.876	-8.832	-1.634	-8.853	-24.141	-13.312	-4.377
Spatial Correlation (Phi)	-0.031	-8.057	-0.090	0.015	-0.056	-9.232	-0.148	0.009
Observations	18965				18965			
Df	18957				18954			
Log likelihood	-29509				-28955			

Table 14: Fatal & Major Injury Crashes MCMC Models [Local Roads (FCE)]

	Model **18** (L-MCMC-2)				Model **19** (C-MCMC-2)			
FATAL & MAJOR INJURY	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Total width (ln)	0.560	4.561	-0.964	2.159	1.433	7.950	-0.936	3.685
Lane count (ln)	-3.870	-10.080	-8.876	0.776	-0.922	-1.552	-8.798	6.356
Median width (ln)	-0.267	-8.147	-0.734	0.159	-0.611	-7.947	-1.719	0.314
VMT (ln)	0.693	33.350	0.428	0.991	0.715	25.888	0.374	1.121
Sinuosity (ln)	2.043	3.474	-5.128	9.534	2.589	3.704	-5.703	11.716
Median income (ln)					-0.477	-7.475	-1.234	0.242
Population density (sq mi, ln)					0.118	1.865	-0.697	0.996
Employment density (sq mi, ln)					-0.020	-0.640	-0.465	0.394
Constant	-5.407	-10.898	-11.329	0.882	-7.262	-9.140	-16.515	1.789
Spatial Correlation (Phi)	-0.056	-7.348	-0.185	0.025	-5.640	-6.050	-18.846	0.014
Observations	18965				18965			
Df	18957				18954			
Log likelihood	-5227.4				-14226			

Table 15: Fatal & Injury Crashes MCMC Models [Local Roads (FCE)]

	Model 20 (L-MCMC-3)				**Model 21** (C-MCMC-3)			
FATAL & INJURY CRASHES	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Total width (ln)	1.569	19.276	0.518	2.636	1.264	15.919	0.240	2.320
Lane count (ln)	-3.222	-13.628	-6.213	-0.234	-2.192	-9.805	-5.080	0.585
Median width (ln)	-0.308	-18.509	-0.530	-0.078	-0.322	-20.181	-0.535	-0.104
VMT (ln)	0.770	79.392	0.643	0.905	0.697	67.083	0.561	0.843
Sinuosity (ln)	-1.078	-3.818	-4.722	2.657	-0.366	-1.304	-3.980	3.373
Median income (ln)					-0.012	-0.361	-0.424	0.374
Population density (sq mi, ln)					0.102	4.973	-0.175	0.377
Employment density (sq mi, ln)					0.097	8.838	-0.051	0.244
Constant	-5.110	-16.145	-9.098	-1.015	-5.847	-14.489	-10.544	-0.819
Spatial Correlation (Phi)	-0.042	-8.797	-0.116	0.011	-0.037	-7.686	-0.115	0.015
Observations	18965				18965			
Df	18957				18954			
Log likelihood	-21417				-21187			

A comparison of the variable coefficients in Table 13, Table 14, and Table 15, from the link-based MCMC models (Model 16 (L-MCMC-1), Model 18 (L-MCMC-2), and Model 20 (L-MCMC-3)), to the combined MCMC models (Model 17 (C-MCMC-1), Model 19 (C-MCMC-2), and Model 21 (C-MCMC-3)) shows that the magnitude reduction from the link-based models to the combined models is more consistent than in the principal arterial models, but still not as consistent as in the models with the entire road network. All the geometric variable coefficients in Model 16 (L-MCMC-1) decreased in magnitude in Model 17 (C-MCMC-1), but in the fatal and major injury models (Table 14), all the geometric variable coefficients except lane count, increase in magnitude from Model 18 (L-MCMC-2) to Model 19 (C-MCMC-2). This increase in coefficient magnitude from link-based to combined model was also seen in the fatal and major injury models for principal arterials (Table 11), supporting my theory that it occurs due to the zero-inflated nature of the fatal and major injury dependent variable. In the fatal and injury models (Table 15), again all the coefficients decrease in magnitude from Model 20 (L-MCMC-3) to Model 21 (C-MCMC-3), except for the median width variable which has a very negligible increase of 0.014.

In the local road models, there is no change in direction of coefficients from the link-based models to the combined models. The negative association of lane count with crash frequency occurs with local roads, just as in the models that included the entire road network (Table 9). This means that the positive association of lane count on crash frequency is not universal to all road types since it was seen only in the models for principal arterial roads.

Notably, in the combined MCMC Model 21 (C-MCMC-3) (fatal and injury models), sinuosity is not found to be significant, and in the combined MCMC Model 19 (C-MCMC-2) (fatal and major injury models), lane count is not found to be significant. All the geometric variable coefficients in the total crashes models are significant. The significance of sinuosity in Model 21 (C-MCMC-3) and lane count in Model 19 (C-MCMC-2) may have been affected by the fact that the fatal and major injury and the fatal and injury models have more observations with zero crash occurrences. The other similarity between the loss of significance in Model 19 (C-MCMC-2) and in Model 21 (C-MCMC-3) besides their exclusion of property-damage crashes is that they are both combined models. In their corresponding link-based models (Model 18 (L-MCMC-2) and Model 20 (L-MCMC-3)), lane count and sinuosity respectively have significant coefficients. A similar result is seen in Table 9 (lane count loses significance from Model 6 (L-MLE-2) to Model 7 (C-MLE-2), and sinuosity loses significance from Model 8 (L-MLE-3) to Model 9 (C-MLE-3)), but not in

Table 48, suggesting the relative unimportance of lane count to fatal and major injury crashes, and sinuosity to fatal and injury crashes in the Pennsylvania road network , with the exception of principal arterials.

The results for the corresponding MLE models are displayed in Table 49 in appendix B. For the total crashes models, when comparing the MCMC models (Model 16

(L-MCMC-1) and Model 17 (C-MCMC-1)) with the MLE models (Model **51** and Model **52**), it is immediately apparent that the variables remain quite consistent. Pavement width, and vehicle miles traveled had positive and significant associations with crashes in both the link-based MLE model, and in the link-based MCMC model. Lane count, median width and sinuosity had negative and significant associations in both models. For the combined MLE and MCMC models, pavement width, lane count, median width, vehicle miles traveled, and sinuosity have significance and the same direction of association seen in the link-based models. Population density, employment density, and median income, did not change direction of association or lose significance from the MLE combined model to the MCMC combined model. These results indicate that spatial correlation is not an important factor in any bias seen in the MLE models. All five geometric variables did have reductions in magnitude of association from the link-based models compared to the combined models, indicating some bias in their coefficients, though not as much bias as is indicated by a change in direction of association, as seen in the models with only principal arterials, and in the models with the entire road network. The negative association of lane count on crashes is contrary to the finding made by Sawalha, Sayed (2001).

Conclusions

The goal of this chapter was to address my first research question. It asks whether there is evidence for specification error in crash frequency modeling, brought about by the omission of contextual variables. In addressing this question, I have examined the magnitude, direction and significance of the coefficients for geometric variables specified in MLE and MCMC models, and the magnitude, direction and significance of the coefficients of the same variables specified along with contextual variables (combined models). I examined these coefficients first for the entire road network, then for just principal arterials in the road network and finally for just local roads in the network. I specified models for total crashes, crashes with fatalities and only major injuries, and then for crashes with fatalities and any kind of injury. In addition to this, I looked at other indicators that I derived from examining the above models. They include the pseudo R2 measure of fit of the MLE models, and the spatial correlation coefficient (Phi) of the MCMC models.

The point of the linked based MLE models is to see if the geometric variables show a reduction in magnitude or a change in direction or significance of their coefficients when the link-based models are compared with the combined MLE models. For the MCMC models I also examine changes seen from the MLE link-based to the MLE combined models, and whether they are present when MCMC link-based models are compared with MCMC combined models. The result of the ability to confirm the MLE models with the MCMC models is the indication that any bias seen in the link-based models are not a result of spatial correlation. Another way to say this is that the ability to confirm the MLE model results will increase confidence in the conclusion that there is specification error in models that exclude contextual variables.

In the first set of models, based on the entire road network (Table 9), the vast majority of geometric variables show a decrease in magnitude of effect on all dependent variables.

Table 16: Total Crashes Summary (Entire Network)

	Model 4 (L-MLE-1)	Model 5 (C-MLE-1)
Pavement width	+ sig	+ sig ↓
Lane count	- sig	- sig ↓
Median width	- sig	- sig ↓
Vehicle miles traveled	+ sig	+ sig ↓
Sinuosity	- sig	- sig ↓
Population density		+ sig
Employment density		+ sig
Median income		+ sig

+: Direction of association Sig: Significant ↓ Direction of change
"ns": not significant at 95% confidence level

Table 17: Fatal & Major Injury Crashes Summary (Entire Network)

	Model 6 (L-MLE-2)	Model 7 (C-MLE-2)
Pavement width	+ sig	+ sig ↓
Lane count	- sig	- sig ↓
Median width	- sig	- sig ↓
Vehicle miles traveled	+ sig	+ sig ↓
Sinuosity	+ ns	+ ↑
Population density		+ sig
Employment density		+ sig
Median income		- sig

+: Direction of association Sig: Significant ↓ Direction of change
"ns": not significant at 95% confidence level

Table 18: Fatal & Injury Crashes Summary (Entire Network)

	Model 8 (L-MLE-3)		Model 9 (C-MLE-3)		
Pavement width	+	sig	+	sig	↓
Lane count	-	sig	-	sig	↓
Median width	-	sig	-	sig	
Vehicle miles traveled	+	sig	+	sig	↓
Sinuosity	-	sig	+	ns	↓
Population density			+	sig	
Employment density			+	sig	
Median income			-	sig	

+: Direction of association Sig: Significant ↓ Direction of change
"ns": not significant at 95% confidence level

The only exception was sinuosity in the fatal and major injury crashes model which showed an increase but with an insignificant coefficient. Virtually all the coefficients were significant except sinuosity in the fatal and major injury crashes models, lane count in the combined MLE model and sinuosity in the combined MLE model for fatal and all injury crashes. For all the combined MLE models in this set, there were slightly higher pseudo R2 values than their corresponding link-based models. I was not able to produce results for the MCMC models for this set because *CrimeStat* failed to run them even after 120 hours of attempting. This may be due to the combination of a large amount of observations and many variables. As a work- around, I ran the models by functional classification groups.

In the second set of MLE models- principal arterials in the Pennsylvania road network (

Table 48), the majority of the geometric variables saw a decrease in magnitude of association for all three dependent variables.

Table 19: Total Crashes Summary (Principal Arterials)

	Model **45** (L-MLE-1)	Model 46 (C-MLE-1)	**Model 10** (L-MCMC-1)	Model 11 (C-MCMC-1)
Pavement width	+ sig	+ sig ↓	+ sig	+ sig ↓
Lane count	- sig	+ sig ↓	- sig	+ sig ↓
Median width	- sig	- sig	- sig	- sig
Vehicle miles traveled	+ sig	+ sig ↓	+ sig	+ sig ↓
Sinuosity	+ sig	+ sig ↑	+ ns	+ sig ↑
Population density		+ sig		+ sig
Employment density		+ sig		+ sig
Median income		+		+ sig

+: Direction of association Sig: Significant ↓ Direction of change
"ns": not significant at 95% confidence level

Table 20: Fatal & Major Injury Crashes (Principal Arterials)

	Model 47 (L-MLE-2)	Model 48 (C-MLE-2)	Model 12 (L-MCMC-2)	Model 13 (C-MCMC-2)
Pavement width	+ sig	+ sig ↓	+ sig	+ sig ↓
Lane count	- ns	+ sig ↑	- ns	+ ns ↑
Median width	- sig	- sig ↓	- ns	- sig ↑
Vehicle miles traveled	+ sig	+ sig	+ sig	+ sig
Sinuosity	+ sig	+ sig ↑	+ sig	+ sig ↑
Population density		+ sig		+ sig
Employment density		+ ns		+ ns
Median income		- sig		- sig

+: Direction of association Sig: Significant ↓ Direction of change
"ns": not significant at 95% confidence level

Table 21: Fatal & Injury Crashes (Principal Arterials)

	Model 49 (L-MLE-3)	Model 50 (C-MLE-3)	Model 14 (L-MCMC-3)	Model 15 (C-MCMC-3)
Pavement width	+ sig	+ sig ↓	+ sig	+ sig ↓
Lane count	- sig	+ sig ↓	- sig	+ sig ↓
Median width	- sig	- sig	- sig	- sig
Vehicle miles traveled	+ sig	+ sig	+ sig	+ sig
Sinuosity	+ sig	+ sig ↑	+ sig	+ sig ↑
Population density		+ sig		+ sig
Employment density		+ sig		+ sig
Median income		- sig		-

+: Direction of association Sig: Significant ↓ Direction of change
"ns": not significant at 95% confidence level

The exceptions are again sinuosity for all the three dependent variables, and lane count for fatal and major injury crashes. All coefficients were significant except for lane count in the MLE link-based fatal and major injury crashes model, which gained significance in the MLE combined model. In all three MLE models, lane count changed direction of association from negative in the link-based models to positive in the combined models. The MCMC models (Table 10, Table 11, and Table 12) largely replicate these results, except on a few points. In the fatal and major injury MCMC models (Table 11), lane count is not significant in either the link-based or the combined model. Median width is also not significant in the link-based model and increases in magnitude in the combined MCMC model. The improvement in the pseudo R2 measure of fit in the combined MLE models for all three dependent variables is larger than the improvement seen in the models with the entire road network. The largest was by almost 6% in the fatal and injury model followed by the total crashes model.

In the local roads models, the vast majority of geometric variables for all three dependent variables decreased in magnitude of association with the exception of

Table 22: Total Crashes Summary (Local Roads)

	Model 51 (L-MLE-1)		Model 52 (C-MLE-1)			Model 16 (L-MCMC-1)		Model 17 (C-MCMC-1)		
Pavement width	+	sig	+	sig	↓	+	sig	+	sig	↓
Lane count	-	sig	-	sig	↓	-	sig	-	sig	↓
Median width	-	sig	-	sig	↓	-	sig	-	sig	↓
Vehicle miles traveled	+	sig	+	sig	↓	+	sig	+	sig	↓
Sinuosity	-	sig	-	sig	↓	-	sig	-	sig	↓
Population density			+	sig				+	sig	
Employment density			+	sig				+	sig	
Median income			+					+	sig	

+: Direction of association Sig: Significant ↓ Direction of change "ns": not significant at 95% confidence level

Table 23: Fatal & Major Injury Crashes Summary (Local Roads)

	Model 53 (L-MLE-2)	Model 54 (C-MLE-2)	Model 18 (L-MCMC-2)	Model 19 (C-MCMC-2)
Pavement width	+ sig	+ sig ↓	+ sig	+ sig ↑
Lane count	- sig	- sig	- sig	- ns ↓
Median width	- sig	- sig	- sig	- sig ↓
Vehicle miles traveled	+ sig	+ sig	+ sig	+ sig
Sinuosity	+ ns	+ ns ↓	+ sig	+ sig
Population density		- ns		+ ns
Employment density		+ ns		- ns
Median income		- sig		- sig

+: Direction of association Sig: Significant ↓ Direction of change
"ns": not significant at 95% confidence level

Table 24: Fatal & Injury Crashes (Local Roads)

	Model 55 (L-MLE-3)	Model 56 (C-MLE-3)	Model 20 (L-MCMC-3)	Model 21 (C-MCMC-3)
Pavement width	+ sig	+ sig ↓	+ sig	+ sig ↓
Lane count	- sig	- sig ↓	- sig	- sig ↓
Median width	- sig	- sig	- sig	- sig
Vehicle miles traveled	+ sig	+ sig ↓	+ sig	+ sig ↓
Sinuosity	- sig	- ns ↓	- sig	- ns ↓
Population density		+ sig		+ sig
Employment density		+ sig		+ sig
Median income		+ ns		-

+: Direction of association Sig: Significant ↓ Direction of change
"ns": not significant at 95% confidence level

pavement width in the fatal and major injury MCMC model (Table 14), which increased in magnitude. Sinuosity in the fatal and major injury MLE combined and link-based models (Model **53** and Model **54**) did not have significance, or in the fatal and injury combined MLE model (Model **56**). Notably, lane count, the only variable to have changed in direction of association (in the principal arterial models, from negative to positive), is negatively correlated with the dependent variables as was the case in the models with the entire road network. The MCMC models (Table 13, Table 14, and Table

15), largely confirm the results from the MLE models (Table 49), sinuosity gains significance in both the link-based and combined fatal and major injury models, but remains insignificant in the combined fatal and injury model. There were small improvements in the pseudo R2 measure of fit of the combined MLE models over the link-based MLE models.

Data quality is an issue in this dataset. I address this issue of data quality in more depth in the next chapter. It is however important to note that this issue should be kept in mind while interpreting model results, since not all data for dependent or independent variables are equal in quality. An example is the total crashes dependent variable. Property damage crashes, which tend to be under-reported to a large degree are a constituent of total crashes, but not of fatal and major injury crashes. Property damage crashes are frequently under-reported because of their relative unimportance, compared to crashes with fatalities and major injuries. Fatal and major injury data tends to be of better quality than total crashes data (Scribner 1994).

On the whole, the re-estimation of the associations of the various variables on crashes, fatal and major injury crashes, and fatal and all injury crashes, using negative binomial autoregressive models so as to be able to control for spatial correlation indicates that the biases seen in the link-based models of the initial estimation are not due to spatial correlation. This means that both overdispersion and spatial correlation have been eliminated as possible explanations and the hypothesis that omitted variable bias is a problem affecting conventional crash prediction models is supported for principal arterials and local roads in the Pennsylvania road system, although the support is strongest for principal arterials, since these models had changes in direction of association of at least one geometric variable in the combined models.

Crash Frequency Analysis for North Carolina State Roads

In this chapter, I visit the question of the impact of measurement error on statistical inferences made from crash frequency prediction, using a dataset containing the attributes of North Carolina roads. By measurement error, I am referring to the introduction of erroneous data through the methods of analysis chosen by the researcher. When analysts are faced with data availability and quality problems, they must work around it in their analysis, and it is in this process that a source of measurement error can be introduced. For this chapter, I examine the impact of measurement error introduced through a process of estimating AADT for observations without AADT data and show how indeterminacy can be caused by the research methods chosen. I compare the results from a crash frequency analysis carried out using a dataset with observed AADT data ("observed" because it was calculated from observed hourly traffic volume data) for all road segments in the dataset, with the results of a crash frequency analysis carried out using a dataset for which some of the observations, chosen randomly, have AADT data estimated. These estimations were made from those observations with observed AADT data. As discussed above, estimated AADT is assumed to contain some error because it is not actually observed, but is only the "best guess" from what has actually been observed. My goal is to assess the impact of "erroneous" AADT (AADT data with measurement error) on inferences. Comparisons made between these two sets of results can inform knowledge on the impact of measurement error.

Measurement error is another significant factor that can lead to erroneous road safety decision-making. Measurement error denotes errors introduced into a dataset during measurement or data collection. It is collecting or measuring data such that what is assumed to be collected is not exactly what has been collected. These errors can be introduced through mistakes made in the data collection process as a result of human error or faulty data collection equipment or may be unavoidable due to uncertainty. For example, samples always contain some measurement error since they cannot be an exact representation of the population being studied. Measurement error is a problem of internal validity (Washington et al. 2011). A known source of measurement error in crash frequency data arises from the mis-identification of persons involved in crashes as pedestrians, for example, the characterization of a driver who has exited the vehicle temporarily as a pedestrian. This and other similar errors that diminish data quality are common in crash data collection.

In addition to data quality, data availability is another significant problem. In crash datasets, certain variables are more prone to data availability issues than others. Annual Average Daily Traffic (AADT) data is one such variable. A true AADT calculation is an average of all hourly traffic counts for the duration of a year (FHWA 2014). This is a very costly and time intensive undertaking, and as such, hourly data is commonly not collected for all road segments within a road network, and what is collected is often affected by inaccuracy (FHWA 2014), thereby increasing unavailability since erroneous data is often disposed of. This means that AADT can only be calculated for a limited number of roads and must be used with caution. It also means that AADT must be extrapolated or estimated for the segments for which it has not already been calculated from hourly traffic volume data.

Data quality and availability are research method issues. This is because the analyst must find ways to render the data useful for analysis, despite any quality or availability issues. To illustrate the data quality and availability problems in crash frequency research, I will briefly discuss some of the methods used in practice to calculate AADT. There are several methods commonly used in AADT calculation from hourly traffic data. There is the simple averaging method which is simply an average of hourly traffic volume data for 365 days. This method is not often used because of the quantity of data required. The most commonly used method is the AASHTO method which uses 84 averages (12 months times 7 days of week) to account for month and daily variability. All that is required is average daily traffic data for each day of any week (they do not necessarily have to be the same week) for the given month. This means that only 1 weekday average for each day (as opposed to four averaged for each day, since each day occurs approximately four times in a month), is needed to come up with the 84 averages required for the AADT calculation (7 weekdays times 12 months e.g. Monday average times twelve months). This obviously does not take all hourly data into account and results are therefore not precise. Exacerbating this issue is the fact that this method underestimates and overestimates the average, because it assumes 4-week months, even though some months consist of 28, 29, 30 or 31 days. In other words, there are still quality and availability issues with the AADT calculation.

In addition to the problem of biased AADT data, there is the problem of missing AADT data. AADT data is usually not available for all segments or all roads in a road network. Since traffic volume data is often not collected for all roads in a dataset due to cost and time limitations as well as data collection errors, many crash datasets have a large amount of observations with missing AADT data. This means that AADT data for road segments that have not been collected must be extrapolated or estimated from segments for which the calculation has been done from hourly data. The process of estimating or extrapolating AADT data then introduces measurement error because the AADT values attributed to segments for which no observations were made are only estimations and will deviate from what the actual observed value would be, to various degrees.

An FHWA study assessed the impact of missing data in the calculation of AADT from hourly traffic data. The bias found as a result was very small for the commonly used AASHTO method. The study found a bias of 0.30% when 60 days' (the largest number of days' worth of missing data tested in the study) worth of data is missing (FHWA 2014). This 16% missing data seems substantial enough to yield more than a .30% bias. While this bias is small, the fact that AADT might still need to be estimated from calculated AADT (which necessarily introduces error) means that the error is compounded.

Data and Methods

In carrying out these analyses for North Carolina highways, I use Highway Safety Information System (HSIS) data. The HSIS is a database developed by the University of North Carolina Highway Safety Research Center (HRSC) in partnership with the Federal Highway Administration (FHWA), using data already collected by participating states, but processed centrally for the purpose of enabling research towards the goal of improving highway safety (FHWA 2016). They collect and clean road safety related data for the states of California, North Carolina, Illinois, Ohio, Maine, Utah, Michigan, Washington and Minnesota. HSIS data is collected in two main groups- crash data (which contains data for the vehicles involved as well as the occupants), and roadway data, which contains data for the road segments on which the crash occurred. Each data group has several variables further identifying the crash frequency such as number of vehicles, type of vehicle, type of crash, etc. I use HSIS data for my North Carolina dataset because it is assumed to be of better quality than most other crash datasets. Below is a description of HSIS by the FHWA.

> "FHWA conducts extensive quality control checks on the data it receives from States. Each year, HSIS analysts examine new data files and compare them to the previous year's data. Then they develop metrics to measure differences in the data. If more than a very small difference is found between the values of a given year and the previous year, the analysts check the variable to understand the difference. The goal is uniformity and consistency year after year." (Fitzgerald 2014)

The reason why HSIS data is likely to be of better quality than DOT data is that it is re-processed centrally at HSIS labs as described above (FHWA 2016), after collection by police departments and subsequent processing by state Departments of Transportation (at which point it has already undergone some level of correction or cleaning). It is important for my starting point in this analysis to be with data that is as clean as possible to isolate the association of measurement error arising from the use of estimated AADT. If there are other sources of measurement error, it would be difficult to know what the specific association of erroneous AADT data is. It is reasonable to assume that two levels of processing will result in better data.

In the same way and for the same reasons as in my Pennsylvania study, I analyze the data in three steps. The first was to model crash frequency on only contextual variables, with block groups as the unit of observation. In the next step, I modeled crashes on only geometric variables in link-based models, and in the final step, I modeled crashes on both geometric and contextual variables in combined models. I carry out these three steps for both my dataset with observed AADT and my dataset with estimated AADT for three dependent variables: crashes, fatal and incapacitating injury crashes, and fatal and injury crashes. Note that AADT is not actually a variable in my models but is used to calculate the vehicle miles traveled variable.

In the previous chapter I made the case that using negative binomial conditional autoregressive models, is the correct way to model this kind of data, rather than simply using negative binomial models. This is in order to account for spatial correlation which can be a significant issue in crash datasets, especially when contextual variables are involved. The results of the analyses carried out using my Pennsylvania dataset show that the MLE model results were closely replicated by the MCMC models. I therefore present and analyze only the results of the negative binomial conditional autoregressive models (MCMC) in this section, while including the results of the negative binomial models in the appendix.

My North Carolina analysis also uses the same data types as my Pennsylvania study. Again, I examined crash frequency data, road geometry data and contextual data. I used a road network shapefile which I obtained from the North Carolina Department of Transportation (NCDOT), containing 340,181 road segments, including interstates, principal arterials, minor arterials, collectors, local roads and ramps. The following is a summary of the road network dataset.

Table 25: Summary of North Carolina Road Network

	Number of segments	% of Total	Total length (miles)
Interstates	8,178	2.40%	1208.40
Principal Arterials	39,615	11.65%	4020.09
Minor Arterials	41,017	12.06%	5290.97
Collectors	68,653	20.18%	16805.32
Local Access Roads	182,718	53.71%	51512.79
Total	340,181	100.00%	78837.57

(NCDOT 2016)

I chose to perform my analyses on interstate highways. The main reason was that this was the functional classification of North Carolina roads that had the least skewed distribution of crashes. Table 26 below shows the distribution of crashes by road functional classification. Most functional classes had around 50% or a higher proportion of their road segments without any crashes at all. This was an important factor to consider because it affected the ability to estimate MCMC models, since the software I used (*CrimeStat)* is severely limited when the dependent variable is very skewed. This problem is noted in the *CrimeStat* software documentation. MCMC models are not appropriate for data with highly skewed dependent variables because this condition may cause the models to fail to converge (Levine 2013b, Levine 2013a). I decided to focus on interstate highways so as to use MCMC models for my estimations, since they account for spatial correlation.

Table 26: Distribution of Crash Occurrences by Road Functional Classification

	Segments with 0 Crashes	Total Segments	% with 0 Crashes
Interstates	2,375	8,178	29.04%
Principal Arterials	15,416	39,615	38.91%
Minor Arterials	16,462	41,017	40.13%
Collectors	28,629	68,653	41.70%
Local Roads	128,796	182,718	70.49%

There were 1208 miles and 8178 segments of interstate highway in the dataset, of which 8071 had observed AADT data. The geometric variables used in my models were from the HSIS North Carolina database. They were pavement width, lane count, median width, shoulder width, and AADT. I again used sinuosity as my measure of curvature and obtained it using the same ArcGIS function used in my Pennsylvania dataset (ESRI 2011). I also calculated vehicle miles traveled by taking the product of AADT and segment length.

I obtained crash frequency data from the Highway Safety Information System (HSIS) for the years 2009 to 2013. As in the Pennsylvania study, 5-year data blocks are used so that crash data is large enough to analyze. Table 27 below summarizes the kinds of crashes that occurred on North Carolina interstates in this period.

Table 27: Summary of Crash Types, 2009- 2013

Crash Type	Frequency
Crashes with property damage only	69,022
Crashes with injuries	21,909
Crashes with fatalities	893
Total crashes	91,824

(FHWA 2016)

Again, just like in the Pennsylvania dataset, the vast majority of crashes were property damage only crashes.

Table 28 below shows geometric variable distributions for interstate highways. Contrary to what might be expected since interstates tend to be designed according to uniform standards, there is substantial variance for total width and median width.

Table 28: Distribution of Geometric Variables for Interstates

Variable	Mean	Standard Deviation
Total Width	80.16	31.83
Lane Count	5.07	1.53
Median Width	50.19	61.70
Shoulder Width	10.57	2.89
Vehicle Miles Traveled	7282.24	14014.66
Sinuosity	0.999	0.0083

(FHWA 2016)

Unlike the data for the other variables which I obtained from the HSIS database, contextual data for the North Carolina study was obtained from the US Census Bureau, since the HSIS database is not a source of contextual data. I obtained data for median income, population, and employment at the block group level and for the years between 2009 and 2013. The employment data consists of employment totals by block group in terms of work location. At the time of analysis, North Carolina had a total of 6155 block groups. Table 29 below shows a summary of the variable values. Both population density and employment density were overdispersed- the standard deviation of their distributions are higher than the mean.

Table 29: Distribution of Contextual Variables

Variable	**Mean**	**Standard Deviation**
Median Income	47,717.58	25,668.94
Population Density	94.08	137.86
Employment Density	716.77	3158.43

Sources: United States Census Bureau

Results and Discussion

In this section, I discuss results from models estimated using observed AADT and models run using estimated AADT for some observations.

Models with Observed AADT data

All Crashes

Table 30 and Table 31 contain models for which the observations had only AADT that was observed and therefore supplied from the data source. This observed AADT was used to calculate the vehicle miles traveled variable. Table 30 shows the negative binomial autoregressive model of crashes on contextual variables only, Table 31 shows the same model type for crashes on geometric variables (link-based models) and for crashes on all variables (combined models). For labeling the models, I continue the sequence I used in the previous chapter. When discussed in the text, the models appear with the same labels used when displayed as tables, as well as another label in parenthesis that shows whether they are spatial (S), link-based (L) or combined models (C), MCMC or MLE models, and the dependent variable of the model represented by a code of 1, 2 or 3. The dependent variables are represented by 1 for total crashes, 2 for fatal and incapacitating injury crashes, and 3 for fatal and injury crashes.

Table 30: Negative Binomial Autoregressive Model with only Contextual Variables (Observed AADT):

	Model **22** (S-MCMC-1)			
CRASHES	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Population Density (ln)	-0.343	-27.950	-0.367	-0.319
Employment Density (ln)	0.136	16.253	0.120	0.152
Median Income (ln)	0.071	7.228	0.053	0.091
Interstate Density (ln)	1.408	18.902	1.263	1.554
Principal Arterials Density (ln)	0.812	20.051	0.733	0.892
Minor Arterials Density (ln)	0.562	14.957	0.489	0.636
Collectors Density (ln)	0.263	6.198	0.179	0.346
Local Roads Density (ln)	0.224	8.740	0.173	0.275
Constant	4.415	39.592	4.184	4.623
Spatial Correlation (phi)	-0.001	-1.414	-0.003	0.000
Observations	6155			
Df	6144			
Log likelihood	-37543.520			

Model 22 (S-MCMC-1) models the effects of the contextual variables only, while accounting for the densities of principal arterial highways, minor arterials, collectors and local roads for all North Carolina block groups. Population density has a negative association with crashes, while employment density and median income have positive associations. All the contextual variable coefficients were found to be significant. This was also true for their associations in my Pennsylvania dataset, for the spatial model with the entire road network (Model 1 (S-MCMC-1)). The population density association with crashes is in line with Ewing, Schieber et al. (2003), Ewing, Dumbaugh (2009) and

Zhou, Sisiopiku (1997), which found that population density is associated with a lower risk of crash occurrence. Since this is the result of just one model, as opposed to a commonality across several models, the positive association between median income and crashes could be a result of model-specific unseen errors and illustrates the problem with indeterminacy in the estimates of crash frequency models.

Table 31: MCMC Models (link-based and combined) (Observed AADT)

	Model 23 (L-MCMC-1)				**Model 24** (C-MCMC-1)			
CRASHES	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Total width (ln)	-0.139	-2.396	-0.648	0.377	-0.161	-2.754	-0.681	0.356
Lane count (ln)	1.293	13.705	0.457	2.128	0.601	6.202	-0.255	1.453
Median width (ln)	-0.100	-6.651	-0.233	0.032	-0.096	-6.422	-0.228	0.035
Shoulder width (ln)	-0.058	-1.428	-0.429	0.288	0.039	1.019	-0.306	0.362
VMT (ln)	0.683	64.950	0.592	0.778	0.677	64.594	0.587	0.771
Sinuosity (ln)	-7.803	-2.619	-31.542	15.973	-7.846	-2.703	-31.093	16.163
Median income (ln)					-0.009	-0.398	-0.222	0.182
Population density (sq mi, ln)					0.059	2.518	-0.146	0.260
Employment density (sq mi, ln)					0.149	10.903	0.029	0.270
Constant	0.694	0.335	-15.819	17.178	0.872	0.433	-15.889	16.966
Spatial Correlation (phi)	-0.038	-4.849	-0.108	0.026	-0.006	-0.983	-0.065	0.038
Observations	8071				8071			
Df	8062				8059			
Log likelihood	-23311.8				-23059.4			

Comparing Model 23 (L-MCMC-1) and Model 24 (C-MCMC-1) shows the changes that occur from the link-based model to the combined model due to the addition of the contextual variables of population density, employment density, and median income. All 8071 interstate segments with observed AADT data were used in these models.

Table 31 show a good deal of stability between the link-based and the combined models for the variables of total width, median width, vehicle miles traveled, and sinuosity. For these four variables, the coefficient magnitudes are approximately the same between Model 23 (L-MCMC-1) and Model 24 (C-MCMC-1) for example, total width goes from -0.14 to -0.16, median width goes from -0.1 to -0.1, vehicle miles traveled from 0.68 to 0.68, and sinuosity from -7.8 to -7.9. The negative association of total width on crash frequency is consistent with the other studies I examined, including Abdel-Aty and Radwan (2000), Labi (2011), Council and Stewart (1999), and Garnowski and Manner (2011). The positive association of vehicle miles traveled (VMT) shows consistency with the traffic volume associations seen in other studies I examined with the lowest in the range being 0.24 (Zeng & Huang, 2014) and the highest being 1.18 (Council & Stewart, 1999). The positive associations of lane count on crash frequency seen in both Model 23 (L-MCMC-1) and Model 24 (C-MCMC-1) are consistent with the results of other studies I examined including Sawalha and Sayed (2001), and Zeng and Huang (2014), although both studies showed much smaller coefficients for lane count with 0.085 and 0.17 respectively.

Shoulder width changes direction of association from negative in Model 23 (L-MCMC-1) to positive in Model 24 (C-MCMC-1). I examined two studies for shoulder width, with both showing negative associations of -0.15 and -0.30 (Sawalha & Sayed, 2001) and (Milton, Mannering 1998)), consistent only with the shoulder width coefficient in Model 24 (C-MCMC-1). It is interesting that Model 24 (C-MCMC-1) and not

Model 23 (L-MCMC-1), should be more consistent with the outside studies, since they did not include contextual variables in their models. Median income and population density take on the opposite direction of association in Model 24 (C-MCMC-1) from their previous direction in Model 22 (S-MCMC-1), while employment density does not change from its positive association.

Fatal & Incapacitating Injury Crashes

Table 32 below shows the crash frequency models for fatal and incapacitating injury crashes, using only observed AADT data for all observations. Model 25 (L-MCMC-2) is the link-based negative binomial autoregressive model for fatal and incapacitating injury crashes, Model 26 (C-MCMC-2) is the combined negative binomial autoregressive model for fatal and incapacitating injury crashes.

Table 32: MCMC Models (link-based and combined) (Observed AADT)

	Model 25 (L-MCMC-2)				Model 26 (C-MCMC-2)			
FATAL & INCAPACITATING INJURY CRASHES	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Total width (ln)	-0.504	-6.545	-1.128	0.100	0.050	0.692	-0.574	0.615
Lane count (ln)	0.243	1.503	-1.019	1.511	-0.275	-1.952	-1.390	0.913
Median width (ln)	0.023	0.615	-0.298	0.345	0.129	3.246	-0.218	0.449
Shoulder width (ln)	0.160	1.757	-0.528	0.911	0.203	2.162	-0.559	0.991
VMT (ln)	0.747	33.137	0.565	0.955	0.749	32.290	0.566	0.958
Sinuosity (ln)	5.885	10.405	1.417	10.281	-7.377	-16.712	-10.897	-3.542
Median income (ln)					-0.016	-0.645	-0.212	0.210
Population density (sq mi, ln)					0.059	0.189	4.084	-0.226
Employment density (sq mi, ln)					0.149	-0.095	-3.373	-0.333
Constant	-11.618	-30.037	-14.453	-8.630	-4.482	-15.393	-6.842	-2.100
Spatial Correlation (phi)	-0.031	-4.766	-0.099	0.013	-0.023	-3.332	-0.090	0.021
Observations	8071				8071			
Df	8062				8059			
Log likelihood	-2448.51				-2474.35			

All the geometric variable coefficients increase in magnitude from Model 25 (L-MCMC-2) to Model 26 (C-MCMC-2), except total width. Lane count and vehicle miles traveled did not increase substantially.

Three variables changed direction of association in Model 26 (C-MCMC-2). Total width became positive, while lane count and sinuosity became negative. Median width and shoulder width became significant in Model 26 (C-MCMC-2). Pavement width, lane count, shoulder width, and median width are not very much in line with the directions of association found in the prior studies I examined. Both Wu et al. (2015) and Kononov et al. (2008) found a negative association for pavement width, while Kononov et. Al. found a positive association for lane count for crashes with fatalities and any level of injury (Wu et al. 2015, Kononov et al. 2008). Both Harwood (2000) et. al. and Haleem et. al. (2012) found a negative association for shoulder width. Alluri et. al. (2012) found a negative association for median width.

Fatal and Injury Crashes

Table 33 below shows the crash frequency models for fatal and all injury crashes, using only observed AADT for all observations. Model 27 (L-MCMC-3) is the link-based negative binomial autoregressive model for fatal and injury crashes, while Model 28 (C-MCMC-3) is the combined negative binomial autoregressive model for fatal and injury crashes.

Table 33: MCMC Models (link-based and combined) (Observed AADT)

	Model 27 (L-MCMC-3)				Model 28 (C-MCMC-3)			
FATAL & INJURY CRASHES	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Total width (ln)	-0.120	-1.994	-0.632	0.424	-0.172	-3.017	-0.658	0.340
Lane count (ln)	1.081	11.356	0.240	1.900	0.348	3.640	-0.484	1.191
Median width (ln)	-0.116	-6.917	-0.263	0.029	-0.071	-4.447	-0.213	0.070
Shoulder width (ln)	-0.097	-2.290	-0.473	0.267	0.001	0.020	-0.336	0.327
VMT (ln)	0.709	63.176	0.612	0.807	0.696	63.671	0.601	0.794
Sinuosity (ln)	-1.111	-0.970	-10.225	7.746	-2.800	-2.627	-11.597	5.490
Median income (ln)					-0.054	-2.203	-0.277	0.144
Population density (sq mi, ln)					0.059	0.139	6.060	-0.062
Employment density (sq mi, ln)					0.149	0.132	9.557	0.010
Constant	-5.112	-6.383	-11.459	1.232	-3.403	-4.655	-9.209	2.610
Spatial Correlation (phi)	-0.036	-6.282	-0.091	0.007	-0.024	-3.445	-0.099	0.023
Observations	8071				8071			
Df	8062				8059			
Log likelihood	-14313.87				-14096.28			

All the variable coefficients decrease in magnitude from Model 27 (L-MCMC-3) to Model 28 (C-MCMC-3) except for sinuosity and total width. The increase in the total width coefficient is not substantial. All the variable coefficients remain the same in direction of association in Model 28 (C-MCMC-3), except for shoulder width which becomes positive. Shoulder width also becomes insignificant in Model 28 (C-MCMC-3). Wu et al. found a negative association for pavement width, while Kononov et. al. found a positive association for lane count for crashes with fatalities and any level of injury (Wu et al. 2015, Kononov et al. 2008), both in line with pavement width and lane count in Model 28 (C-MCMC-3). Both Harwood et. al. and Haleem et. al. found a negative association for shoulder width (Haleem, Gan et al. 2012, Harwood, Council et al. 2000) contrary to the direction of the shoulder width coefficient in Model 28, which was not statistically significant. Alluri et. al. found negative association for median width (Alluri, Ogle 2012), in line with Model 28 (C-MCMC-3).

Models with Estimated AADT data

All Crashes

In the previous sections, I discussed the results of models run using observed AADT. In the next few sections, I estimate AADT for 30% of the observations, chosen at random, using the remaining 70% or 5650 observations from the total 8071. The choice to explore the effect of using estimated AADT data for 30% of the dataset was not completely arbitrary. This 30% represents a third of the dataset, and the extent to which this relatively small proportion of the dataset has a substantial impact on inferences will be an indication of the importance of the problem of measurement error introduced by analyst methods of dealing with data problems. Table 34 below shows the results of this estimation.

Table 34: OLS estimation of AADT using 70% of AADT observations

VARIABLES	AADT
Population	32.36***
	(1.934)
Employment	1.778***
	(0.108)
Pavement Width	144.8***
	(14.06)
Lane Count	10,259***
	(313.8)
Constant	-24,095***
	(1,289)
Observations	5,650
R-squared	0.537
Robust standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

The AADT estimation using population, employment, pavement width and lane count as the independent variables resulted in a 53.7% R^2, meaning that the model explains 53.7% of the variation in the data. This is a moderately good fit.

The parameters from this model were then used to predict AADT for the remaining 30% for which AADT was assumed to be missing. As some of the variables used in the AADT estimation were also used in the crash models, I added the error terms in the crash models as well, as seen in the Table 35 models below. Table 35 shows the negative binomial autoregressive model of crashes for both link-based and combined models, or Model **29** (L-MCMC-1) and Model **30** (C-MCMC-1) respectively.

Table 35: Negative Binomial Autoregressive Models with Estimated AADT (link-based and combined)

	Model **29** (L-MCMC-1)				Model **30** (C-MCMC-1)			
CRASHES	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Total width (ln)	-0.095	-1.631	-0.601	0.414	-0.256	-4.587	-0.736	0.234
Lane count (ln)	1.243	13.983	0.458	2.015	0.634	6.704	-0.203	1.454
Median width (ln)	-0.104	-7.198	-0.232	0.022	-0.078	-5.151	-0.211	0.052
Shoulder width (ln)	0.018	0.468	-0.331	0.349	-0.083	-2.186	-0.427	0.243
VMT (ln)	0.654	63.844	0.564	0.746	0.641	62.027	0.553	0.735
Sinuosity (ln)	-18.221	-13.173	-28.964	-6.909	7.226	5.095	-3.588	18.336
Residuals (AADT)	1.20E-05	14.325	5.00E-06	1.90E-05	8.00E-06	9.439	1.00E-06	1.50E-05
Median income (ln)					0.057	2.572	-0.137	0.247
Population density (sq mi, ln)					0.019	0.831	-0.179	0.215
Employment density (sq mi, ln)					0.163	12.214	0.048	0.281
Constant	7.814	8.100	-0.160	15.225	-9.344	-9.561	-17.054	-1.691
Spatial Correlation (phi)	-0.006	-1.412	-0.050	0.028	0.000	-0.057	-0.049	0.041
Observations	8071				8071			
Df	8061				8058			
Log likelihood	-23438.37				-23192.53			

Lane count, median width, VMT, and sinuosity decreased in magnitude in Model **30** (C-MCMC-1), when compared with their values in Model **29** (L-MCMC-1), while total width and shoulder width increased. While most variables did not change their direction of association, shoulder width became negative, and sinuosity became positive in Model **30** (C-MCMC-1). Total width and median width were negative in both models, while lane count, and VMT were positive in both models. Two variables, total width and shoulder width were not significant in Model **29** (L-MCMC-1), but became significant in Model **30** (C-MCMC-1). The negative association of total width was also seen in Abdel-Aty and Radwan (2000), Labi (2011), Council and Stewart (1999), and Garnowski and Manner (2011). The positive association of VMT was consistent with Zeng & Huang (2014) and Council & Stewart (1999) and that of lane count was consistent with Sawalha and Sayed (2001), and Zeng and Huang (2014). The negative association of shoulder width in Model **30** (C-MCMC-1) was consistent with the Sawalha & Sayed (2001) and (Milton, Mannering 1998) studies. The studies I examined for median width found opposite associations. Abdel-Aty (2000) found a negative association, agreeing with my results, while Malyshkina and Mannering (2010) found a positive association. In general, my results agree with the other crash frequency studies I examined, even for the combined model (Model **30** (C-MCMC-1)). This implies that measurement error from the omission of contextual variables is not an issue.

Again, the goal of this re-estimation of crash frequency, using estimated AADT is to compare the results with the models where only observed AADT was used, to assess how measurement error can affect the inferences made. The key thing is to note how substantial the differences are. Seeing a substantial difference in the associations will show that data availability and quality issues can have a far-reaching effect on inferences made. In comparing Model **30** (C-MCMC-1), which uses estimated AADT for 30% of the dataset, with Model 24 (C-MCMC-1), which uses observed AADT only, there are some substantial differences. Shoulder width and sinuosity, which were positive and negative respectively in Model 24 (C-MCMC-1), changes direction of association to negative for shoulder width, and positive for sinuosity in Model **30** (C-MCMC-1). While only shoulder width is insignificant in Model 24 (C-MCMC-1), both shoulder width and sinuosity are significant in Model **30** (C-MCMC-1). These results show that using estimated data, for even just a small proportion of the dataset (a third) can result in very different results and highlights the impact of data quality and availability. Using estimated AADT as in Model **30** (C-MCMC-1), the inference is that shoulder width has a negative association with crashes, while sinuosity has a positive association. The inference is the exact opposite when using observed AADT as in Model 24 (C-MCMC-1).

These results raise the important question of how data quality may affect the reliability of the inferences that can be drawn about the associations of geometric variables on crash frequency. This is an important question because these inferences form the basis of plans to implement road safety design changes. Safety treatments, as much as possible, should not be applied based on inferred associations that might in reality have no significance, or show the wrong direction due to bias.

Fatal & Incapacitating Injury Crashes

Table 36 below shows the crash frequency models for fatal and incapacitating injury crashes, using estimated AADT for 30% of the observations. Model 31 (L-MCMC-2) is the link-based negative binomial autoregressive model for fatal and incapacitating injury crashes, while Model 32 (C-MCMC-2) is the combined negative binomial autoregressive model for fatal and incapacitating injury crashes.

Table 36: Negative Binomial Autoregressive Models (link-based and combined) (Estimated AADT)

	Model 31 (L-MCMC-2)				Model 32 (C-MCMC-2)			
FATAL & INCAPACITATING INJURY CRASHES	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Total width (ln)	-0.204	-1.973	-1.067	0.628	-0.229	-2.552	-0.965	0.477
Lane count (ln)	0.394	2.274	-1.108	1.882	0.034	0.185	-1.482	1.516
Median width (ln)	0.115	2.525	-0.257	0.520	0.149	3.457	-0.208	0.534
Shoulder width (ln)	0.075	0.800	-0.664	0.899	0.122	1.210	-0.712	1.036
VMT (ln)	0.751	29.154	0.540	0.976	0.712	27.035	0.499	0.939
Sinuosity (ln)	-6.504	-9.574	-11.417	-0.626	-21.417	-30.099	-27.049	-15.788
Residuals (AADT)	4.30E-06	2.316	-1.20E-05	2.10E-05	6.20E-06	3.13	-1.10E-05	2.40E-05
Median income (ln)					-0.152	-3.747	-0.455	0.190
Population density (sq mi, ln)					0.113	2.171	-0.360	0.561
Employment density (sq mi, ln)					-0.034	-1.088	-0.303	0.233
Constant	-4.888	-9.824	-8.860	-1.205	7.687	14.780	3.749	11.461
Spatial Correlation (phi)	-0.041	-5.536	-0.114	0.009	-0.010	-1.859	-0.066	0.027
Observations	8071				8071			
Df	8061				8058			
Log likelihood	-2481.25				-2493.45			

There is a lot of stability between Model 31 (L-MCMC-2) and Model 32 (C-MCMC-2). All the geometric variables remain unaffected in direction of association when they are specified with contextual variables in Model 32 (C-MCMC-2). Most of the geometric variables, except for lane count and sinuosity did not change very much in magnitude of association in Model 32 (C-MCMC-2). Again, this does not support my hypothesis, which is that geometric variable coefficients are affected by bias when they are specified without contextual variables. This was not the case when the fatal & incapacitating injury models using only observed AADT were compared (Model 25 (L-MCMC-2) and Model 26 (C-MCMC-2).

Comparing the combined model for fatal & incapacitating injury crashes using observed AADT (Model 26 (C-MCMC-2)), with the corresponding model using 30% estimated AADT Model 32 (C-MCMC-2), the impact of using estimated data becomes more apparent. Two variables switch direction of association from (Model 26 (C-MCMC-2) which uses observed AADT to Model 32 (C-MCMC-2) which uses some estimated AADT. Total width becomes negative and gains significance, while lane count becomes positive and remains insignificant in Model 32 (C-MCMC-2). While shoulder width does not change direction between those two models, it becomes insignificant in Model 32 (C-MCMC-2).

Fatal & Injury Crashes

Table 37 below shows the crash frequency models for fatal and injury crashes, using estimated AADT for 30% of the observations. Model 33 (L-MCMC-3) is the link-based negative binomial autoregressive model for fatal and injury crashes, while Model 34 (C-MCMC-3) is the combined negative binomial autoregressive model for fatal and injury crashes.

Table 37: Negative Binomial Autoregressive Models (link-based and combined) (Estimated AADT)

	Model 33 (L-MCMC-3)				Model 34 (C-MCMC-3)			
FATAL & INJURY CRASHES	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Total width (ln)	-0.295	-4.739	-0.847	0.266	-0.307	-4.943	-0.848	0.261
Lane count (ln)	1.287	13.123	0.405	2.150	0.661	6.641	-0.207	1.553
Median width (ln)	-0.100	-5.721	-0.255	0.055	-0.069	-4.045	-0.219	0.082
Shoulder width (ln)	-0.062	-1.481	-0.439	0.300	-0.037	-0.893	-0.403	0.322
VMT (ln)	0.644	55.667	0.544	0.748	0.651	58.416	0.555	0.752
Sinuosity (ln)	6.384	4.488	-4.259	17.869	8.789	6.771	-1.098	19.523
Residuals (AADT)	1.10E-05	12.893	4.00E-06	1.90E-05	7.4E-6	8.497	0	1.50E-05
Median income (ln)					-0.027	-1.113	-0.243	0.182
Population density (sq mi, ln)					0.075	3.178	-0.134	0.282
Employment density (sq mi, ln)					0.149	10.394	0.024	0.276
Constant	-9.571	-9.612	-17.547	-2.049	-11.116	-12.634	-18.406	-4.464
Spatial Correlation (phi)	-0.012	-2.614	-0.062	0.022	-0.023	-3.713	-0.083	0.021
Observations	8071				8071			
Df	8061				8058			
Log likelihood	-14331.67				-14148.95			

None of the geometric variable coefficients in Model 33 (L-MCMC-3) change direction of association in Model 34 (C-MCMC-3). There is also not much change in magnitude of association between these two models. This does not support my hypothesis which is that the absence of contextual variables in crash frequency models introduces statistical errors that can be indicated by bias in the geometric variable coefficients. This was also the case for the fatal and major injuries models when the link-based model was compared with the combined model for estimated AADT. This was not found to be the case for the total crashes models.

The comparison between the combined fatal & injury model using observed AADT (Model 28 (C-MCMC-3)) with the combined fatal & injury model using estimated AADT Model 34 (C-MCMC-3) does indicate that measurement error can be impactful. While most geometric variables have the same direction of association, and show only small variations in magnitude of association, shoulder width becomes negative, and sinuosity becomes positive in Model 34 (C-MCMC-3).

Conclusions

The goal of this chapter was to explore the impact of data quality and availability on crash frequency models. To do this, I ran crash frequency models with geometric and contextual variables, using a dataset with AADT generated from observed hourly traffic volume data. I then ran these models again using a dataset for which AADT data for a randomly chosen 30% of the observations was assumed to be missing, and therefore estimated from observed AADT. The rationale here is that estimated AADT will have some error since it is a deviation from whatever the true value would be. I compared results from the models with only observed AADT to the results from the models with 30% estimated AADT. The differences which are due to the errors in the estimation, can be indicative of the consequence of measurement error. It shows how data quality and availability can impact crash frequency modeling, and in turn, impact the decisions that are made based on these models. The following are summary tables of these comparisons between the combined models with observed AADT and the combined models with estimated AADT for the dependent variables of total crashes, fatal and incapacitating injury crashes and fatal and injury crashes.

Table 38: Summary of total crashes models

	Model 24 (C-MCMC-1)			Model 30 (C-MCMC-1)		
	(Observed AADT)			**(Estimated AADT)**		
Total width (ln)	↑	-	sig	↑	-	sig
Lane count (ln)	↓	+	sig	↓	+	sig
Median width (ln)	↓	-	sig	↓	-	sig
Shoulder width (ln)	↓	+	ns	↑	-	sig
VMT (ln)	↓	+	sig	↓	+	sig
Sinuosity (ln)	↑	-	sig	↓	+	sig
Median income (ln)		-	ns		+	sig
Population density (sq mi, ln)		+	sig		+	ns
Employment density (sq mi, ln)		+	sig		+	sig

The summary tables show two columns for each model. The first column shows the change in magnitude of association from the corresponding link-based model. For example, Table 38 (total crashes) shows that total width and sinuosity increased in magnitude in Model 24 (C-MCMC-1) from Model 23 (L-MCMC-1), which is not shown here. The individual magnitude changes for each variable is not very important because for most of the variables the changes in magnitude are very small. The second column for Model 24 (C-MCMC-1) which was run using observed AADT, is more important because it shows the direction of association of the model, and will be compared to the second column of the corresponding estimated AADT model, Model 30 (C-MCMC-1). For the total crashes models (Model 24 and Model 30), three variables change direction of association. Shoulder width takes on a negative association, while sinuosity and median income both take on a positive association in Model 30 (C-MCMC-1) which is specified with estimated AADT.

Table 39: Summary of fatal and incapacitating injury crashes models

	Model 26 (C-MCMC-2)			Model 32 (C-MCMC-2)		
	(Observed AADT)			**(Estimated AADT)**		
Total width (ln)	↓	+	ns	↑	-	sig
Lane count (ln)	↑	-	ns	↓	+	ns
Median width (ln)	↑	+	sig	↑	+	sig
Shoulder width (ln)	↑	+	sig	↑	+	ns
VMT (ln)	↑	+	sig	↓	+	sig
Sinuosity (ln)	↑	-	sig	↑	-	sig
Median income (ln)		-	ns		-	sig
Population density (sq mi, ln)		+	ns		+	sig
Employment density (sq mi, ln)		+	ns		-	ns

For the fatal and incapacitating injury crashes models, Model 26 (C-MCMC-2) and Model 32 (C-MCMC-2), total width and lane count change direction of association. Total width takes on a negative association, and lane count takes on a positive association with crash frequency in Model 32 which is specified with estimated AADT.

Table 40: Summary of fatal and injury crashes models

	Model 28 (C-MCMC-3			Model 34 (C-MCMC-3)		
	(Observed AADT)			**(Estimated AADT)**		
Total width (ln)	↑	-	sig	↑	-	sig
Lane count (ln)	↓	+	sig	↓	+	sig
Median width (ln)	↓	-	sig	↓	-	sig
Shoulder width (ln)	↓	+	ns	↓	-	ns
VMT (ln)	↓	+	sig	↑	+	sig
Sinuosity (ln)	↑	-	sig	↑	+	sig
Median income (ln)		-	sig		-	ns
Population density (sq mi, ln)		+	ns		+	sig
Employment density (sq mi, ln)		+	ns		+	sig

For the fatal and injury models, Model 28 (C-MCMC-3) and Model 34 (C-MCMC-3), shoulder width and sinuosity change direction of association, with shoulder width taking on a negative association, and sinuosity taking on a positive association with crash frequency. For each of the three dependent variables, the effect of re-running the models using estimated AADT was to change the direction of association of the coefficients of roughly one-third of the geometric variables specified.

An FHWA study assessed the impact of missing hourly data on AADT estimation. The study estimated bias by comparing AADT estimates with no missing days, to estimates with 1 to 60 missing days. The maximum bias found was 0.30%, for their dataset with 60 days-worth of missing data (Krile, Robert 2014). This is obviously a very negligible amount of bias, showing that missing data (16% of the data is missing) may not have a large impact on AADT estimation. My results, which assume a larger proportion of the dataset is missing (30%), show the opposite- that missing data can be quite impactful. Figure 2 below shows the results from the FHWA study.

Figure 2: FHWA study findings on the impact of missing hour data

Days Excluded	**Method 1- Simple Averaging**		**Method 2- AASHTO**		**Method 3- AASHTO Adjusted**	
	Median Percent Bias	% Increase on Method 2 %Bias CI	Median Percent Bias	95% CI on % Bias	Median Percent Bias	% Increase on Method 2 %Bias CI
1	0.00	3.07	-0.05	(-0.42, 0.25)	0.00	-23.86
3	0.00	25.54	-0.05	(-0.57, 0.42)	0.00	-8.81
7	0.00	27.44	-0.04	(-0.86, 0.68)	0.00	-5.10
14	0.02	22.79	-0.04	(-1.38, 1.17)	0.00	-1.07
All But 7	-0.03	0.00	-0.03	(-2.24, 1.99)	0.00	0.74
30 day	-0.11	86.75	-0.08	(-1.30, 1.06)	-0.02	1.03
60 day	-0.39	43.79	-0.30	(-3.60, 2.53)	-0.26	-0.97
2 x 30 day	-0.17	82.41	-0.13	(-1.88, 1.57)	-0.07	0.97

The results of my study show the impact of a slightly different kind of data availability issue- missing AADT data (as opposed to missing hourly traffic volume data, which is used to calculate AADT). In both instances, AADT must be calculated, but in my study, the estimation of AADT from observed AADT which has been calculated from hourly data and potentially impacted by data availability issues may further bias results. In other words, being able to have AADT data for all road segments in a road network might involve using biased AADT data calculated from hourly observations to estimate missing AADT (which inadvertently can only be done with some error). Crash frequency analysis subsequently carried out using such data can result in significant bias with coefficients showing opposite directions of associations than they would show with observed AADT data, as seen in the model comparisons in Table 38, Table 39, and Table 40 . The only other choice is to not perform crash frequency analysis for those roads for which there is no observed AADT data.

Adding More Contextual Variables: Age, Precipitation and Elevation

All Crashes

In my literature review, I discussed some empirical research carried out to investigate the associations of other contextual variables such as age of the population, and weather on crash occurrences. A number of studies found that younger drivers are associated with higher crash risk because of their relative inexperience or high risk-taking tendency (Chen et al. 2006, Klauer et al. 2006, Deery 1999, Clarke et al. 2006). The association of precipitation to increase crash risk is also well documented (Eisenberg, Warner 2004, Strong, Ye et al. 2010). In this section, I show the results of some additional estimates using my North Carolina dataset with age, precipitation and elevation variables. I use my North Carolina dataset for this additional analysis because North Carolina has a somewhat varied landscape, being landlocked on its western end, with the ocean to its east.

I used elevation data and 30-year average precipitation data obtained from the Oregon State University PRISM Climate database (Oregon State University 2018). The elevation range for North Carolina is from 0 to 1575 feet above sea level, while the 30-year average annual precipitation range is from 928 to 2205 inches. I geocoded precipitation data by appending the precipitation attributes to the North Carolina block group that was under each precipitation zone. One limitation of using 30-year average precipitation data is that it does not account for variation over a 30-year period, although it will account for variation between wet and drier climates. It is more a measure of climate than it is of weather, since it doesn't take seasonal variation into account. I looked at two age variables- the proportion of the population age sixty-five and up (%65-up), and the proportion of those between ages 18-24 (%18-24). The range for those age 65 and up was from 0 to 14.5%, and that of those ages 18-24 was from 0 – 50.6%.

Table 41: Negative Binomial Autoregressive Models with Observed AADT (Contextual Variables)

Model 35 (S-MCMC-1)

CRASHES	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Population Density (ln)	-0.39	-31.89	-0.42	-0.37
Employment Density (ln)	0.15	18.54	0.13	0.16
Median Income (ln)	0.13	13.42	0.11	0.15
% Age 18-24 (ln)	0.32	11.32	0.26	0.37
% 65 and UP (ln)	-0.30	-9.47	-0.36	-0.24
Precipitation (ln)	0.17	0.97	-0.09	0.41
Elevation (ln)	-0.05	-4.48	-0.07	-0.03
Interstate Density (ln)	1.41	19.15	1.27	1.56
Principal Arterials Density (ln)	0.80	20.20	0.72	0.88
Minor Arterials Density (ln)	0.60	16.18	0.53	0.67
Collectors Density (ln)	0.29	6.97	0.21	0.38
Local Roads Density (ln)	10.25	9.68	0.20	0.30
Constant	2.96	2.29	1.18	4.92
Spatial Correlation (phi)	0.00	-1.53	0.00	0.00
Observations	6155			
Df	6140			
Log likelihood	-37393.9			

Model 35 (S-MCMC-1) shows the results of the spatial model, with the additional variables on total crashes for all North Carolina block-groups. All variables have significant associations except precipitation. Population density, employment density

and median income have the same direction of associations as in Model **22** (S-MCMC-1), where they are specified also with observed AADT, but without the additional variables. There is also not a great deal of variation in the magnitudes of association of these variables with the population density, employment density and median income coefficients at -0.39, 0.15 and 0.13 respectively in Model 35 (S-MCMC-1), compared with -0.34, 0.13, and 0.07 respectively in Model **22** (S-MCMC-1). The addition of the age, precipitation and elevation variables does not alter the model very much. The age variable %18-24 does have a positive association with total crashes, while %65-up has a negative association with total crashes. These results may be indicative of the 18-24 age group's relative driving inexperience (Chen et al. 2006, Klauer et al. 2006, Deery 1999, Clarke et al. 2006). The negative association of %65-up may be indicative of a low risk-taking tendency among seniors. Precipitation, while not significant has a positive association with total crashes, in line with Eisenberg, Warner (2004).

Table 42: Negative Binomial Autoregressive Models with Observed AADT (Combined Model)

Model 36 (C-MCMC-1)

CRASHES	**Coeff.**	**t-stat**	**2.5th Percentile**	**97.5th Percentile**
Total width (ln)	-0.115	-2.159	-0.586	0.362
Lane count (ln)	0.568	6.227	-0.219	1.374
Median width (ln)	-0.081	-5.559	-0.210	0.046
Shoulder width (ln)	-0.048	-1.270	-0.385	0.274
VMT (ln)	0.696	69.311	0.610	0.787
Sinuosity (ln)	-9.483	-8.045	-18.664	-0.212
Median income (ln)	0.069	3.061	-0.132	0.266
Population density (sq mi, ln)	0.156	11.894	0.042	0.272
Employment density (sq mi, ln)	-0.036	-1.554	-0.244	0.162
% Age 18-24 (ln)	0.257	5.182	-0.177	0.695
% Age 65 up (ln)	0.171	2.895	-0.346	0.692
Precipitation (ln)	-0.055	-0.412	-1.097	0.927
Elevation (ln)	-0.174	-9.151	-0.343	-0.008
Constant	2.934	3.042	-4.491	10.359
Spatial Correlation (Phi)	-0.003	-0.608	-0.048	0.037
Observations	8071			
Df	8055			
Log likelihood	-23112.83			

Model 36 (C-MCMC-1) shows the results of the combined model with the additional variables. The geometric variables have the same direction of association as those in Model 24 (C-MCMC-1), where they are specified also with observed AADT but without the additional variables. The exception is shoulder width which has a positive association with total crashes in Model 24 (C-MCMC-1), but a negative association with total crashes in Model 36 (C-MCMC-1). The negative association in the latter model is consistent with Sawalha & Sayed (2001) and (Milton, Mannering 1998). There is not

much variation between the coefficient magnitudes either, for the geometric variables when these two models are compared. This implies that the addition of the age, precipitation and elevation variables does not have a great impact on the previous models. The contextual variables of median income, population density and employment density do however vary somewhat between the two models. In Model 36 (C-MCMC-1), both median income and population density have positive associations while employment density has a negative association. In Model 24 (C-MCMC-1), median income has a negative association, while population and employment density have positive associations with total crashes. Therefore, of the three main contextual variables, only population density did not change direction of association upon the addition of the age, precipitation and elevation variables. Also unlike in the spatial model, %65-up has a positive association with crashes, while precipitation and elevation have negative associations.

Fatal & Incapacitating Injury Crashes

Model 37 (C-MCMC-2)

FATAL & INCAPACITATING INJURY CRASHES	Coeff.	t-stat	2.5th Percentile	97.5th Percentile
Total width (ln)	-0.072	-0.747	-0.873	0.693
Lane count (ln)	-0.152	-0.906	-1.573	1.120
Median width (ln)	0.096	2.152	-0.260	0.474
Shoulder width (ln)	0.053	0.513	-0.759	0.957
VMT (ln)	0.762	31.405	0.557	0.971
Sinuosity (ln)	6.990	9.793	1.862	12.868
Median income (ln)	0.001	0.030	-0.264	0.255
Population density (sq mi, ln)	0.096	1.849	-0.342	0.527
Employment density (sq mi, ln)	0.000	-0.009	-0.255	0.255
% Age 18-24 (ln)	-0.204	-1.577	-1.325	0.951
% Age 65 up (ln)	0.245	1.845	-0.926	1.415
Precipitation (ln)	-0.645	-10.229	-1.104	-0.151
Elevation (ln)	0.004	0.088	-0.344	0.342
Constant	-9.740	-20.011	-13.501	-6.221
Spatial Correlation (Phi)	-0.022	-3.054	-0.096	0.025
Observations	8071			
Df	8055			
Log likelihood	-2478.16			

Model 37 (C-MCMC-2) shows the results of the combined model for fatal and incapacitating injury crashes with the additional variables. Almost all the geometric variables have the same direction of association as those in Model 26 (C-MCMC-2), where they are specified also with observed AADT but without the additional variables. The exceptions are total width which takes on a negative association, and sinuosity which takes on a positive association in Model 37 (C-MCMC-2). The negative association of total width is consistent with the Wu, Han et al. (2015) study. This implies that the

addition of the age, precipitation and elevation variables does not have a great impact on the model. The contextual variables of median income, population density and employment density do however vary somewhat between the two models. In Model 37 (C-MCMC-2), median income takes on a positive association, while population and employment density do not change in direction of association. In this model, %18-24 has a negative association and %65-up has a positive association. Precipitation has a negative association, and elevation, a positive association.

Fatal & Injury Crashes

Model 38 (C-MCMC-3)

FATAL & INJURY CRASHES	Coeff.	t-stat	2.5th Percentile	97.5th Percentile
Total width (ln)	-0.353	-5.757	-0.888	0.198
Lane count (ln)	0.600	5.964	-0.287	1.488
Median width (ln)	-0.026	-1.536	-0.175	0.124
Shoulder width (ln)	-0.046	-1.117	-0.415	0.318
VMT (ln)	0.696	61.425	0.598	0.798
Sinuosity (ln)	-9.902	-7.832	-19.592	0.483
Median income (ln)	-0.024	-0.870	-0.258	0.223
Population density (sq mi, ln)	0.110	4.507	-0.108	0.326
Employment density (sq mi, ln)	0.144	9.925	0.016	0.271
% Age 18-24 (ln)	0.090	1.612	-0.404	0.587
% Age 65 up (ln)	0.042	0.650	-0.538	0.623
Precipitation (ln)	-0.135	-1.114	-1.142	0.742
Elevation (ln)	-0.052	-2.436	-0.241	0.137
Constant	2.618	2.728	-4.755	10.016
Spatial Correlation (Phi)	-0.015	-2.779	-0.075	0.023
Observations	8071			
Df	8055			
Log likelihood	-14069.67			

Model 38 (C-MCMC-3) shows the results of the combined model with the additional variables for fatal and injury crashes. Almost all the geometric variables have the same direction of association as those in Model 28 (C-MCMC-3), where they are specified also with observed AADT but without the additional variables. The exception is shoulder width which has a positive association with crashes in Model 28 (C-MCMC-3), but a negative association in Model 38 (C-MCMC-3). This negative association is consistent with both Haleem et al. (2012), and Harwood et al. (2000). There is not much variation

between the coefficient magnitudes for the geometric variables when these two models are compared. This implies that the addition of the age, precipitation and elevation variables does not have a great impact on the model. The contextual variables of median income, population density and employment density also do not vary between the two models in direction of association. The age variables have positive associations while precipitation and elevation have negative associations. The addition of elevation, precipitation and the age variables yielded meaningful change in the model coefficients. At least one variable in each model changed direction of association. These results underscore the problem of indeterminacy arising from different model specifications.

Possible Variable Interactions

In the previous chapter, I found that with certain models, for example, Model 26 (C-MCMC-2), variables like sinuosity, lane count median and shoulder width increased in magnitude in combined models, when compared to the corresponding link-based model, Model 25 (L-MCMC-2). This also happened to a much smaller degree and with far fewer variables in Model 24. Many crash frequency studies have found strong interactions between various road characteristics in their datasets (Wang, Simandl et al. 2016). I discussed that one possible reason for this could be interactions between these variables and some contextual variables. In this section I explore four possible interactions for interstates and principal arterials. I explored interactions between pavement width and population density, pavement width and elevation, lane count and %65-up, lane count and %18-24, and finally, sinuosity and %65-up. The following table shows the correlations between these variables.

Table 43: Correlations between potentially interacting variables

	Pavement Width	Population Density	Elevation	Lane Count	%65-up	%18-24	Sinuosity
Pavement Width	1.00						
Population Density	0.21	1.00					
Elevation	-0.18	0.04	1.00				
Lane Count	0.62	0.29	-0.09	1.00			
%65-up	-0.24	-0.30	0.29	-0.22	1.00		
%18-24	0.14	0.25	-0.08	0.23	-0.15	1.00	
Sinuosity	0.06	0.02	-0.14	0.03	-0.05	0.03	1.00

Pavement width and population density may have a relationship because roads tend to be narrower when going through dense city centers with high population density, although the correlation coefficient shown in Table 43 is not very high. Pavement width and elevation also do not correlate very highly, but may have a relationship as roads going through hills tend not to be wide. Lane count and %65-up may interact to increase the likelihood of crashes, and the same may be said for lane count and %18-24. Lane count does not correlate very highly with either age variable.

Since all these interactions involve contextual variables, I ran regressions for combined models as opposed to link-based models, with total crashes as the dependent variable. Model 39 shows the results for the MLE model with the pavement width-population density interaction for interstates.

Table 44: Interaction of Population Density with Pavement Width (Interstates)

	***Model** 39*
VARIABLES	**Crashes**
Total width (ln)	-.350***
	(0.085)
Lane count (ln)	.693***
	(0.137)
Median width (ln)	-.071***
	(0.024)
Shoulder width (ln)	-0.076
	(0.062)
Vehicle Miles Traveled (ln)	.632***
	(0.013)
Sinuosity (ln)	-5.589***
	(2.120)
Median income (ln)	.0118
	(0.020)
Employment density (sq mi, ln)	.154***
	(0.014)
Moderate Population Density	-.054
	(0.559)
High Population Density	-.693
	(0.646)
Moderate Population Density* Pavement Width	.014
	(0.131)
High Population Density * Pavement Width	0.215
	(0.146)
Constant	0.476
	(1.496)
Observations	8,178
Log likelihood	-23035.4
Ll Constant Only	-25648
LR Chi2	4627
Pseudo_R2	0.128
Robust standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

For Model 39, I categorized the population density variable into 3 levels, for low-density, moderate-density and high-density. The low-density category was between 0 and 30 people per square mile, the moderate-density category was between 30 and 80 people per square mile, and the high-density category was 80 people per square mile and higher. The low-density category was the reference category for this model. The results show that none of the population density categories had significant associations with total crashes. The same was true for the interaction of these two categories with pavement width.

None of the other interactions I explored for the age, elevation, lane and sinuosity variables had any significant associations with total crashes. Results are shown in the appendix.

Table 45: Interaction of Population Density with Pavement Width (Principal Arterials)

	***Model** 40*
VARIABLES	Crashes
Total Width (ln)	-0.330***
	(0.0525)
Lane count (ln)	0.726***
	(0.0724)
Median Width (ln)	-0.130***
	(0.00822)
Shoulder Width (ln)	-0.0780***
	(0.0132)
Vehicle Miles Traveled (ln)	0.552***
	(0.0142)
Sinuosity (ln)	-6.745***
	(1.269)
Median Income (ln)	0.0223
	(0.0144)
Employment Density (ln)	0.162***
	(0.00845)
Moderate Population Density	-0.297
	(0.247)
High Population Density	-1.005***
	(0.228)
Moderate Population Density * Total Width (ln)	0.106*
	(0.0636)
High Population Density * Total Width (ln)	0.339***
	(0.0574)
Constant	1.740*
	(0.934)
Observations	39,615
Log likelihood	-93400
Ll Constant Only	-102361
LR Chi2	6711
Pseudo_R2	0.0875
Robust standard errors in parentheses	

*** p<0.01, ** p<0.05, * p<0.1

Model **40** above shows the results for the principal arterial MLE model with the pavement width-population density interaction. I categorized population density for principal arterials in the same way I categorized population density in the dataset with just interstate highways. For Model **40**, the low-density category was between 0 and 30 people per square mile, the moderate-density category was between 30 and 80 people per square mile, and the high-density category was 80 people per square mile and higher. The low-density category was the reference category for this model. The first thing to note is that all the geometric variable coefficients have the expected direction of association. Total width, shoulder width, median width and sinuosity have negative associations with total crashes, while lane count and vehicle miles traveled have a positive association. The negative association of total width with crash frequency is consistent with Abdel-Aty and Radwan (2000), Labi (2011), Council and Stewart (1999), and Garnowski and Manner (2011). The positive association of vehicle miles traveled (VMT) is consistent with Zeng & Huang (2014) and Council & Stewart (1999). The positive association of lane count is consistent with Sawalha and Sayed (2001), and Zeng and Huang (2014). The positive association of shoulder width is consistent with Sawalha & Sayed (2001) and (Milton, Mannering 1998).

One reason for exploring interactions between certain variables was to see if the unexpected increase in magnitude of geometric variables in certain combined models, when compared with their corresponding link-based model was as a result of unexplored variable interactions. An example is seen in the comparison between the link-based and combined models for total crashes using observed AADT- Model 24 (C-MCMC-1) and Model 23 (L-MCMC-1). Another important finding from the Model **40** results above, is that compared with Model 23 (L-MCMC-1), is that this unexpected increase in magnitude is actually more pronounced, with the exception of vehicle miles traveled and sinuosity. This means that while the interactions in Model **40** are somewhat significant, they are not the reason for the unexpected increase in coefficient magnitude. Also, only one population density category is significant at the 99% confidence level. This raises the question of whether population density is actually interacting with total width at all.

Table 46: Interaction of %18-24 with Lane Count (Principal Arterials)

	***Model** 41*
VARIABLES	**Crashes**
Total Width (ln)	-0.176***
	(0.0410)
Lane Count (ln)	0.470***
	(0.0920)
Median Width (ln)	-0.132***
	(0.00800)
Shoulder Width (ln)	-0.0827***
	(0.0132)
Vehicle Miles Traveled (ln)	0.552***
	(0.0141)
Sinuosity (ln)	-7.312***
	(1.307)
%18to24 (ln)_moderate	-0.344**
	(0.147)
%18to24(ln)_high	-0.576***
	(0.151)
%18to24 (ln)_moderate * Lane count (ln)	0.311***
	(0.0956)
%18to24(ln)_high * Lane count (ln)	0.520***
	(0.0974)
Median Income (ln)	0.00863
	(0.0159)
Population Density (ln)	0.0989***
	(0.0189)
Employment Density (ln)	0.146***
	(0.0114)
Constant	1.827*
	(0.967)
Observations	39,615
Log likelihood	-93393
Ll Constant Only	-102361
LR Chi2	6626
Pseudo_R2	0.0876
Robust standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

In Model 41 above, %18-24 has been interacted with lane count. The reference category is %18-24 of under 1%, and consists of 15,033 observations. The next category (%18-24_moderate) denotes a composition of between 1% and 1.5% of those age 18 to 24 and consists of 11,762 observations. The last category (%18-24_high) denotes a composition of 1.5% or higher and consists of 12,820 observations. Both Model 41 and Model 40 show increases in the magnitude of the geometric variables, as was seen in Model 24 (C-MCMC-1) which was specified without any variable interactions when compared with the link-based counterpart- Model 23 (L-MCMC-1). This also means that the interactions explored in Model 40 and Model 41 are not the reasons for the increases seen in the geometric variable coefficients. The geometric variables in Model 41 have the expected directions of association and are consistent with outside studies. The two categories of %18-24 show a negative association with total crashes which is unexpected. Their interactions with lane count on the other hand, show positive associations, with %18-24_high having a higher magnitude of association than %18-24_high. This means that having a higher percentage of people age 18-24 increases total crashes multiplicatively.

Conclusions

In this chapter, I examined measurement error in my North Carolina dataset, by comparing the results of models run with observed AADT data with results for models run with estimated AADT data. The data subset with the estimated AADT can be compared with datasets that are affected by measurement error, because as an estimate or "best guess" of AADT for each road segment, it deviates from what the actual AADT would be if observed, to an unknown degree. Inferences made about the differences in the results can be useful in gaining an understanding of the potential of analyst data processing decisions to introduce measurement error and in turn, influence inferences from crash frequency models, and the road safety decisions that they inform. The results showed that data quality and availability issues can impact variable coefficients to the extent of changing the direction of association from what they would be should data quality and availability issues be absent. In practice, variables that go into crash frequency models for the determination of safety treatments can be produced or updated frequently or occasionally. Each update is an opportunity for data processing to be evaluated to control the potential to introduce measurement error and address the problem of indeterminacy of crash modification factors.

In this chapter, I also examine the impact of adding more contextual variables that are not commonly looked at in crash frequency research, and finally, I examine the possibility of variable interactions. The impact of adding the new contextual variables of precipitation, elevation, %18-24 and %65-up on the geometric variables was found to be minimal, although these variables had significant associations with total crashes. The implication is that there is some benefit to specifying these additional variables. In exploring interactions, I found that there were no interactions for interstates, but that there were interactions between lane count and %18-24 for principal arterials. An important conclusion from this chapter is that the results show that data quality and data processing decisions have a large impact on model outcomes.

Safety Decision-making

The previous sections examined specification and measurement errors in crash frequency models. I used two datasets, one from Pennsylvania and one from North Carolina to estimate link-based models (with only geometric variables) and combined models (with both geometric and contextual variables). My findings indicate that models specified without contextual variables are affected by specification error, as indicated by the changes in coefficient value of the geometric variables. The coefficients of most of the geometric variables reduced in magnitude upon the addition of contextual variables to the models, and the coefficients of some geometric variables changed in direction. Similar changes occurred when I ran models using a dataset with only observed AADT data, and compared it with results from models I ran using a dataset with some of the AADT data estimated by linear regression. My findings showed the importance of data quality and availability to crash frequency models.

My second research question addresses the use of these models in decision-making by agencies that are concerned with highway safety. It asks, *"How might a better understanding of the impact of error in the Highway Safety Manual affect decision-making to improve road safety?"* The perspective of road safety decision makers is key to addressing this research question. Since this question concerns the human experience regarding the subjective issue of how a thing can be used, it is best addressed using qualitative research methods. I discussed the importance of this question in my problem statement by explaining how road safety decisions impact crash related fatalities and the efficiency of public spending.

Since I make the case that much of the modeling that undergirds this decision-making is done through using the *Highway Safety Manual*, I will start by re-introducing the HSM. The *Highway Safety Manual* is a publication of the American Association of State Highway and Transportation Officials (AASHTO), with the intention of providing an optional guide for use in making and implementing road safety decisions. It is AASHTO's answer to the absence of a single authoritative document in road safety decision-making (Babar, Parkhill 2006). The *Parking Generation Manual* published by the Institute of Transportation Engineers does the same thing for parking, AASHTO's *Highway Capacity Manual* for highway capacity problems and AASHTO's *A Policy on Geometric Design of Highways and Streets (Green Book)* for highway design.

The *Highway Safety Manual* is published in three volumes. The first covers mainly systematic analysis, including network screening, diagnosis, and introduces countermeasure selection, economic appraisal, project prioritization and evaluation. The second volume covers the predictive method for rural and other two-lane roads, rural multi-lane roads, and urban and suburban arterials. The third volume specifies the crash modification factors for road segments, intersections, and other parts of road networks.

In my literature review, I examined the problems associated with using the *Highway Safety Manual,* including the well documented problems of indeterminacy and non-transferability. The crash modification factors, which are derived from the coefficients of the variables regressed on crashes in crash frequency models are treated as being determinate, as opposed to being random variables with values that vary widely (Hauer et. al., 2012). The transferability problem treats the crash modification factors as generally applying to different contexts with superficial similarities, for example, a single crash modification factor for rural two-lane roads is expected to apply to safety analysis on rural two-lane roads in both Iowa and Pennsylvania. In my research, I hypothesized that in addition to these two related problems, the omission of contextual variables from crash frequency models, and the presence of measurement error in data used for these models, serve to further limit the inferential capability of these models from which the crash modification factors are derived. It is my goal to understand how decisions are made using resources like the *Highway Safety Manual,* so as to be able to determine the ways in which these problems are transferred to road safety decision-making outcomes.

In order to better understand how specification and measurement errors affect decision-making, I asked the following questions:

a. How is the *Highway Safety Manual* used in road safety decision-making?
b. Are decision makers aware of possible problems associated with the use of the *Highway Safety Manual*?
c. How do transportation officials account for the possible problems with the use of the *Highway Safety Manual*?

d. How can better modeling practices gain ground?

I address these questions by analyzing answers from interviews I conducted of transportation officials. I also participated in a roundtable discussion organized by NJDOT's Statewide Traffic Records Coordinating Committee (STRCC), with about 28 public agency practitioners in attendance. After having conducted a number of interviews and observing that the responses to the questions did not vary, I determined that it was not necessary or even helpful to conduct my analysis of the interview responses using qualitative analysis software. I conducted one-on-one interviews with four practitioners from public agencies and three practitioners from private establishments. In selecting interview participants, I used a combination of purposeful sampling and snowball sampling. I sought participants at the Transportation Research Board Annual Meeting. A number of participants recruited from this source pointed me out to some of their colleagues from public agencies and private enterprises who they believed would be suitable as participants. The public agencies were mostly state Departments of Transportation and the private establishments were mainly consultancy firms that carry out safety analysis for public agencies. I determined the number of interviews to use for my analysis by allowing my data to reach saturation, which is the point at which I found that there was no new information surfacing.

While I expected to find that many transportation agencies use the *Highway Safety Manual* in the way intended by AASHTO, I found from my analysis of interviewee responses, that this was not the case. The method intended by AASHTO starts with the application of local traffic volume values to the appropriate safety performance function supplied in the HSM in order to determine a baseline crash figure, and then applying a crash modification factor also supplied in the HSM to this baseline crash figure in order to determine the association of a specific safety treatment. Certain entities did not use statistical analysis at all (Interviewee E. 2017), and some only used the CMFs from the HSM, and applied it to their own safety performance functions (Interviewee E. 2017). I also expected that the transportation agencies are using the HSM did so without addressing such issues as the omission of contextual variables. One reason for this is because the *Highway Safety Manual* itself does not give any treatment to this problem, even though it is mentioned in passing (AASHTO, Vol 2, pg C-19). Since the HSM, which is considered expert and authoritative knowledge in highway safety does not treat this issue, it is likely to be assumed by transportation agencies to be a non-issue, especially with uncritical use of the HSM. I discuss these findings in more detail in the following sections.

Findings

The use of the *Highway Safety Manual* in practice.

Here, I address the question of how various agencies, public and private, use the HSM in road safety decision-making. The use of statistical modeling, and more specifically, the *Highway Safety Manual* (HSM) by agencies and private consultants depends on a number of factors. As mentioned before, the use of statistical modeling is not explicitly a requirement for any state (Babar, Parkhill 2006) and as such I wanted to understand how standardization in highway safety practice could be achieved in spite of this. Most respondents believed that as a result of not being a requirement, the use and results from the *Highway Safety Manual* can vary widely (Interviewee A. 2017, Interviewee C. 2017, Interviewee D. 2017).

I gathered from all respondents that one of the most important factors affecting how the HSM is used is the nature of the project. Some projects might primarily be safety projects, while others might be projects undertaken mainly for the purpose of capacity or other non-safety related improvements, for which assessing safety is only one of several components (Interviewee D. 2017). When the former is the case, the entity carrying out the analysis is most likely experienced in carrying out statistical modeling or accustomed to having it done by a partnering research establishment. This means that when safety needs to be assessed, it is likely to be done using statistical modeling methods. When the latter is the case however, safety may be assessed using generally less advanced methods than statistical modeling. One method that I found is used by one of the private establishments I interviewed is the examination of the number of crashes in the study period by mile post (Interviewee D. 2017). This method allows for the qualitative assessment of possible explanatory factors by examining the physical and contextual attributes of the place and time of the crash frequency (Interviewee D. 2017). An example of this might be examining the curvature or median width on a segment that has been identified to be characterized by fifty crashes in the given year, compared with another segment with only five crashes.

One participant discussed expertise as another important factor that determines the use and extent of statistical modeling for safety assessment (Interviewee E. 2017). This includes expertise with using and interpreting statistical models. According to this respondent, many state agencies partner with universities or outsource safety analysis to private consultants to get around this problem with internal expertise (Interviewee E. 2017).

Uniformity in the procedure used in assessment and implementation was another factor mentioned by a respondent that can affect the use of statistical modeling, and the use of the *Highway Safety Manual* in particular (Interviewee B. 2017). The *Highway Safety Manual* does not provide SPFs and guidelines for all road types and all contexts. This means that the HSM is limited in its applicability to entire road networks (AASHTO 2010a). The outcome is that within road networks, there will be certain portions of the network for which a different method might be required for safety analysis. This raises many questions for consistency of results and the ability to compare across such portions of the network and as such, agencies may often hesitate to use the HSM for safety planning across entire road networks (Interviewee E. 2017).

To the more specific question of whether or not an agency uses the *Highway Safety Manual* in its statistical modeling, I learned from at least one agency that the potential of the project to win federal funding is a factor (Interviewee E. 2017).

To the question of how the *Highway Safety Manual* is actually used, most respondents addressed in accordance with my expectation. They use the HSM safety performance functions and crash modification factors primarily for countermeasure assessment (Interviews. 2017). This means that for each of their localities, they obtain traffic volume data, which they plug into the HSM baseline safety performance function, and then apply the applicable crash modification factor in order to determine the best safety treatment, from a number of identified alternatives, is best for achieving the goal of safety. This predictive method, along with the collection of crash modification factors are the most frequently used part of the HSM, even though some agencies also use the network screening procedure to determine what parts of their road network require the most attention (Interviewee E. 2017).

In summary, I found that the primary objective of the project, the availability of statistical expertise, the ability to use the *Highway Safety Manual* network-wide, and the funding source were all factors determining the use of the HSM. These factors should not be seen as the only factors determining the use of the HSM, but should instead be seen as an itemization of the most common factors I encountered while interviewing transportation agencies.

Availability of expertise is one of the identified needs that the *Highway Safety Manual* was published to address (AASHTO 2010a). It is therefore interesting that several interviewees mentioned this as a factor affecting whether or not transportation agencies use the HSM. While the HSM was published so that transportation agencies could use it as a guide in the statistical analysis of crashes, it limits its own use because it requires an understanding of statistical models. FHWA runs several training programs in the use of the HSM in response to this need (AASHTO 2013). The *Highway Safety Manual* does help to reduce the burden of needed expertise since the local jurisdiction need not come up with its own safety performance functions, or crash modification factors. I discuss findings as to whether or not the HSM meets the overall goal of putting statistical analysis within reach of practitioners in the *Gains and Challenges* section below. Another important question about the limited expertise to carry out statistical safety analysis is whether or not the non-statistical methods used instead are rigorous enough in making sound safety assessments.

Gains made through the use of the *Highway Safety Manual*

The question of possible problems or challenges with the use of the HSM, and the awareness of problems such as the potential for erroneous inferences due to specification error is addressed in the next section. In this section, I discuss the gains made as an important preface.

One of the most important gains discussed by respondents that the emergence of the *Highway Safety Manual* has brought about is the extension of the analytic and policy making capabilities of public and private transportation agencies with limited expertise in safety analysis (Interviewee E. 2017). Since the publication of the *Highway Safety Manual*, the number of public road safety agencies that carry out statistical modeling for safety analysis is on the rise (Interviewee E. 2017). A related gain to the extension of the analytic capability of transportation agencies is the improvement of methods used. Methods undergirded by the use of statistical modeling have proven to be superior to methods of safety analysis that utilize averages, rates and other relatively simple forms of analysis since they avoid the problem of regression to the mean (LDOTD 2012a). One important gain accrued from the use of the HSM is therefore the improvement of safety analysis from previous methods. While this is an improvement, there are still issues with the quality of analysis that the HSM allows, since the findings of both my studies using Pennsylvania and North Carolina data show that the omission of contextual variables can adversely affect crash models and the inferences made from them. There are also the issues of indeterminacy and transferability.

A related gain is the issue of scale. At least one state transportation agency I interviewed stated that the publication of the HSM allowed their agency to carry out larger scale capital projects (Interviewee E. 2017). This particular agency had plenty of experience using statistical analysis already, but for spot treatments within their state's road network, and for a limited variety of safety treatment types. They carried out these kinds of safety improvements using internally developed crash modification factors, which due to the cost and effort required, necessarily limited the number of CMFs that could be developed.

Another important gain for agencies is the potential for carrying out safety analysis with increased confidence about the safety outcomes of road improvement projects (Interviews. 2017), and for obtaining federal funding (Interviewee E. 2017). Carrying out large scale safety projects became possible with the emergence of the *Highway Safety Manual*, since there are now hundreds of crash modification factors formerly unavailable for these kinds of analysis. In addition to this, it became possible to implement these projects in a defensible way, since the HSM crash modification factors were developed by an authoritative agency, using large amounts of data, and tested both internally and externally. In other words, the *Highway Safety Manual* is also used by agencies for defending the need for funding for capital projects, since the fact that it is backed by a well-known and authoritative organization with a presumably high level of research rigor, adds to the credibility of HSM-based projects. There is also the fact that the precursor to the FAST Act, SAFETEA-LU, which established the Highway Safety Improvement Program, required states to develop a Strategic Highway Safety Plan (SHSP) that was to be data driven and involve countermeasure analysis, as a requirement for receiving federal funding (Federal Highway Administration 2005). Both the Highway Safety Improvement Program and its requirement for the development of an SHSP continue under the FAST Act.

The extension of analytic capabilities through the *Highway Safety Manual* is primarily through its safety performance functions and crash modification factors. In providing safety performance functions and crash modification factors for various road and safety treatment types, the *Highway Safety Manual* allows agencies to conduct safety analysis more easily, using these ready-to-apply but necessary elements of statistical safety analysis. Safety performance functions are very technical, and their development requires a knowledge of statistics, as well as a large variety and amount of data. These skill and data limitations are significant enough to preclude the use of statistical analysis at the agency level.

Challenges created by the use of the *Highway Safety Manual*

Several interviewees cited data availability and quality issues as significant challenges to the use of the HSM (Interviewee D. 2017). First is the issue of the data requirements necessitated by the use of the *Highway Safety Manual*. While the *Highway Safety Manual* provides crash modification factors, each jurisdiction using them must apply them to their own traffic volume and road geometry data. This makes data an important precursor to the use of the HSM. Even without the use of the HSM, the maintenance of applicable data is an important part of safety analysis (Ogle 2007). At the federal government level, guidelines are in place to create uniformity and compliance with data collection through the Model Minimum Uniform Crash Criteria (MMUCC) for crash data (NHTSA 2008), the Model Inventory of Roadway Elements (MIRE) for roadway data and the National EMS Information System (NEMSIS) for crash victim data (Council, Harkey et al. 2007). Collecting crash data, especially based on any of these guidelines is a very labor intensive and costly undertaking, making this a barrier to the use of data-driven and statistical safety analysis. The cost of owning and maintaining data collection or measurement equipment such as GPS devices is a practical example of such a limitation. These factors have negatively impacted the quantity and quality of data collected. For example, in many states, the crash reporting thresholds are very high, so that many property damage crashes are not reported

(Council, Harkey 2006). This makes using the dependent variable of total crashes yield less accurate results than using crashes with fatalities as a dependent variable.

As seen in the North Carolina and Pennsylvania datasets, traffic volume data is often not available for all segments of all roads, due to the cost and labor required to collect it. For an agency that is accustomed to less sophisticated methods of safety analysis, this data requirement can be a limitation to the use of statistical modeling, even with the *Highway Safety Manual* as a guide (Interviewee E. 2017).

Another data issue is related to the problem of crash thresholds. Crash thresholds are lower property damage value limits that are used to determine what automobile crashes are reported in the statewide repository of automobile crashes that state Departments of Transportation collect. In many states, reported crash totals for various portions of the road network do not reflect the actual crash totals because some crashes do not reach the established thresholds in value (Interviewee E. 2017). Thresholds vary widely by states, for example, four states have a $0 threshold, mandating that all crash occurrences be reported, one state has a $100 threshold, 10 states have a $500 threshold, 20 states have a $1000 threshold, and two states have $2000 and $3000 crash thresholds (NJDOT 2017). It is reported that in some states, up to 20% of crash occurrences are not reported because they are deemed to fall below the crash threshold by the responding officer (NJDOT 2017).

Data issues are further complicated by the difficulty of collecting crash site data (Council, Harkey 2006). When a police officer arrives at the site of a crash occurrence, they are met with many time-sensitive needs. They will need to collect locational data, often including street names, physical descriptions and the geographic coordinates or mile posts of the crash occurrence. They will also need to collect names, license, registration and insurance information of the drivers involved in the crash, assuming there are no injuries or fatalities that need to be attended to. They will need to call for the removal of disabled vehicles, or for assistance in managing traffic flow in the event that the crash has caused an obstruction, or for emergency vehicles in the event that there are injuries or fatalities from the crash. In addition to performing as many of these tasks as needed, the police officer or officers will need to prioritize them in order of importance, so that it will be possible to carry out as many of them that are needed as possible. This is very challenging, and the difficulty often leads to the failure to collect the right kind, quality and amount of data, especially when there are more pressing needs like ensuring the provision of medical assistance in the case of an injury (Interviewee E. 2017).

Other problems such as the technological requirements of collecting data present challenges for safety improvement analysis came up in discussion with respondents. An example is in the collection of geographic coordinates for crash occurrences, which requires the use of devices enabled for geographic positioning (Interviewee E. 2017). Some police departments are equipped with such devices while others are not.

While the *Highway Safety Manual* has the benefit of extending the expertise and statistical analysis of transportation agencies, it can also in the same vein, limit it. Dependence on a publication like the *Highway Safety Manual* can reduce the motivation to truly understand the workings behind the processes being used, since it can be used by plugging in the ready-to-apply statistical elements it contains and results can easily be interpreted using guidelines in the HSM (Interviewee E. 2017). This means that it becomes easy to be unaware of adjustments that need to be made when new findings about methodology emerge, since publications like the *Highway Safety Manual* cannot be updated at the pace by which the research industry puts out findings in peer reviewed journals. The *Highway Safety Manual* was published in 2010, and since then, a vast amount of research produced by independent researchers exploring the validity and applicability of the HSM's recommendation has emerged with the potential to greatly improve safety outcomes for projects where safety treatments were implemented based on the use of the *Highway Safety Manual*. This means that an agency with the right knowledge and expertise could have procedures in place that are more advanced than those recommended by the *Highway Safety Manual*, making it potentially outdated for their purposes (Interviewee E. 2017).

Varying crash thresholds can present an internal validity problem. Much of the research that informed the crash modification factors in the *Highway Safety Manual* are based in states that have high thresholds for reporting crashes. This means that the models used in estimating these CMF, have more zero-crash road segments than in actuality, since many crashes are not reported (crashes with property damage values below the reporting threshold are not reported). The CMF are thus likely to be affected by any bias introduced due to the assumption of more zero crash occurrences than there would have been in reality. Applying such crash modification factors to various jurisdictions which in turn, have differing thresholds than those used in the research that undergirds the *Highway Safety Manual* SPFs will in turn yield bias. This problem of crash reporting thresholds is a difficult one because these thresholds exist for a defensible reason. They exist because of the limited time and resources that challenge the collection of data at the crash site.

Awareness and response to the specification error problem

Having discussed gains and problems of using the HSM, as well as the awareness of the specification error problem, I address the responses of transportation professionals in this section. It was my finding that the problem of specification error due to the omission of contextual variables was not a high priority issue for many public or private agencies. One respondent acknowledged the potential and need to improve statistical models for crash analysis, with the caveat that the knowledge and skills of practitioners could be a limiting factor, creating the need to keep models as simple as possible, with regards to methodological considerations such as model type and model specification (Interviewee E. 2017). The above-discussed problem of data availability is another factor affecting the use of models that include contextual variables.

Another respondent questioned the importance of accuracy. Many state agencies use the *Highway Safety Manual* to make comparisons between two or more safety countermeasures that are being considered for implementation. For many state agencies, several sites within the road network vie for limited safety funds for the implementation of adequate safety treatments. Models showing the associations of safety treatments at the various sites are compared and the site showing the highest level of crash reduction due to a specific safety treatment is usually chosen from among the alternatives for safety treatment implementation (Interviewee E. 2017). This raises the question of the importance of accuracy, if an agency is implementing the safety treatment that yields the highest crash reduction out of several safety treatment options, versus implementing a safety treatment in response to a targeted safety issue.

The issue of limiting the complexity of statistical models is tied to the goal of using models as forecasting tools instead of as inferential tools. Agencies are not motivated to spend time, effort and funding on collecting data for variables that they cannot act on (Interviewee E. 2017). This means collecting data for primarily geometric variables and not so much for contextual variables since action can be taken on design changes to such variables as median, pavement, and shoulder width, lane count and curvature for road segments if the models forecast such changes to be effective at reducing crashes. In the face of limited time and resources, the importance of avoiding bias by including contextual variables seems like a mere theoretical exercise that cannot be justified.

On the issue of the importance of accuracy when comparing the relative efficacy of alternative safety treatments, the assumption is that the models showing the crash reduction effects of the various alternatives being considered are comparable. If this assumption is true, then the bias found in the various model alternatives may be expected to be the same, as opposed to potentially yielding different effects on crash frequency. On the other hand, if the bias is different, then it is possible that the crash reduction effects may be affected, and the wrong conclusions drawn from a comparison. For instance, it is possible that an agency is considering making safety improvements on two different portions of a road network but will eventually choose only one location, due to limited funding. The choice will often be based on the magnitude of crash reduction that can be expected at each of those sites. Theoretically, even if the safety treatments being considered at those sites are different, e.g. site A is being assessed for a median barrier installation and site B is being assessed for a reduction in number of lanes, it should be possible to compare results, if the right variables have been specified and controlled for. Where specification error comes into play here and potentially complicates comparability, is where omitted variables may differentially affect safety performance functions. This is most clearly seen in the case of interaction terms in safety performance functions. For instance, if the contextual variable of median income has been omitted from both site A and site B functions, and it is known to interact with the lane count variable, but not with the median width variable, then the site A SPF has been adequately specified, while the site B SPF has not, meaning that comparison between the crash reduction in site A and B will yield some problematic inferences.

While I did not find interactions between the variables I explored in my North Carolina dataset, many crash frequency studies have explored and found interactions suggesting that in many cases, road characteristics are not simply additive (Wang et al. 2016). Even without the potential for interaction between omitted and specified variables, the fact that a variable such as median income might be different for site A and site B will mean that its omission will have differential effects on the site A and B SPFs, rendering them less comparable. An agency could be choosing the less optimal alternative for implementation.

Evaluation of road safety decisions based on the *Highway Safety Manual*

Safety project evaluation is important for addressing the question of how better modeling practices can gain ground. As part of the *Highway Safety Manual* road safety management process, the evaluation of the effectiveness of implemented safety treatments is recommended. The HSM contains guidelines for three main observational methods. They include the observational before/after studies (naïve), observational before/after using the Empirical Bayes method and observational studies using the comparison group method. There are also guidelines for conducting experimental before/after studies (AASHTO 2010c).

Not many of the agencies I interviewed had carried out evaluations of their statistical modeling based on the *Highway Safety Manual* guidelines. Some respondents discussed time as a limiting factor, stating that not enough time had elapsed since their adoption of

HSM guidelines through safety design implementation, for evaluation to take place (Interviewee E. 2017). One respondent was very experienced in carrying out evaluations, as a system was already in place prior to their adoption of the HSM (Interviewee E. 2017). In addition to using the HSM to forecast the safety effects of future projects, this agency uses the HSM for evaluation as an added dimension to various kinds of before-after observational studies including naïve before and after analysis, Empirical Bayes before and after analysis, and comparison group before and after analysis.

In a naïve before and after study, the evaluation is carried out by taking the ratio of the value of the performance measure, for instance, crash frequency in the after period, to its value in the before period. A key assumption is that the performance measure will not have changed during the after period from its before period value, without the implementation of the countermeasure (FHWA 2011). This assumption is obviously an oversimplification since factors such as traffic volume growth, and better vehicle safety design can cause the performance measure to change as well. This ratio of the after-period performance measure to the before-period performance measure can be used to determine crash reduction factors.

As mentioned above, this is an overly simplistic way to calculate crash modification factors, since it assumes that a change in crashes seen in the after-period is due only to the safety design implementation. An Empirical Bayes before and after analysis can be used to obtain more defensible findings on the impact of a safety countermeasure (FHWA 2011). In calculating a crash modification factor this way, the CMF is taken as a ratio of the observed crashes in the after-period, to the crashes in the after period, assuming no treatment was implemented. This number of crashes, assuming no treatment in the after period is calculated by dividing the predicted crashes assuming treatment in the after period (obtained using a safety performance function) by the predicted crashes assuming no treatment in the after period, and then multiplying by the crashes in the before period (which is a function of observed and predicted crashes in the before period).

In the case of comparison group before and after studies, the crash modification factor is obtained by calculating the ratio of the averaged observed crashes for the project sites in the after period to the average observed crashes for the control sites in the after period.

The North Carolina Department of Transportation is one agency with some experience in carrying out evaluations, with publicly available evaluation reports. Currently, over a thousand completed evaluations can be found online (NCDOT 2017). A quick perusal reveals that most of these evaluations have been carried out using the naïve before and after method, which does not incorporate safety performance functions into its assessment. In other words, even though a specific site and specific countermeasure may have been selected using statistical analysis and more specifically, SPFs, there is no post-countermeasure implementation assessment of such SPFs. There is a foregone assumption of the soundness of SPF use in the forecasting stage where a site and a specific countermeasure is chosen, and the focus is then on assessing the degree of countermeasure success or lack thereof through the comparison of the pre-countermeasure observed crash totals to the post-countermeasure observed crash totals.

Using the Empirical Bayes method however, can begin to assess the SPF used in forecasting, because this method incorporates predicted crashes into its calculation of the countermeasure crash modification factor.

One method that researchers have used to evaluate the *Highway Safety Manual* SPFs is through the process of model validation. I discussed a study by Cunto et al (2015) in the literature review section that outlined this process. The aim of the study was to specifically assess the applicability of the SFPs in a real life context, or the SPFs transferability. The study used comparison groups to carry out the evaluation. The authors used the calibration group to come up with a complete safety performance function that included a crash modification factor for the presence of roadway lighting and a calibration factor (derived by taking the ratio of observed crashes to predicted crashes for the calibration sample) for the adjustment of the model parameters to the local jurisdiction being studied, which was Fortaleza, Brazil. The same crash modification factor derived from the HSM, and the calculated calibration factor were then applied to the validation sample in order to assess how much in line the predicted crashes were, with the observed crashes from the validation sample. The results showed the validation to have been unsuccessful, leading the study authors to conclude that a combination of other variables or other model forms might have worked better to improve the crash predictions (Cunto et al, 2015). The HSM recommends a similar procedure that derives calibration factors by comparing predicted crashes to observed crashes. The Cunto et al, (2015) study takes it a step further, by testing the derived calibration factors against a sample with known observed crashes. I discuss the Cunto et al, (2015) method in my literature review. This kind of evaluation is necessary because it allows an agency to know whether or not the safety analysis methods they are using are adequate, or in need of modification.

Summary of findings

The following is a summary of the findings from my interviews about road safety decision-making and how specification and measurement error affect it. This summary is based on my research question of how safety decisions are made using inferences from safety models.

1. The *Highway Safety Manual* is used to expand analytic capacity beyond the use of simple analytic methods that may introduce such bias as regression to the mean, for more statistically sound inferences and potentially improved decision-making.
2. It is used as a forecasting tool, for the estimation of expected crashes that any particular safety design improvement might yield.
3. It is used in the evaluation of implemented safety design improvements, in conjunction with, and in addition to other methods such as naïve before-after and Empirical Bayes before-after safety evaluations.
4. It is used in building defensible cases for garnering federal and other public funding for capital projects.

5. Data requirements, particularly for traffic volume and observed crashes can be a problem for agencies, who often don't collect traffic volume data on all their roads.
6. The requirement for the application of calibration factors to the safety performance functions in the forecasting stage, can be a deterrent to use because data collection can prove to be an expensive effort, in addition to the fact that it can have varying levels of success or lack thereof, in making the safety performance functions better fitted to the local jurisdiction.
7. The specific problem of omitted variable bias is not necessarily a priority for many agencies at this moment. Some agencies acknowledged this as a possibility but expressed some confidence that the use of calibration factors will account for it. Other agencies expressed the need to keep safety performance functions as simple as possible, and the cost in time, money and efforts to collect more data in response to this problem are limiting factors.
8. The range of transportation agency response to the possibility of specification error from omitted variable bias is from reliance on calibration factors, to no response.
9. Most professionals interviewed believe that improvements can be best made in the area of data collection and expertise.

Conclusions

This dissertation research addresses the problem of indeterminacy in the methods used in crash frequency modelling, and in the crash modification factors they inform, due to omitted variable bias and data problems. The issue of indeterminacy has been previously addressed in crash frequency studies. The nature of crash modification factors has been examined by a number of researchers, and there is some consensus that they are not constants, even though they are treated as such in the *Highway Safety Manual*. The issue of the omission of spatial factors, leading to the problem of omitted variable bias is however largely unexamined. In addition to examining this problem, I also address the issue of methodological complexities that arise in the process of accounting for spatial factors. These methodological complexities can introduce other problems into the models, and they show that a lot of indeterminacy comes from analyst research decisions in circumnavigating such complexities. Some examples of these decisions include the following:

- What level of geography should the spatial variables be based on (zonal, census tract, census block group etc.). This is also known as the modifiable areal unit problem.
- Do all spatial factors affect crashes occurring on road segments at the same level of geography?
- What combination of spatial variables should be specified in models? This is also a valid question for models specifying only geometric variables.

- What is the best way to treat missing data? Methods commonly include estimating from observations with data or dropping all observations with missing data. Each option affects outcomes in a different way.
- Should interactions be explored? The *Highway Safety Manual* uses crash modification factors from studies which mostly leave out possible interactions.

Finally, I examined the question of how indeterminacy affects decision making.

The research questions I investigated include the question of specification and measurement error in crash frequency models, the impact of data availability and quality problems on the ability to make correct inferences, and the impact of awareness of these indeterminacy problems on decision making at the transportation agency level.

In order to determine if specification and measurement error are important factors affecting road safety decision-making, I examined the results of crash frequency analyses using two different datasets from Pennsylvania and North Carolina. I also conducted qualitative interviews of transportation professionals working for both public and private entities. For the both datasets, I compared models specified with only geometric variables (link-based models) with models specified with both geometric variables and contextual variables (combined models) in order to determine the effect of the presence of those contextual variables on the geometric variables. In the second dataset, I also compared models with observed traffic volume data, with models with estimated traffic volume data, in order to determine how important measurement error is. Finally, using qualitative interviews, I explored the ways by which transportation agencies use inferences from crash frequency analysis in road safety decision-making.

For my Pennsylvania dataset, I first examined crash frequency models using the entire network. I then modeled subsets of the data- principal arterial highways and local roads. I also modeled each subset with the negative binomial regression (MLE), and then with conditional autoregressive models (MCMC), using *Crimestat,* to control for any possible spatial autocorrelation.

For my North Carolina dataset, I used only interstate highways because most of the functional classification groups of highways were skewed towards zero crash occurrences, and interstates were least affected by this problem. I explored the issue of data availability in this dataset using AADT. AADT is typically not collected for all roads in any dataset and as such is the most likely variable to be affected by data availability issues. I carried out the research in two stages, the first stage was with 8071 observations, all with data for observed AADT, and the second stage with 8071, but with 30% of the observations with AADT estimated. For each stage, I ran negative binomial models, and then ran conditional autoregressive models to account for spatial autocorrelation. I compared these two stages in order to see if there is any effect from the use of estimated AADT.

Following this, I explored the impact of specifying more spatial variables. In addition to the previously studied variables of population and employment density and median income, I explored the impacts of elevation, precipitation and two age variables-the proportion of people between ages 18 and 24 (%18-24) and the proportion of people 65 and older (%65-up). Finally, I explored the impact of accounting for possible interactions between certain variables including interactions between pavement width and population density, pavement width and elevation, lane count and %65-up, lane count and %18-24, and finally, sinuosity and %65-up.

In the last chapter of analysis, I interviewed several public and private transportation agency practitioners in order to assess the level of awareness of the problem of indeterminacy, and their approach to making the best possible safety decisions, despite this problem. I carried out a series of one-on-one interviews, and also attended a safety workshop.

From the first set of analyses in the Pennsylvania dataset (MLE models with the entire road network), all the combined models showed decreases in the magnitudes of the geometric variables, when compared to the link-based models, except for sinuosity in the fatal and incapacitating injury model, which showed a small increase. None of the geometric variables in the combined models show a change in the direction of association except for sinuosity in the fatal and injury model which changes from a negative association in the link-based model to a positive association in the combined model. Since my hypothesis is that the omission of spatial variables causes bias that is reflected as decreased magnitudes of association, or as changes in the direction of association in the geometric variables when specified in models with spatial variables, my conclusion for this first set of analysis is that my hypothesis is not disproved.

For the second set of analyses for my Pennsylvania dataset where I modelled the three dependent variables using only principal arterials, the majority of the geometric variables decreased in magnitude of association, except sinuosity in all three MCMC models, and lane count and median width in the fatal and major injury crashes model. For all three dependent variables, lane count changes direction of effect from negative in the link-based model to positive in the combined model. The reduction in magnitude of effect generally seen in most of the geometric variables and the change in direction of effect in the lane count variable, support my hypothesis of specification error from the omission of spatial variables. The reduction in magnitude of association and the change in direction of lane count were almost exactly replicated by the MLE models. Finally, for the third set of analysis (local roads) for the Pennsylvania dataset, the majority of the geometric variables decreased in magnitude of association, but no variables changed direction of association.

All three sets of analyses support my hypothesis that specification error from the omission of spatial variables is a problem in crash frequency models. The bias seen in crash frequency models will vary based on the model specification, thereby resulting in indeterminate crash modification factors. The analyses in my Pennsylvania chapter show how analyst decisions on model specification contributes to indeterminacy.

For the North Carolina dataset, I again ran link-based models and combined models to examine the specification error problem. In addition to this, I examined the impact of measurement error, by simulating missing AADT data and replacing the missing data with estimated AADT for 30% of the observations I modeled. The models with estimated AADT were compared with the models with 100% observed AADT. The results showed that there were substantial differences. The total crashes model showed different directions of magnitude change for several variables when the estimated AADT model was compared with the observed AADT model. It also showed opposite directions of effect for shoulder width, and for sinuosity in the estimated AADT model when compared with the observed AADT model.

In the fatal and major injury crashes models, the direction of magnitude change in the model with observed AADT was also not consistent with that of the estimated AADT model. Several variables showed increases, and others showed decreases. Total width and lane count also show completely different directions of association (total width is positive in the observed AADT model and negative in the estimated AADT model, and lane count is negative in the observed AADT model and positive in the estimated AADT model. In the fatal and injury models, the same variables- shoulder width and sinuosity have different directions of association between the observed AADT and estimated AADT models, just as was seen in the total crashes model. There is more consistency between these two models for the fatal and injury crashes model in terms of direction of change in magnitude of association than for any other dependent variable in the dataset. These results largely show how widely crash modification factors can be impacted, even when just a small proportion (in this case 30%) of the dataset is affected by measurement error introduced through analyst methods.

To further assess the issue of indeterminacy due to analyst decisions, I examined the combined model with estimated AADT with elevation, precipitation, the proportion of people between ages 18 and 24 (%18-24) and the proportion of people 65 years and over (%65up). An average of two variables per model changed direction of association for all three models when compared with the corresponding model without the additional variables. In the total crashes model, shoulder width changed to a positive association, in the fatal and incapacitating injury crashes model, total width takes on a negative association, and sinuosity takes on a positive association. For the fatal and injury model, shoulder width changes to a positive association, as was seen in the total crashes model. The results from my North Carolina study showed that analyst methods of addressing data availability problems can contribute to indeterminacy.

In the final section of my North Carolina study, I examined the possible interactions between different sets of geometric and contextual variables. I examined possible interactions between pavement width and population density, pavement width and elevation, lane count and %65-up, lane count and %18-24, and sinuosity and %65-up for interstate highways. None of these combinations showed any interactions. To further explore the possibility of interactions, I examined the population density-pavement width combination, using only principal arterials. The results showed that there may be some interaction. The interaction terms are significant, with the high population density-total width interaction significant at the 99% level, and the moderate population density-total width significant at the 90% level. I also examined a possible interaction between %18-24 and lane count. All the interaction terms for this combination were significant at the 99% level. The coefficients of the moderate and high categories of %18-24 suggests that a higher composition of people aged 18-24 interacts with lane count to increase crash frequency.

The main goal of my dissertation study has been to explore the impact of indeterminacy on highway safety decision making. I examined this indeterminacy problem by investigating the impact of context and of research method decisions. The results of my research show how the coefficients of crash frequency models can easily change direction of association with crash frequency with different variable specifications. This is important because the combination of variables that models are specified with is dependent on the analyst decisions. While there are variables that are commonly specified across studies, the exact combination of variables specified always differ. Factors such as data availability play an important role in a researcher's decision about model specification. My results also show that it is possible that these models are typically under-specified because they omit contextual variables which have been shown to be strongly associated with crash frequency in a number of studies. I also found that data availability problems not only determine what variables a researcher specifies in a crash frequency model, they also determine the representativeness of variables that are included. Certain variables like AADT are considered important to models such that when there are data availability problems for these variables, it is no longer a question of whether to omit or include the variable, but a question of how to best estimate missing data for inclusion in crash frequency models. My results show that the methods employed in estimating these variables can also have a substantial contribution to the indeterminacy of crash modification factors. The overall implication of my results is that highway safety improvement, using crash modification factors that have not been locally estimated is a black box, since many decisions that have been key determinants of the crash modification factors are unknown. It is therefore impossible to know where errors have been introduced or how far reaching the errors are. Locally developed crash modification factors may not be any less erroneous than those developed with the goal of transferability to various geographies. They however have the advantage of traceable errors, since the research method decisions can be more easily deciphered. Knowing the research methods decisions also makes it possible to cumulatively improve upon methodology.

My dissertation has relevance for research, and for highway safety practice. It revisits old questions surrounding the usefulness and feasibility of standardized practice and the capacity of research to enable standardization. Standardization serves to facilitate practice by reducing the need to carry out certain processes, for example, the substitution of crash modification factors from the *Highway Safety Manual* in theory, removes the need for localities to develop their own crash modification factors. It also reduces the time and money cost associated with various methods for assessing safety, including developing local crash modification factors. My research has however shown that this goal of standardization depends on the nature of the input (in this case crash modification factors) and the nature of the research used to determine the input (crash frequency modeling). My research is directly relevant to the question of the adequacy of crash modeling research to the standardization of highway safety decision making. The overall importance lies in the high cost that crashes have on safety and on safety spending. With thousands of fatalities, millions of injuries and billions spent in improving safety, this cost is very high.

My dissertation research points to a number of interesting future research directions. One possible direction would be to explore the impact of the source of data on indeterminacy. Traditional means for data collection are subject to time and effort limitations. For example, police officers who collect crash related data such as geocoded location of crashes, number of fatalities or injuries, pedestrian involvement, are limited by time, the capability of measurement and recording equipment such as GPS devices, and human error. On the other hand, the use of more advanced means that are typically used in the collection of big data, for example, the use of cell phone data to determine traffic volume, or average speed, may significantly improve the representativeness the data used in crash frequency modeling.

My findings point to a few policy recommendations. One recommendation is to review the criteria for qualifying safety projects for federal funding. Many states currently have mandates for the use of the *Highway Safety Manual* for safety projects, through their Highway Safety Improvement Programs (HSIP). This means that these state agencies must use the Highway Safety Manual for safety improvement projects, to qualify for federal funding. These mandates do not of course preclude the development of local safety performance functions by state transportation agencies, but they can have the effect of discouraging it. This is because transportation agencies depend on federal funding for the implementation of many safety projects. A review of these mandates is important for the purpose of ensuring that the use of local safety performance functions does not prevent access to federal funding for safety projects. Another policy recommendation is a more uniform collection of data on key variables such as traffic volume through the use of emerging data collection methods such as satellite or personal device positioning.

Appendix

Appendix A: Interview Guide

I have identified 4 questions that arise from the research question and form the basis for the interview questions I used in gaining the perspectives of the decision makers and consultants.

1. How is the *Highway Safety Manual* used in road safety decision-making?

   Some interview questions to help address this question might be:

   i. How does your agency use statistical modeling in road safety decision-making?

   ii. Is statistical modeling a required procedure in road safety decision-making for your jurisdiction? (if so, by whom?)

   iii. What tools, resources or manuals has your agency used besides the *Highway Safety Manual* for statistical modeling?

   iv. How does your agency incorporate the use of the *Highway Safety Manual* as part of its decision-making process?

   v. How long has your agency, or consultants working with your agency, used the *Highway Safety Manual* in predicting crash frequency?

   vi. What are some gains made to your processes by the use of the *Highway Safety Manual*?

vii. What are some of the challenges created that have emerged through the use of the *Highway Safety Manual*?

2. Are decision makers aware of possible problems associated with the use of the *Highway Safety Manual*? Some interview questions might be:
    i. What other general problems arise in making decisions about road geometry modifications?
    ii. How would you rank these general problems and the challenges created by the *Highway Safety Manual*?
    iii. What is your view of the effectiveness of the road safety decisions your agency has made in the past 10 years?
    iv. What is the basis for your view? (e.g. technical basis such as CBA, post-project evaluation or other)
    v. How important is possible specification or measurement error to realizing the expected safety associations pertaining to a specific road safety modification?
    vi. How important is possible specification or measurement error to efficient spending on safety modifications?

3. How are transportation officials accounting for the possible problems with the use of the *Highway Safety Manual*? Some interview questions might be:

i. In what ways are the processes you have in place designed to account for the impact of safety data quality problems? (This question assumes a consciousness of possible data issues typically found in data used in modeling crash frequency)

ii. In what ways are the processes you have in place designed to minimize the impact of specification and measurement error? (This question assumes the agency is conscious of possible specification and measurement error problems from the HSM. It is irrelevant if the agency is not aware of this or other sources of specification or measurement error in its procedures).

4. How can better modeling practices gain ground? Some interview questions might be:

   i. How important is decision-making backed by an institutional authority, such as is the case with decision-making based on the *Highway Safety Manual* recommendations?

   ii. How important is the relative simplicity that the *Highway Safety Manual* offers?

iii. What is the potential of modeling procedures that account for specification error to be widely used in place of the *Highway Safety Manual*?

Appendix B: Pennsylvania Study MLE Models

Table 47: Spatial MLE Models (Entire network)

	***Model** 42*	***Model** 43*	***Model** 44*
VARIABLES	Crashes	Fatal & Major Injury Crashes	Fatal & Injury Crashes
Population density (sq mi, ln)	-0.448***	-0.547***	-0.425***
	(0.0109)	(0.0143)	(0.0110)
Employment density (sq mi, ln)	0.212***	0.132***	0.203***
	(0.00581)	(0.00787)	(0.00602)
Median income (ln)	0.0434*	0.0646**	0.0129
	(0.0252)	(0.0258)	(0.0211)
Interstates density (ln)	0.343***	0.261***	0.370***
	(0.0397)	(0.0463)	(0.0383)
Principal arterials density (ln)	0.0834***	0.0917***	0.167***
	(0.0174)	(0.0239)	(0.0167)
Minor arterials density (ln)	0.0200	0.0166	0.100***
	(0.0165)	(0.0233)	(0.0159)
Collectors density (ln)	-0.00971	-0.0400*	0.0516***
	(0.0162)	(0.0240)	(0.0158)
Local access roads density (ln)	0.00591	0.0825	0.0809*
	(0.0410)	(0.0525)	(0.0416)
Interchange ramps density (ln)	0.449**	0.285	0.573***
	(0.208)	(0.203)	(0.218)
Constant	4.730***	1.960***	4.254***
	(0.278)	(0.285)	(0.236)
	-0.708***	-0.764***	-0.673***
	(0.0166)	(0.0364)	(0.0166)
Observations	9,740	9,740	9,740
Log likelihood	-49266	-18321	-43342
Ll Constant Only	-51691	-20472	-45266
LR Chi2	3617	4110	2414
Pseudo_R2	0.0469	0.105	0.0425
Robust standard errors in parentheses			
*** p<0.01, ** p<0.05, * p<0.1			

Table 48: MLE Models with only Principal Arterial Roads (FC B)

	***Model** 45*	***Model** 46*	***Model** 47*	***Model** 48*	***Model** 49*	***Model** 50*
VARIABLES	**Crashes**	**Crashes**	**Fatal & Major Injury Crashes**	**Fatal & Major Injury Crashes**	**Fatal & Injury Crashes**	**Fatal & Injury Crashes**
Total width (ln)	1.182***	0.362***	0.781***	0.315***	1.431***	0.451***
	(0.0330)	(0.0362)	(0.0595)	(0.0640)	(0.0371)	(0.0375)
Lane count (ln)	-0.527***	0.257***	-0.0310	0.361***	-0.599***	0.256***
	(0.0580)	(0.0666)	(0.109)	(0.109)	(0.0652)	(0.0693)
Median width (ln)	-0.103***	-0.103***	-0.131***	-0.124***	-0.121***	-0.130***
	(0.00463)	(0.00533)	(0.00959)	(0.0107)	(0.00523)	(0.00615)
VMT (ln)	0.389***	0.354***	0.417***	0.402***	0.390***	0.355***
	(0.00640)	(0.0107)	(0.0147)	(0.0155)	(0.00755)	(0.0109)
Sinuosity (ln)	0.213**	0.576***	0.461**	0.633***	0.379***	0.830***
	(0.0985)	(0.0880)	(0.207)	(0.171)	(0.115)	(0.108)
Median income (ln)		0.00514		-0.166***		-0.0750***
		(0.0179)		(0.0211)		(0.0169)
Population density (sq mi, ln)		0.180***		0.168***		0.226***
		(0.00968)		(0.0182)		(0.0105)
Employment density (sq mi, ln)		0.176***		0.0181		0.174***
		(0.00656)		(0.0119)		(0.00715)
Constant	-3.795***	-3.801***	-6.921***	-4.814***	-5.254***	-4.196***
	(0.0985)	(0.223)	(0.206)	(0.303)	(0.114)	(0.218)
	-0.135***	-0.558***	-0.465***	-0.635***	0.0578***	-0.413***
	(0.0114)	(0.0142)	(0.0587)	(0.0669)	(0.0125)	(0.0157)
Observations	17,859	17,859	17,859	17,859	17,859	17,859
Log likelihood	-63640	-60209	-14589	-14387	-53281	-50022
Ll Constant Only	-66634	-66634	-15420	-15420	-56002	-56002
LR Chi2	5988	14896	1663	1893	5441	14249
Pseudo_R2	0.0449	0.0964	0.0539	0.0670	0.0486	0.107
Standard errors in parentheses						
*** p<0.01, ** p<0.05, * p<0.1						

Table 49: MLE Models with only Local Roads (FC E)

	***Model** 51*	***Model** 52*	***Model** 53*	***Model** 54*	***Model** 55*	***Model** 56*

VARIABLES	Crashes	Crashes	Fatal & Major Injury Crashes	Fatal & Major Injury Crashes	Fatal & Injury Crashes	Fatal & Injury Crashes
Total width (ln)	1.675***	1.178***	0.835***	0.767***	1.569***	1.114***
	(0.0727)	(0.0820)	(0.194)	(0.186)	(0.0868)	(0.0924)
Lane count (ln)	-2.897***	-1.737***	-2.373***	-2.314***	-3.034***	-1.991***
	(0.181)	(0.228)	(0.434)	(0.469)	(0.214)	(0.247)
Median width (ln)	-0.242***	-0.214***	-0.255***	-0.266***	-0.299***	-0.279***
	(0.0123)	(0.0145)	(0.0299)	(0.0308)	(0.0143)	(0.0184)
VMT (ln)	0.728***	0.631***	0.666***	0.686***	0.742***	0.659***
	(0.00760)	(0.00995)	(0.0220)	(0.0250)	(0.00929)	(0.0112)
Sinuosity (ln)	-0.691***	-0.474**	0.107	0.0389	-0.446**	-0.206
	(0.177)	(0.236)	(0.548)	(0.459)	(0.224)	(0.223)
Median income (ln)		0.238***		-0.446***		0.00473
		(0.0390)		(0.0996)		(0.0444)
Population density (sq mi, ln)		0.0761***		-0.0308		0.0884***
		(0.0172)		(0.0427)		(0.0201)
Employment density (sq mi, ln)		0.147***		0.0266		0.124***
		(0.00911)		(0.0246)		(0.0108)
Constant	-5.056***	-7.689***	-6.289***	-1.337	-5.502***	-5.643***
	(0.250)	(0.526)	(0.630)	(1.353)	(0.296)	(0.584)
	-0.360***	-0.485***	-0.124	-0.175	-0.305***	-0.406***
	(0.0231)	(0.0290)	(0.137)	(0.142)	(0.0304)	(0.0370)
Observations	18,965	18,965	18,965	18,965	18,965	18,965
Log likelihood	-29337	-28842	-5137	-5125	-21299	-21063
Ll Constant Only	-35382	-35382	-5899	-5899	-25859	-25859
LR Chi2	12091	15002	1525	1691	9119	10998
Pseudo_R2	0.171	0.185	0.129	0.131	0.176	0.185
Standard errors in parentheses						
*** p<0.01, ** p<0.05, * p<0.1						

Appendix C: North Carolina Study

Table 50: MLE Models (Spatial variables only)

VARIABLES	Total Crashes	Fatal and Incapacitating Injury Crashes	Fatal and Injury Crashes
Population Density (ln)	-0.342***	-0.327***	-0.492***
	(0.0200)	(0.0183)	(0.0158)
Employment Density (ln)	0.136***	0.126***	0.0324***
	(0.0133)	(0.0119)	(0.0101)
Median Income (ln)	0.0723***	0.101***	0.146***
	(0.0280)	(0.0269)	(0.0258)
Interstate Density (ln)	1.406***	1.238***	0.859***
	(0.0686)	(0.0630)	(0.0676)
Principal Arterials Density (ln)	0.811***	0.864***	0.653***
	(0.0461)	(0.0432)	(0.0443)
Minor Arterials Density (ln)	0.562***	0.536***	0.300***
	(0.0445)	(0.0407)	(0.0407)
Collectors Density (ln)	0.262***	0.281***	0.264***
	(0.0446)	(0.0406)	(0.0441)
Local Roads Density (ln)	0.224***	0.227***	0.326***
	(0.0252)	(0.0238)	(0.0281)
Constant	4.398***	2.908***	0.399
	(0.294)	(0.286)	(0.272)
Observations	6,155	6,155	6,155
Log likelihood	-37543	-30203	-12211
Ll Constant Only	-38443	-31148	-13332
LR Chi2	1157	1212	1960
Pseudo_R2	0.0234	0.0304	0.0840
Robust standard errors in parentheses			
*** p<0.01, ** p<0.05, * p<0.1			

Table 51: MLE Link-based and Combined Models (Observed AADT)

VARIABLES	Crashes	Crashes	Fatal & Incapacitating Injury Crashes	Fatal & Incapacitating Injury Crashes	Fatal & Injury Crashes	Fatal & Injury Crashes
Pavement Width (ln)	-0.205***	-0.247***	-0.167	-0.179	-0.226***	-0.273***
	(0.0542)	(0.0703)	(0.130)	(0.125)	(0.0595)	(0.0714)
Median Width (ln)	-0.124***	-0.0828***	0.0550	0.0487	-0.101***	-0.0518**
	(0.0139)	(0.0237)	(0.0419)	(0.0435)	(0.0160)	(0.0234)
Shoulder Width (ln)	-0.104***	-0.0486	-0.143	-0.147	-0.175***	-0.107
	(0.0350)	(0.0632)	(0.0885)	(0.0953)	(0.0376)	(0.0675)
Lanes (ln)	1.360***	0.681***	-0.0802	-0.0316	1.252***	0.576***
	(0.0864)	(0.137)	(0.203)	(0.220)	(0.0938)	(0.134)
Sinuosity (ln)	3.090	-5.680***	-1.900	-1.247	3.351	-5.369**
	(3.172)	(2.058)	(3.518)	(3.375)	(3.102)	(2.255)
Vehicle Miles Traveled (ln)	0.650***	0.641***	0.724***	0.729***	0.656***	0.647***
	(0.00930)	(0.0130)	(0.0260)	(0.0302)	(0.0108)	(0.0136)
Population Density (ln)		0.0278		0.0555		0.101***
		(0.0289)		(0.0510)		(0.0257)
Employment Density (ln)		0.170***		-0.0456		0.149***
		(0.0160)		(0.0286)		(0.0148)
Median Income (ln)		0.0134		-0.101***		-0.0153
		(0.0220)		(0.0315)		(0.0198)
Constant	-6.172***	-0.107	-6.102**	-5.477**	-7.436***	-1.290
	(2.203)	(1.453)	(2.467)	(2.402)	(2.157)	(1.604)
Observations	8,071	8,071	8,071	8,071	8,071	8,071
Log likelihood	-23251	-23010	-2420	-2416	-14249	-14020
Ll Constant Only	-25510	-25510	-2892	-2892	-16185	-16185
LR Chi2	4518	4646	943.2	675.5	3872	3943
Pseudo_R2	0.0886	0.0980	0.163	0.165	0.120	0.134
Standard errors in parentheses						
***** p<0.01, ** p<0.05, * p<0.1**						

Table 52: MLE Link-based and Combined Models (Estimated AADT)

VARIABLES	Total Crashes	Total Crashes	Fatal & Incapacitating Injury Crashes	Fatal & Incapacitating Injury Crashes	Fatal & Injury Crashes	Fatal & Injury Crashes
Pavement Width (ln)	-0.206***	-0.248***	-0.179	-0.193	-0.241***	-0.285***
	(0.0548)	(0.0701)	(0.131)	(0.126)	(0.0596)	(0.0707)
Median Width (ln)	-0.121***	-0.0812***	0.0514	0.0469	-0.0965***	-0.0520**
	(0.0141)	(0.0234)	(0.0418)	(0.0431)	(0.0161)	(0.0230)
Shoulder Width (ln)	-0.0806**	-0.0364	-0.139	-0.140	-0.149***	-0.0954
	(0.0353)	(0.0682)	(0.0885)	(0.0948)	(0.0378)	(0.0712)
Lanes (ln)	1.380***	0.705***	-0.0614	-0.0538	1.288***	0.617***
	(0.0870)	(0.138)	(0.204)	(0.222)	(0.0939)	(0.134)
Sinuosity (ln)	-1.074	-8.214***	-2.373	-1.992	-0.966	-7.877***
	(3.571)	(2.435)	(3.609)	(3.482)	(3.445)	(2.663)
VMT (ln)	0.601***	0.608***	0.705***	0.710***	0.604***	0.612***
	(0.00955)	(0.0133)	(0.0267)	(0.0311)	(0.0109)	(0.0140)
Residuals	1.07e-05***	6.74e-06***	2.76e-06	2.47e-06	1.13e-05***	7.49e-06***
	(7.56e-07)	(1.12e-06)	(1.73e-06)	(1.75e-06)	(7.94e-07)	(1.05e-06)
Population Density (ln)		0.0217		0.0583		0.0832***
		(0.0290)		(0.0538)		(0.0264)
Employment Density (ln)		0.170***		-0.0342		0.152***
		(0.0159)		(0.0289)		(0.0149)
Median Income (ln)		0.0249		-0.102***		0.00193
		(0.0207)		(0.0337)		(0.0191)
Constant	-3.011	1.728	-5.596**	-4.781*	-4.123*	0.533
	(2.480)	(1.720)	(2.540)	(2.477)	(2.395)	(1.890)
Observations	8,071	8,071	8,071	8,071	8,071	8,071
Log likelihood	-23328	-23109	-2436	-2432	-14293	-14088
Ll Constant Only	-25510	-25510	-2892	-2892	-16185	-16185
LR Chi2	4363	4404	911.5	677.3	3785	3755
Pseudo_R2	0.0855	0.0941	0.158	0.159	0.117	0.130
Standard errors in parentheses						
*** p<0.01, ** p<0.05, * p<0.1						

Works Cited

AASHTO, 2010a. Highway Safety Manual. Washington, D.C.

AASHTO, 2010b. An Introduction to the Highway Safety Manual. Washington, DC.: AASHTO.

ABDEL-ATY, M. and RADWAN, A.E., 2000. Modeling traffic accident occurrence and involvement. Accident Analysis and Prevention, 32, pp. 633-642.

ABDEL-ATY, M., LEE, J., SIDDIQUI, C. and CHOI, K., 2013. Geographical unit based analysis in the context of transportation safety planning. Transportation Research Part A: Policy and Practice, 49, pp. 62-75.

Asch, P., and Levy, D.T. “Young Driver Fatalities: The Roles of Drinking Age and Drinking Experience.” Southern Economic Journal, vol. 57, no. 2, 1990, pp. 512–520.

BABAR, G. and PARKHILL, M., 2006. Ongoing development of a Highway Safety Manual.

BLOMQUIST, G., 1986. A utility maximization model of driver traffic safety behavior. Accident Analysis and Prevention, 18(5), pp. 371-375.

CARPENTER, C.S., and STEHR, M. "The effects of mandatory seatbelt laws on seatbelt use, motor vehicle fatalities, and crash-related injuries among youths." Journal of Health Economics 27.3 (2008): 642-662.

CHEN, L.H., BAKER, S.P. and LI, G.H., 2006. Graduated driver licensing programs and fatal crashes of 16year-old drivers: A national evaluation.

CLARKE, D.D., WARD, P., BARTLE, C. and TRUMAN, W., 2006. Young driver accidents in the UK: The influence of age, experience, and time of day. Accident Analysis and Prevention, 38, pp. 871-878.

COHEN, A., and EINAV, L. "The effects of mandatory seat belt laws on driving behavior and traffic fatalities." Review of Economics and Statistics 85.4 (2003): 828-843.

COUNCIL, F.M. and STEWART, J.R., 1999. Safety associations of the conversion of rural two-lane to four-lane roadways based on cross-sectional models. Transportation Research Record, (1665), pp. 35-43.

DAVIS, G. A. (2004). Possible aggregation biases in road safety research and a mechanism approach to accident modeling. Accident Analysis & Prevention, 36(6), 1119-1127.

DEERY, H.A., 1999. Hazard and Risk Perception among Young Novice Drivers. Journal of Safety Research, 30(4), pp. 225-236.

DERRIG, R. A., SEGUI-GOMEZ, M., ABTAHI, A., & LIU, L. L. (2002). The effect of population safety belt usage rates on motor vehicle-related fatalities. Accident Analysis & Prevention, 34(1), 101-110.EVANS, L. "The dominant role of driver behavior in traffic safety." American Journal of Public Health 86.6 (1996): 784-786.

EWING, R. and DUMBAUGH, E., 2009. The Built Environment and Traffic Safety A Review of Empirical Evidence.

EWING, R., SCHIEBER, R.A. and ZEGEER, C.V., 2003. Urban sprawl as a risk factor in motor vehicle occupant and pedestrian fatalities. American Journal of Public Health, 93(9), pp. 1541-1545 5p.

FEDERAL HIGHWAY ADMINISTRATION, 2016-last update, Fixing America's Surface Transportation Act or "FAST Act". Available: https://www.fhwa.dot.gov/fastact/2016].

FHWA 2016. *Highway Safety Information System*. Available: http://www.hsisinfo.org/background.cfm

FLYVBJERG, B., 2001. Making Social Science Matter: Why Social Inquiry Fails and How It Can Succeed Again. Cambridge: Cambridge Univ Pr.

FLYVBJERG, B., 1988. Rationality and Power. Chicago: Univ of Chicago Press.

FOTHERINGHAM, A.S. and WONG, D.W.S., 1991. The Modifiable Areal Unit Problem in Multivariate Statistical Analysis. Environment and Planning A, 23(7), pp. 1025-1044.

FREEDMAN, D. (1997). From association to causation via regression. Advances in applied mathematics, 18(1), 59-110.

FULLER, R., 2005. Towards a general theory of driver behaviour. Accident Analysis and Prevention, 37, pp. 461-472.

GARBACZ, C. (1991). IMPACT OF THE NEW ZEALAND SEAT BELT LAW. Economic Inquiry, 29(2), 310–316. https://doi.org/10.1111/j.1465-7295.1991.tb01273.x

GARNOWSKI, M. and MANNER, H., 2011. On factors related to car accidents on German Autobahn connectors. Accident Analysis and Prevention, 43, pp. 1864-1871.

GRAHAM, D.J. and GLAISTER, S., 2003. Spatial Variation in Road Pedestrian Casualties: The Role of Urban Scale, Density and Land-use Mix. Urban Studies (Routledge), 40(8), pp. 1591.

HARVEY, A. C., & DURBIN, J. (1986). The effects of seat belt legislation on British road casualties: A case study in structural time series modelling. Journal of the Royal Statistical Society: Series A (General), 149(3), 187-210.

HAUER, E., BONNESON, J.A., COUNCIL, F., SRINIVASAN, R. and ZEGEER, C., 2012. Crash Modification Factors Foundational Issues.

HEDLUND, J., 2000. Risky business: Safety regulations, risk compensation, and individual behavior. Injury Prevention, 6(2), pp. 82-89.

KLAUER, S.G., DINGUS, T.A., NEALE, V.L., SUDWEEKS, J.D. and RAMSEY, D.J., 2006. The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data.

Krile, R., Feng, J., & Schroeder, J. (2014). Assessing Roadway Traffic Count Duration and Frequency Impacts on Annual Average Daily Traffic Estimation: Assessing Accuracy Issues with Current Known Methods in AADT Estimation from Continuous Traffic Monitoring Data. FHWA.

LABI, S., 2011. Efficacies of roadway safety improvements across functional subclasses of rural two-lane highways. Journal of Safety Research, 42, pp. 231-239.

LIPSKY, M. (2010). Street-level bureaucracy: Dilemmas of the individual in public service. Russell Sage Foundation.

LOEB, P. D. (2001). The effectiveness of seat belt legislation in reducing driver-involved injury rates in Maryland. Transportation Research Part E: Logistics and Transportation Review, 37(4), 297-310.

Lord, D., Persaud, B.N., 2000. Accident prediction models with and without trend: application of the generalized estimating equations procedure.
Transportation Research Record 1717, 102–108.

LORD, D. and MANNERING, F., 2010. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. Elsevier Ltd.

LOUKAITOU-SIDERIS, A., SUNG, H.-.(.1.). and LIGGETT, R.(.2.)., 2007. Death on the crosswalk: A study of pedestrian-automobile collisions in Los Angeles. Journal of Planning Education and Research, 26(3), pp. 338-351.

MACHADO-LEÓN, J.L., DE OÑA, J., DE OÑA, R., EBOLI, L. and MAZZULLA, G., 2016. Socio-economic and driving experience factors affecting drivers’ perceptions of traffic crash risk. Transportation Research Part F: Psychology and Behaviour, 37, pp. 41-51.

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, pp. 131-139.

MICHON, J.A., 1985. Critical View of Driver Behavior Models: what do we Know, what should we Do? , / 01 / 01 / 1985, Plenum Press (GM Symposium Series), pp. 485-524.

MILTON, J.(.1.). and MANNERING, F.(.2.)., 1998. The relationship among highway geometrics, traffic related elements and motor-vehicle accident frequencies. Transportation, 25(2-4), pp. 395-413.

MITCHELL, T., 2002. Rule of Experts: Egypt, Techno-Politics, Modernity. Berkeley: U. of California.

MITRA, S. and WASHINGTON, S., 2012. On the significance of omitted variables in intersection crash modeling. Accident Analysis and Prevention, 49, pp. 439-448.

NÄÄTÄNEN, R. and SUMMALA, H., 1974. A model for the role of motivational factors in drivers' decision-making*. Accident Analysis and Prevention, 6(3-4), pp. 243-261.

NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION, 2010. The Economic and Societal Impact of Motor Vehicle Crashes. DOT HS 812 013. NHTSA.

NOLAND, R.B., 2013. From theory to practice in road safety policy: Understanding risk versus mobility. Research in Transportation Economics, 43, pp. 71-84.

NOLAND, R.B., KLEIN, N.J. and TULACH, N.K., 2013. Do lower income areas have more pedestrian casualties? Accident Analysis and Prevention, 59, pp. 337-345.

NOLAND, R.B. and QUDDUS, M.A., 2004. A spatially disaggregate analysis of road casualties in England. Accident Analysis and Prevention, 36, pp. 973-984.

O'NEILL, B., 1977. A decision-theory model of danger compensation. Accident Analysis and Prevention, 9(3), pp. 157-165.

OPENSHAW, S., 1984. Ecological fallacies and the analysis of areal census data. Environment & Planning A, 16(1), pp. 17-31.

PELTZMAN, S., 1975. The Associations of Automobile Safety Regulation. The University of Chicago Press.

PETERSON, S., HOFFER, G., and MILLNER, E. "Are drivers of air-bag-equipped cars more aggressive? A test of the offsetting behavior hypothesis." The Journal of Law and Economics 38.2 (1995): 251-264.

PETRIDOU, E., MOUSTAKI, M. "Human factors in the causation of road traffic crashes." European journal of epidemiology 16.9 (2000): 819-826.

PIETRZAK, M.B., 2014. Redefining the Modifiable Areal Unit Problem within Spatial Econometrics: The Case of the Scale Problem. Equilibrium: Quarterly Journal of Economics and Economic Policy, 9(2), pp. 111-132.

RANNEY, T.A., 1994. Models of driving behavior: A review of their evolution. Accident Analysis and Prevention, 26(6), pp. 733-750.

SABEY, B.E., STAUGHTON, G.C. "Interacting roles of road environment vehicle and road user in accidents." Ceste I Mostovi (1975).

SEN, A. "An empirical test of the offset hypothesis." The Journal of Law and Economics 44.2 (2001): 481-510.

SAWALHA, Z. and SAYED, T., 2001. Evaluating Safety of Urban Arterial Roadways. Journal of Transportation Engineering, 127(2), pp. 151.

Shankar, V.N., Albin, R.B., Milton, J.C., Mannering, F.L., 1998. Evaluating median cross-over likelihoods with clustered accident counts: an empirical inquiry using random effects negative binomial model. Transportation Research Record 1635, 44–48.

SHOUP, D.C. 2005, *The High Cost of Free Parking,* Chicago. American Planning Association, Planners Press.

Sittikariya, S., Shankar, V., 2009. Modeling Heterogeneity: Traffic Accidents, vol. 80. VDM-Verlag.

SMEED, R.J., 1949. Some Statistical Aspects of Road Safety Research. Royal Statistical Society.

SUMMALA, H., 1988. Risk control is not risk adjustment: The zero-risk theory of driver behaviour and its implications. Ergonomics, 31(4), pp. 491-506.

TAYLOR, D.H., 1964. Drivers' Galvanic Skin Response and the Risk of Accident. Ergonomics, 7(4), pp. 439.

TREAT, J.R. "A study of precrash factors involved in traffic accidents." HSRI Research Review (1980).

WASHINGTON, S.P., KARLAFTIS, M.G. and MANNERING, F.L., 2011. Statistical and econometric methods for transportation data analysis. CRC Press, Boca Raton, FL.

WILDE, G.J.S., 1982. The Theory of Risk Homeostasis: Implications for Safety and Health. Risk Analysis: An International Journal, 2(4), pp. 209.

YANNIS, G., KANELLOPOULOU, A., AGGELOUSSI, K. and TSAMBOULAS, D., 2005. Modelling driver choices towards accident risk reduction. Safety Science, 43, pp. 173-186.

Washington, S.P., Karlaftis, M.G., Mannering, F.L., 2010. Statistical and Econometric Methods for Transportation Data Analysis, second ed. Chapman Hall/ CRC, Boca Raton, FL.

ZENG, Q. and HUANG, H., 2014. Bayesian spatial joint modeling of traffic crashes on an urban road network. Accident Analysis and Prevention, 67, pp. 105-112.

ZHOU, M.(.1.). and SISIOPIKU, V.P.(.2.)., 1997. Relationship between volume-to-capacity ratios and accident rate