Feasibility of Forecasting Highway Safety in Support of Safety Incentive and Safety Target Programs

Final Report 597

Prepared by:
Simon Washington, Ph.D.
and
Do-Gyeong Kim
Department of Civil and Environmental Engineering
Arizona State University
Tempe, Arizona

November 2007

Prepared for:
Arizona Department of Transportation
206 South 17th Avenue
Phoenix, Arizona 85007
and
Arizona Governor's Office of Highway Safety
In cooperation with
U.S. Department of Transportation
Federal Highway Administration
National Highway Traffic Safety Administration

Technical Report Documentation Page

<table>
<tr><td>1. Report No.:
FHWA-AZ-06-597</td><td>2. Government Accession No.</td><td>3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
Feasibility of Forecasting Highway In Support of Safety Incentive and Safety Target Programs</td><td>5. Report Date:
November 2007</td></tr>
<tr><td>6. Performing Organization Code</td></tr>
<tr><td colspan="2">7. Authors: Drs. Simon Washington and Do-Gyeong Kim</td><td>8. Performing Organization Report No.</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address
Arizona State University
Department of Civil & Environmental Engineering
Tempe, AZ 85287-5306</td><td>10. Work Unit No.</td></tr>
<tr><td>11. Contract or Grant No.
SPR-PL-1-(61) 597</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address
Arizona Department of Transportation
206 S. 17th Avenue
Phoenix, AZ 85007</td><td>13.Type of Report & Period Covered
Final Report</td></tr>
<tr><td>14. Sponsoring Agency Code</td></tr>
<tr><td colspan="3">15. Supplementary Notes:
Prepared in cooperation with the Arizona Governor's Office of Highway Safety and the U.S. Department of Transportation, Federal Highway Administration & National Highway Traffic Safety Administration</td></tr>
<tr><td colspan="3">16. Abstract
Using the frequency of fatal crashes from the current observation period (e.g. month, year, etc.) as the prediction of expected future performance does not account for changes in safety that result from increases in exposure (population, additional roads, new drivers, etc.). This effect is especially pronounced in rapidly growing regions, where safety changes are anticipated in the absence of safety investment programs. The objective of this study was to examine the feasibility of predicting future fatal motor vehicle crashes given changes in future risk exposure, so that reasonable safety 'targets' can be established in support of a safety incentive or safety target programs. Safety incentive or target programs can be used to set future safety targets (i.e. fatal crashes) for jurisdictions in Arizona.</td></tr>
</table>

<table>
<tr><td colspan="2">17. Key Words
Safety forecasting, fatal crashes, safety planning, safety management</td><td colspan="2">18. Distribution Statement
Document is available to the U.S. public through the National Technical Information Service, Springfield, Virginia 22161</td><td rowspan="2">23. Registrant's Seal</td></tr>
<tr><td>19. Security Classification
Unclassified</td><td>20. Security Classification
Unclassified</td><td>21. No. of Pages
41</td><td>22. Price</td></tr>
</table>

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS					APPROXIMATE CONVERSIONS FROM SI UNITS				
Symbol	When You Know	Multiply By	To Find	Symbol	Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH					**LENGTH**		
in	inches	25.4	millimeters	mm	mm	millimeters	0.039	inches	in
ft	feet	0.305	meters	m	m	meters	3.28	feet	ft
yd	yards	0.914	meters	m	m	meters	1.09	yards	yd
mi	miles	1.61	kilometers	km	km	kilometers	0.621	miles	mi
		AREA					**AREA**		
in^2	square inches	645.2	square millimeters	mm^2	mm^2	Square millimeters	0.0016	square inches	in^2
ft^2	square feet	0.093	square meters	m^2	m^2	Square meters	10.764	square feet	ft^2
yd^2	square yards	0.836	square meters	m^2	m^2	Square meters	1.195	square yards	yd^2
ac	acres	0.405	hectares	ha	ha	hectares	2.47	acres	ac
mi^2	square miles	2.59	square kilometers	km^2	km^2	Square kilometers	0.386	square miles	mi^2
		VOLUME					**VOLUME**		
fl oz	fluid ounces	29.57	milliliters	mL	mL	milliliters	0.034	fluid ounces	fl oz
gal	gallons	3.785	liters	L	L	liters	0.264	gallons	gal
ft^3	cubic feet	0.028	cubic meters	m^3	m^3	Cubic meters	35.315	cubic feet	ft^3
yd^3	cubic yards	0.765	cubic meters	m^3	m^3	Cubic meters	1.308	cubic yards	yd^3
	NOTE: Volumes greater than 1000L shall be shown in m^3.								
		MASS					**MASS**		
oz	ounces	28.35	grams	g	g	grams	0.035	ounces	oz
lb	pounds	0.454	kilograms	kg	kg	kilograms	2.205	pounds	lb
T	short tons (2000lb)	0.907	megagrams (or “metric ton”)	mg (or “t”)	mg	megagrams (or “metric ton”)	1.102	short tons (2000lb)	T
		TEMPERATURE (exact)					**TEMPERATURE (exact)**		
°F	Fahrenheit temperature	5(F-32)/9 or (F-32)/1.8	Celsius temperature	°C	°C	Celsius temperature	1.8C + 32	Fahrenheit temperature	°F
		ILLUMINATION					**ILLUMINATION**		
fc	foot candles	10.76	lux	lx	lx	lux	0.0929	foot-candles	fc
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2	cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE AND PRESSURE OR STRESS					**FORCE AND PRESSURE OR STRESS**		
lbf	poundforce	4.45	newtons	N	N	newtons	0.225	poundforce	lbf
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa	kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380

Table of Contents

List of Tables

List of Figures

List of Acronyms

ADOT	Arizona Department of Transportation
AADT	Annual Average Daily Traffic
DOT	Department of Transportation
GHSA	Governors’ Highway Safety Association
MPO	Metropolitan Planning Organization
PPHH	Persons Per Household
TAZ	Traffic Analysis Zone
VMT	Vehicle Miles Traveled

Motivation for Study

Approximately 43,000 people die on the nation's roads each year. In addition, motor vehicle crashes are the leading cause of death or injury for persons from age 2 through 33. These traffic fatality and injury statistics have led to significant interest in highway safety investments that will save lives.

One relatively unexplored area of research is the setting of safety targets, incentives, or milestones for jurisdictions. For example, a region may want to achieve a measurable decrease in pedestrian-involved fatal crashes in a future time period. Using simply the baseline (e.g., the current year's) crash frequencies (e.g., fatal crashes, injury crashes, etc.) to set performance targets is inadequate, especially in rapidly developing states such as Arizona, since the impacts of growth alone will affect the expected safety of a region or jurisdiction. Specifically, increased vehicle miles traveled (VMT), increased population, new drivers, and new facilities will lead to an increase in expected crashes. Under these circumstances, the current frequency of crashes is not a reliable estimate of future expected crashes.

Crash prediction models based on statistical or econometric modeling techniques have been developed for a variety of purposes; most commonly to estimate the expected crash frequencies from various roadway entities (highways, intersections, interstates, etc.) and also to identify geometric, environmental, and operations factors that are associated with crashes. The vast majority of these models have been developed to forecast crashes on roadway links and intersections—a scale that is too fine and too cumbersome for forecasting crashes at the jurisdiction or regional scale.

Agencies, such as departments of transportation (DOTs) and the Governors' Highway Safety Association (GHSA), may benefit from tools that enable the setting of future safety targets. These targets may be used to support incentive-based programs within a jurisdiction—offering incentives based on how many motor-vehicle–related fatalities (and/or injuries) are reduced by a region's safety investment program.

To support an incentive based program, it is necessary to be able to accurately forecast what safety targets are expected to be in a region or jurisdiction in a future time period, given changes in road mileage, population, land area, VMT, and so on. This can be accomplished by estimating crash prediction models using aggregate jurisdiction-level characteristics as predictors. Of course a defensible and comprehensive set of predictors must be collected in order to produce a reliable and precise forecasting model.

The purpose of this research is to examine the feasibility of developing a safety forecasting model in Arizona to support both or either a safety incentive or a safety target program. A safety incentive program in concept would offer incentives (e.g., project funds) for jurisdictions able to show reductions in crashes (e.g., fatal) due to implementation of effective safety programs. It is not known whether safety incentive programs would be adopted or are attractive to jurisdictions or to Metropolitan Planning Organizations (MPOs)—but the concept has been raised and discussed in professional forums. It is known, however, that in order to set reasonable safety goals—such as goals established in a statewide safety management plan—one must be able to forecast the expected total crashes and fatal crashes within jurisdictions given the planned growth in population, road mileage, etc. expected over various growth time horizons. Thus, the need to forecast safety targets has largely been overlooked in Arizona and the U.S. as a whole.

Motor vehicle crash data along with exposure-related safety data obtained from the U.S. Census, ADOT, and through questionnaires sent to Arizona jurisdictions were used to develop the forecasting models and methodology for this project. The models are intended to enable transportation safety practitioners and decision makers to better understand the relationship between fatalities and exposure-related safety variables, especially in jurisdictions with rapid growth, so that future safety targets can be set.

The remainder of this report provides additional background and review of previous research regarding jurisdiction-level crash forecasting models. The data collected in support of this research effort are then described. A description of the modeling approach used in this research is followed by a discussion of modeling results. The procedure to apply the safety incentive or safety target forecasting model is then given, with two example applications. Finally, conclusions and recommendations are provided.

Background and Relevant Research

The safety profession is replete with models that predict crashes at the microscopic level—say for intersections or for road segments (see Bibliography for extensive examples). These models, however, do not address the forecasting of crashes at more aggregate levels. Very little research has addressed this issue, and much work is ongoing in this area of research (Washington et al. 2006). Other research efforts on aggregate level forecasting are currently underway at Ryerson Polytechnic University, Purdue University, and the University of British Columbia. Despite the little research that has been conducted in this area, much has been learned from prior research on microscopic crash models.

A reasonable question to ask is: "Are macroscopic, or jurisdiction-level, statistical models defensible and logically feasible?" The following arguments, based on accepted principles and logic from the road safety and statistics communities, support the use of aggregate level safety prediction models.

1. Crashes are largely random events. Much research has shown that human errors account for 60% to 90% of crashes. Thus, many crashes are more a function of human-related factors rather than roadway-related factors. As simple examples, crashes that result from a driver tuning a radio, answering a cell phone, following another vehicle too closely, speeding, or running a red light are events that occur somewhat randomly on a network. It is easy to understand, then, that modeling crashes at the segment or intersection level is challenging, because there is a large random component to crashes that is not explained by local road characteristics. At a more aggregate level, in contrast, crashes are related to aggregate predictors, such as population demographics, 'high risk' driving populations, the general classes of road facilities, etc., and assigning crashes to specific links or segments is not necessary. Thus, by aggregating the transportation system at the Traffic Analysis Zone(TAZ)[1] level or higher , some of the difficulties caused by 'lumpiness' of random events that we see across intersections or across road segments are reduced.

2. Aggregate safety differences are substantiated by research. Much research supports that safety is related to aggregate measures of exposure. First order effects are revealed as more VMT and crashes, and population and crashes are strongly and positively correlated. Older drivers suffer from reduced reaction and perception times, as well as reduced vision and flexibility. Younger drivers suffer from inexperience and aggressiveness. Minorities have been shown to wear safety restraints less than whites, and restraint use in rural areas is less than in urban areas. Interstates have relatively low crash rates, while rural roads with high speeds have more serious-injury crashes. Crashes in urban areas are attended by emergency medical services more quickly than crashes in rural areas. Intersections are locations of complex traffic movements and thus have greater

[1] TAZ is the unit of analysis used in metropolitan planning level travel demand models.

numbers of crashes than road segments. Increasing traffic congestion tends to reduce crash severity. School zones are associated with bicycle and pedestrian crashes. These well supported aggregate relationships, and others not listed here, are the relationships captured in aggregate level prediction models. The aggregate relationships described above form the basis for the statistical modeling at the TAZ level. It is the reliance on these 'average' relationships, and characteristics measured at the TAZ level, on which model predictions are based.

3. Models for predicting have fewer restrictions than models for explaining. Intersection and road-segment level accident prediction models are usually held to a high standard, as they are often used both to predict the expected performance of such facilities but also to explain relationships between variables. Often, and sometimes wrongly, these microscopic models are used to infer the effects of countermeasures, such as the safety effect of the presence of a left-turn lane on angle crashes. When a model is used simply for prediction, however, and not inference, there is greater flexibility in model estimation and variable selection choices. The PLANSAFE model is intended only for prediction, not explanation. (See Washington et al. 2006 for a discussion of the PLANSAFE model used for forecasting safety.) Thus, for example, if a population variable is used to predict fatal crashes per TAZ, its estimated coefficient is used solely in the prediction equation but is not interpreted to have specific explanatory marginal effects.

These three arguments, or justifications, form the basis for the development of jurisdiction-level accident prediction models. A consequence of these arguments, however, is that the models cannot be used for explanation of crash causation or for the assessment of roadway-specific countermeasures. The aggregate relationships modeled are suitable for predicting a hypothetical or future outcome should the set of predictors be changed. This restriction is not too dissimilar from the restriction placed on travel demand models, whose primary purpose is to predict demand for roadway space of motor vehicles in hypothetical or future scenarios.

We must also recognize that aggregation reduces the variability and can lead to ecological correlation, which can lead to interpretation problems. Thus, we must proceed carefully to examine the forecasting models, to interpret them with caution, and to apply them as they are intended—to predict future trends in safety given aggregate changes in exposure.

Data Collection and Description

Estimating jurisdiction-level crash prediction models to predict safety requires aggregate information such as socio-economic, demographic, and transportation related data. In particular, data related to established risk exposure variables are needed.

The analysis unit of this research is the jurisdiction (cities and towns), since the objective of this research is to examine how jurisdictional characteristics influence safety. Arizona consisted of 87 jurisdictions within 15 counties (as of 2005), which are divided into the six regional Councils of Governments (COG) for multi-jurisdictional regional planning as shown in Figure 1: Central Arizona Association of Governments (CAAG), Maricopa Association of Governments (MAG), Northern Arizona Council of Governments (NACOG), Pima Association of Governments (PAG), South Eastern Arizona Governments Organization (SEAGO), and Western Arizona Council of Governments (WACOG). Of course these COG boundaries change over time and currently new COGs have been formed within the state.

Figure 1: Location of Counties and COGs/MPOs in Arizona

As mentioned previously, the required input data for the analysis are aggregated by jurisdiction. The aggregate data came from three sources: crash data, census data, and through mail surveys. Crash data were collected from the 2000 Arizona Crash Facts which was published by the Arizona Department of Transportation (ADOT 2001). Jurisdiction-level characteristics used for predictors were obtained from Census 2000 data, which includes a variety of socio-demographic information related to people, business, and geography that is maintained by the U.S. Census Bureau. Since some of the census data were available only for the year 2000, the number of fatalities that occurred in 2000 was used in the analysis. In addition, the survey response rate was quite unsatisfactory (17 out of 87 surveys were returned, or about 20% response rate, some of which were incomplete), and so additional variables thought to be important for safety forecasting were dropped from the analysis. Thus, the analysis is based on results from 2000 crash and census data. The questionnaire sent to jurisdictions is shown in Appendix A.

Jurisdiction-Level Safety Predictors

Although numerous jurisdiction-level safety predictors can be obtained from the U.S. Census Bureau 2000 data, this research employed nine characteristics for the analysis as predictors: population change, population density, the percentage of elderly people, the percentage of young people, the proportion of minorities, the number of dwelling units per acre, persons per household, number of employees, and mean travel time to work. Table 1 shows the abbreviated names of the variables and their units of measurement. In the following section, the characteristics of these nine variables are described in relation to safety.

VARIABLE	DESCRIPTION
PERCHAN	The percent change in population from 1990 to 2000
POPDEN	Population density (population/square mile)
POPELDER	Persons aged 65 years old or more as a percentage of the total population
POPYOUNG	Persons aged 17 years old or less as a percentage of the total population
POPMINOR	Total number of minorities as a portion of the total population
HUDEN	Number of housing units per square mile
PPHH	Persons per household
EMPLOY	The number of employees as a percentage of the total population
MTT	Mean travel time to work (minutes), workers age 16+

Table 1: Description and Measurement of Variables Used in the Analysis

Population Change

Crashes are likely to be affected by growth because a jurisdiction is coping with its growth of infrastructure and population above and beyond what is captured by population alone. For example, a rapidly growing city is more likely to have new construction projects, new housing projects, new drivers, business growth, and improvements to the transportation infrastructure (upgrading of segments and intersections). These attributes are more likely to be associated with additional crashes compared to a city with no population growth. The percent change in population was used to capture this effect.

The population of Arizona in 1990 was 3,665,339, while the population in 2000 was 5,130,632, an increase of 40.0% over ten years. Maricopa County had the largest population in Arizona, but with respect to the percent population change, Mohave County had the highest as shown in Table 2. In contrast, Greenlee County had the smallest change in population—a modest 6.7% increase.

County	Population In 1990	Population In 2000	Change (%)
Mohave County	93,497	155,032	65.8
La Paz County	13,844	19,715	42.4
Yuma County	106,895	160,026	49.7
Pima County	666,957	843,746	26.5
Apache County	61,591	69,423	12.7
Coconino County	96,591	116,320	20.4
Navajo County	77,674	97,470	25.5
Yavapai County	107,714	167,517	55.5
Maricopa County	2,122,101	3,072,149	44.8
Gila County	40,216	51,335	27.6
Pinal County	116,397	179,727	54.4
Cochise County	97,624	117,755	20.6
Graham County	26,554	33,489	26.1
Greenlee County	8,008	8,547	6.7
Santa Cruz County	29,676	38,381	29.3
TOTAL	3,665,339	5,130,632	40.0

Table 2: The Percent Change in Arizona Population from 1990 to 2000 by County

As shown in Figure 2, more than 75% of the people in Arizona in 2000 lived in Pima and Maricopa counties (16.45% and 59.88%, respectively). In contrast, 0.17% and 0.38% of population lived in Greenlee and La Paz counties, respectively. With respect to the population by jurisdiction, approximately 35% of Arizona residents lived in the cities of Phoenix and Tucson. Phoenix has the highest population, while Jerome has the smallest population. The population and percent change by jurisdiction is shown in Appendix B.

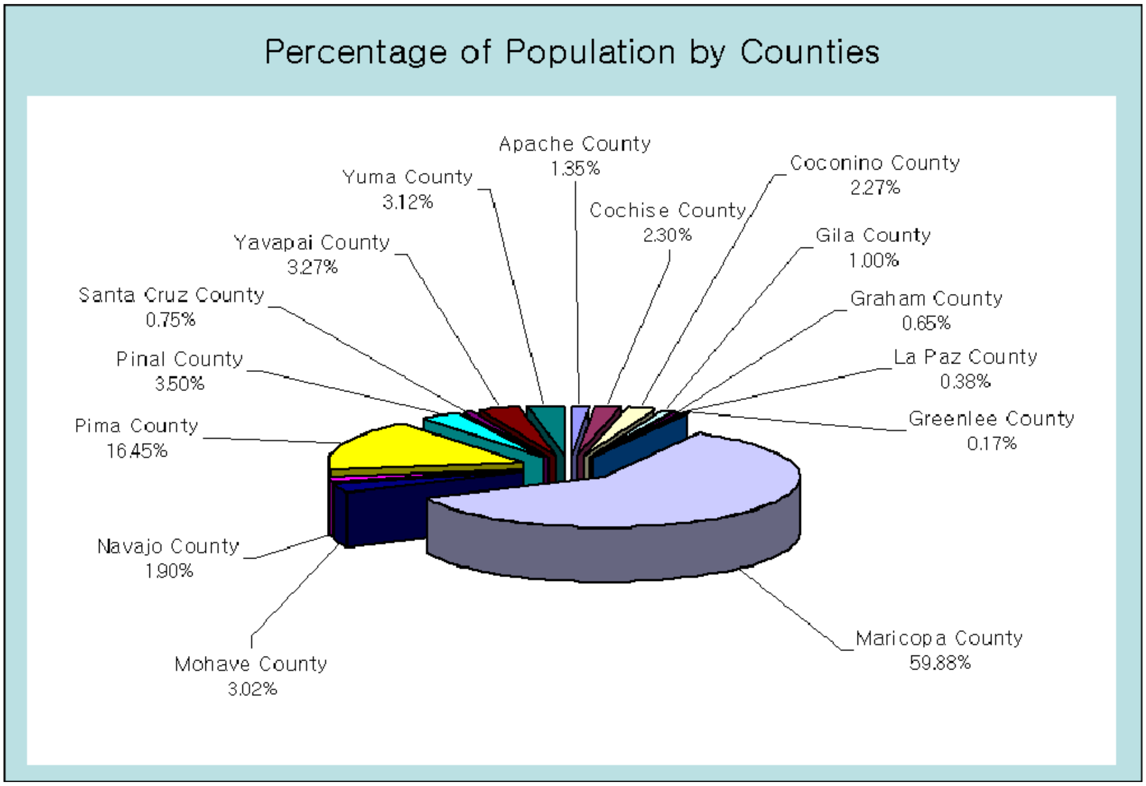

Figure 2: Percentage of Total Arizona Population by County in 2000

Population Density

It is reasonable to believe that crashes are more likely to occur in urban areas rather than rural areas because urban areas have higher accident exposures with regard to Annual Average Daily Traffic (AADT) and VMT than rural areas. However, when fatal crashes are examined the opposite relationship is expected. Rural areas are associated with higher speeds and generally non-median roadways (i.e., highways vs. freeways), and thus crashes tend to be more severe on these facilities. Also, congestion in urban areas tends to limit speeds, which in turn reduces crash severities. Thus, degree of urbanization should be negatively associated with fatal crashes and positively associated with total crashes. The population density of a jurisdiction is calculated by dividing the population by land area of the jurisdiction.

Arizona covers 114,006 square miles (113,642 square miles of land areas and 364 square miles of water areas), making it the sixth largest of the 50 states. Despite the vast area, Arizona has a relatively small number of counties (15). Coconino is the largest county (18,617.4 square miles); Santa Cruz is the smallest (1,237.6 square miles). Table 3 (p.8) shows the counties' land areas. .

Councils of Governments	County	Land Area
WACOG	Mohave County	13,311.6
	La Paz County	4,500.0
	Yuma County	5,514.1
PAG	Pima County	9,186.3
NACOG	Apache County	11,204.9
	Coconino County	18,617.4
	Navajo County	9,953.2
	Yavapai County	8,123.3
MAG	Maricopa County	9,203.1
CAAG	Gila County	4,767.7
	Pinal County	5,369.6
SEAGO	Cochise County	6,169.4
	Graham County	4,629.3
	Greenlee County	1,847.0
	Santa Cruz County	1,237.6
TOTAL		113,634.6

Table 3: Arizona Land Areas by County in 2000

The average population density of Arizona in 2000 was 45.2 persons per square mile. Maricopa County had the highest of the 15 counties (333.8/mi^2), whereas La Paz County had the lowest (4.4/mi^2) Table 4 shows the population density of all the counties. In terms of population density by city, Guadalupe had the highest density (6,813.9), while Buckeye had the lowest density (44.8). Population densities by jurisdiction—the unit of analysis used for modeling—are summarized in Appendix C.

County	Land Area	Population	Population Density
Mohave County	13,311.6	155,032	11.6
La Paz County	4,500.0	19,715	4.4
Yuma County	5,514.1	160,026	29.0
Pima County	9,186.3	843,746	91.8
Apache County	11,204.9	69,423	6.2
Coconino County	18,617.4	116,320	6.2
Navajo County	9,953.2	97,470	9.8
Yavapai County	8,123.3	167,517	20.6
Maricopa County	9,203.1	3,072,149	333.8
Gila County	4,767.7	51,335	10.8
Pinal County	5,369.6	179,727	33.5
Cochise County	6,169.4	117,755	19.1
Graham County	4,629.3	33,489	7.2
Greenlee County	1,847.0	8,547	4.6
Santa Cruz County	1,237.6	38,381	31.0
TOTAL	113,634.6	5,130,632	45.2

Table 4: Arizona Population Densities by County in 2000

The Percentage of Elderly and Young Populations

The percentages of elderly and young people are factors that are related to crashes, as mentioned previously. Young drivers are generally inexperienced, while elderly drivers suffer from reduced perception and reaction times as well as crash survivability. A total of 667,839 people, about 13.0% of the population, were 65 years old or more in Arizona in 2000. A total of 1,150,466 people, about 22.4% of the population, were 17 years old or less in 2000. Of the 15 counties, Maricopa County had the highest number of both elderly and young people, and with respect to jurisdictions, Phoenix had the highest number of both elderly and young people.

Besides these age-related factors, this research employs the proportion of minorities as a predictor of severe crashes. Minorities, on average, wear safety restraints less than whites (see White et al. 2001 and Washington et al. 1999) and tend to drive older vehicles with less extensive safety features (compared to new vehicles). Phoenix was found to have the highest number of minorities. Table 5 shows the number of elderly, young, and minorities by county.

County	Elderly People	Young People	Minority
Mohave County	20,801	19,063	15,786
La Paz County	2,139	977	1,967
Yuma County	12,388	26,681	62,785
Pima County	66,911	111,925	236,186
Apache County	1,002	2,555	2,488
Coconino County	3,689	13,239	19,779
Navajo County	3,798	8,791	11,318
Yavapai County	22,056	16,699	14,401
Maricopa County	276,354	672,338	1,000,267
Gila County	5,664	4,657	6,289
Pinal County	16,017	19,974	42,397
Cochise County	10,191	15,538	31,654
Graham County	2,327	3,740	5,410
Greenlee County	371	922	1,824
Santa Cruz County	2,448	6,159	20,105
TOTAL	667,839	1,150,466	1,856,374

Table 5: Arizona Population of Elderly, Young, and Minorities by County in 2000

Residential Dwelling Density and Persons per Household

Residential dwelling density (the number of housing units per square mile) and persons per household were also used as predictors to model jurisdiction-level crash prediction models (see Table 6). Residential dwelling density is a surrogate for the compactness of development—which corresponds to high intersection intensity (the questionnaire in Appendix A tried to capture this more explicitly). Persons per household might reflect socio-economic status, which may correlate with vehicle age and safety equipment, employment, etc. The total number of housing units over 87 jurisdictions within 15 counties was 1,660,557. Of the 15 counties, Maricopa County and La Paz County had the highest and smallest house densities, respectively. With respect to residential density by jurisdiction, South Tucson was the highest (2,059.0) whereas Buckeye was the lowest (16.1).

The number of persons per household (PPHH) for the 87 Arizona jurisdictions ranged from 1.74 (Youngtown) to 7.51 (Colorado City), while the average number of persons per household was 2.79. Unlike residential density, Colorado City was found to have the highest and Youngtown the lowest persons per household respectively.

County	Number of Housing Units	Housing Density	Average Number of PPHH
Mohave County	50,509	1,273.4	2.45
La Paz County	4,343	140.4	2.32
Yuma County	40,911	2,419.7	2.86
Pima County	232,563	3,734.6	2.47
Apache County	4,001	440.4	3.41
Coconino County	25,661	582.6	2.80
Navajo County	14,768	874.8	3.17
Yavapai County	45,725	2,267.8	2.33
Maricopa County	1,139,705	14,507.4	2.67
Gila County	11,663	2,001.0	2.50
Pinal County	46,014	3,682.0	2.68
Cochise County	30,344	2,272.2	2.55
Graham County	5,880	1,089.0	2.99
Greenlee County	1,471	220.6	2.73
Santa Cruz County	6,999	727.5	3.23
TOTAL	1,660,557	36,233.4	2.79

Table 6: Arizona Number of Housing Units, Density, and Persons per Household by County in 2000

Number of Employees and Mean Travel Time to Work

Relatively large numbers of employees lead to an increase in the number of trips, which increases traffic volume, and the increased trips and traffic volume may also increase exposure to the risk of motor-vehicle-related crashes. Also, work trips tend to include aggressive and/or distracted drivers. Consequently, the higher number of employees might be highly correlated with the higher crash frequencies. Due to the potential correlation between the number of employees and crash frequencies, the number of employees was included as an independent variable.

Similarly, crash frequencies might be expected to increase as mean travel time to work grows because the longer travel times suggest higher exposure levels and greater levels of driver fatigue. The mean travel time to work in Arizona is 22.1 minutes, as shown in Table 7.

County	Number of Employees	Mean Travel Time to Work (minutes)
Mohave County	78,758	20.6
La Paz County	5,453	17.2
Yuma County	73,596	18.6
Pima County	420,263	23.9
Apache County	6,600	28.0
Coconino County	49,627	19.0
Navajo County	23,894	22.9
Yavapai County	80,491	22.5
Maricopa County	2,148,360	26.1
Gila County	19,535	20.3
Pinal County	77,672	27.4
Cochise County	53,114	19.8
Graham County	11,276	22.6
Greenlee County	2,421	20.3
Santa Cruz County	14,982	19.7
TOTAL	3,006,042	22.1

Table 7: Arizona Number of Employees and Mean Travel Time to Work by County in 2000

Jurisdiction-Level Crash Data

As mentioned previously, crash data used in this research were obtained from the 2000 Arizona Crash Facts. In that year, 131,368 crashes and 891 fatal crashes occurred in Arizona. Excluding crashes that occurred within a specific county but not within a specific city (e.g., state rural roads in Apache County experienced 358 total crashes and 10 fatal crashes in 2000), 108,176 total crashes and 392 fatal crashes occurred in the 87 Arizona jurisdictions (cities or towns). Of the 87 jurisdictions, the City of Phoenix had the greatest number of total crashes and fatal crashes, with 44,146 crashes and 168 fatal crashes reported. Zero crashes were reported in eight jurisdictions, whereas a total of 47 jurisdictions reported zero fatal crashes in 2000. Table 8 presents the frequency of total crashes and fatalities across the 87 Arizona jurisdictions in 2000 (excluding crashes occurring on state rural roads or other rural roads outside a jurisdiction but within the county).

As shown in table 8, crash frequencies by jurisdiction for fatal crashes are quite small relative to total crashes. As a result, statistical models based on total crashes will tend to be more reliable than those on fatal crash data. Also of note is the preponderance of zeroes reported in the fatal crash column. A large number of zeroes is a common phenomenon with fatal crash data across analysis units. These zeroes tend to raise challenges in the estimation of statistical models, as described by Lord et al. (2004). In the analysis that follows we generally follow the advice of Lord et al. (2004) in dealing with the 'excess' zeroes.

County	Jurisdiction	Total Crashes	Fatal Crashes
Cochise County	Benson	36	0
	Bisbee	48	1
	Douglas	321	1
	Huachuca City	7	1
	Sierra Vista	775	2
	Tombstone	19	0
	Willcox	39	0
Graham County	Pima	14	0
	Safford	113	0
	Thatcher	48	0
Greenlee County	Clifton	30	0
	Duncan	0	0
Santa Cruz County	Nogales	411	1
	Patagonia	0	0
TOTAL		1,861	6
Mohave County	Bullhead City	673	4
	Colorado City	17	0
	Kingman	400	1
	Lake Havasu City	525	6
La Paz County	Parker	16	0
	Quartzsite	38	0
Yuma County	San Luis	0	0
	Somerton	16	0
	Wellton	1	0
	Yuma	1,489	3
TOTAL		3,175	14

Table 8 Arizona Total and Fatal Crashes by Jurisdiction in 2000

County	Jurisdiction	Total Crashes	Fatal Crashes
Pima County	Marana	545	1
	Oro Valley	263	2
	Sahuarita	40	0
	South Tucson	181	0
	Tucson	14,822	52
TOTAL		15,851	55
Apache County	Eagar	44	0
	Saint Johns	25	0
	Springerville	1	0
Coconino County	Flagstaff	2,480	7
	Fredonia	3	0
	Page	123	2
	Williams	53	2
Navajo County	Holbrook	83	1
	Pinetop-Lakeside	96	0
	Show Low	133	0
	Snowflake	80	1
	Taylor	6	0
	Winslow	162	1
Yavapai County	Camp Verde	89	0
	Chino Valley	83	3
	Clarkdale	0	0
	Cottonwood	196	1
	Jerome	10	0
	Prescott	883	2
	Prescott Valley	357	2
	Sedona	0	0
TOTAL		4,907	22
Maricopa County	Avondale	473	0
	Buckeye	6	2
	Carefree	9	0
	Cave Creek	10	0
	Chandler	3,056	4
	El Mirage	114	3
	Fountain Hills	69	0
	Gila Bend	18	0
	Gilbert	1,352	7
	Glendale	4,997	27
	Goodyear	249	4

Table 8 continued

County	Jurisdiction	Total Crashes	Fatal Crashes
Maricopa County (continued)	Guadalupe	35	0
	Litchfield Park	0	0
	Mesa	11,019	30
	Paradise Valley	239	0
	Peoria	1,554	1
	Phoenix	44,146	168
	Queen Creek	0	0
	Scottsdale	4,555	19
	Surprise	244	3
	Tempe	8,453	16
	Tolleson	125	0
	Wickenburg	97	2
	Youngtown	0	0
TOTAL		80,820	286
Gila County	Globe	141	0
	Hayden	5	1
	Miami	24	0
	Payson	139	0
	Winkelman	2	0
Pinal County	Apache Junction	331	3
	Casa Grande	661	2
	Coolidge	82	0
	Eloy	83	2
	Florence	60	1
	Kearny	4	0
	Mammoth	9	0
	Superior	21	0
TOTAL		1,562	9

Table 8 continued

Methodological Approach and Modeling Results

Methodological Approach

Two different approaches are generally used to estimate crash predictions, Poisson regression and negative binomial regression (Jovanis and Chang 1986; Washington et al. 2003), although various other modeling approaches are possible (see for example Lord et al. 2004). Crash counts are approximated well by a Poisson process (Joshua and Garber 1990), since crash counts are discrete, positive integers. The Poisson regression model requires that the variance of the crash frequency is approximately equal to its mean. In much of the observed crash data, however, the variance of the crash frequency is greater than the mean and overdispersion occurs (Miaou et al. 1992). Miaou et al. introduced the negative binomial distribution for modeling traffic safety which accommodates greater variance in the data than allowed by the Poisson distribution. The overdispersion typically arises from variation in crash means across sites. As a result, the negative binomial regression model is the preferred modeling approach when overdispersion is present (Washington et al. 2003). This research employs the negative binomial regression model to develop jurisdiction-level crash prediction models. The negative binomial regression model is briefly summarized in the following section.

The Negative Binomial regression model specifies a relationship between the expected number of crashes occurring at the *i*th element and the *q* parameters, X_{i1}, X_{i2}, …, X_{iq}, as follows:

$$E(y_i) = \mu_i = \exp(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots\cdots + \beta_q X_{iq} + \varepsilon_i)$$

In addition, the Negative Binomial regression model includes a quadratic term in the variance to reflect overdispersion in the model variance. As a result, the Negative Binomial regression model takes the following form:

$$P(y_i) = \frac{(y_i + \alpha - 1)!}{y_i!(\alpha - 1)!} \frac{\mu^{y_i}}{(1+\mu)^{y_i+\alpha}},$$

where α is the overdispersion parameter and the variance is:

$$Var(y_i) = \mu_i + \alpha(\mu_i)^2.$$

The overdispersion or extra-Poisson variation is generally due to variables omitted from the model that explain variation in crashes between sites. If α is equal to 0, the Negative Binomial reduces to a Poisson model. The value of α corresponds with the degree of overdispersion over and above that associated with the mean μ_i.

Similar to previous studies, the final model structure used is

$$E(y_i) = \exp \sum \beta_j x_{ij}$$

where $E(y_i)$ is the expected number for jurisdiction *i*, x_{ij}'s are variables describing safety-related exposure variables for jurisdictions, and the β_j are estimated parameters or effects of the predictor variables.

Modeling Results

First, we examined the correlation between independent variables, since one of the problems with multiple regression is that explanatory variables may be correlated, thus confounding the effects of variables with one another. In particular, regression coefficients that indicate the effect of one factor may change when some other factor is added or removed from the model.

Figure 3 shows the correlation between independent variables used in the study. Using statistical tests of significant correlation, it was found that the percentage of young people (POPYOUNG) and persons per household (PPHH) are highly positively correlated with ρ =0.8545, whereas the percentage of young people (POPYOUNG) and the number of employees (EMPLOY) are highly negatively correlated with ρ =-0.972.

Correlation Matrix

	perchan	popden	popelder	popyoung	popminor	huden	pphh	emppop	mtt
perchan	1.0000								
popden	0.0610	1.0000							
popelder	0.0093	-0.1262	1.0000						
popyoung	-0.0863	0.1185	-0.7084	1.0000					
popminor	-0.1699	0.3595	-0.4187	0.4282	1.0000				
huden	-0.0940	0.7464	-0.0131	-0.0360	0.2525	1.0000			
pphh	-0.0071	0.1794	-0.5845	0.8545	0.3720	-0.0172	1.0000		
emppop	0.0986	-0.1061	0.7401	-0.9720	-0.4379	0.0433	-0.8188	1.0000	
mtt	0.4228	0.1901	-0.0255	0.0895	-0.0534	0.1674	0.1382	-0.0653	1.0000

Figure 3: Correlation Matrix of Predictor Variables

The explanatory variables described in Table 1 were used in the software program LIMDEP (copyright William Greene) to choose the significant variables for the jurisdiction-level fatal crash model. One of the disadvantages with the LIMDEP software is that the LIMDEP does not automate the process of removing insignificant variables from the model, one at a time, until only significant variables are left in the model. Thus, various models were tested to estimate the possible nature of the relationships between the independent variables and fatal crashes, beginning with a ‘full’ model with many variables. Table 9 presents the estimation results for a fatal crash model with a ‘full’ (complete) set of predictor variables.

Among the 9 variables, only five variables were found to be statistically significant at the 10% significance level (shown in bold): PERCHAN, POPDEN, POPELDER, HUDEN, and PPHH. These variables, therefore, serve as predictors of fatal crashes. Two statistically significant variables, PERCHAN and HUDEN, are positively associated with fatal crashes, while the other three variables are negatively related with fatal crashes.

Variable	Coefficient	Standard Error	*t*-statistics	*p*-value
Constant	2.4299	17.345	0.140	0.889
PERCHAN	0.0082	0.004	2.054	**0.040**
POPDEN	-0.0004	0.000	-1.885	**0.059**
POPELDER	-0.1258	0.041	-3.033	**0.002**
POPYOUNG	0.1141	0.202	0.565	0.572
POPMINOR	0.0009	0.014	0.067	0.947
HUDEN	0.0033	0.001	4.782	**0.000**
PPHH	-2.5289	1.224	-2.066	**0.039**
EMPLOY	0.0428	0.174	0.247	0.805
MTT	-0.0258	0.046	-0.564	0.573
α (overdispersion parameter)	1.9937	0.513	3.886	0.000
Number of observations	87			
Log-likelihood at zero	-385.4672			
Log-likelihood at convergence	-143.3673			

Table 9: Negative Binomial Model Estimation Results of Fatal Crashes with Complete Set of Predictors

Table 10 shows the results of three new models retaining significant predictor variables for fatal crashes by jurisdiction in Arizona. The table also compares these models to previously reported and aggregate fatal crash models estimated through other research efforts. It should be noted that the effort by Van Schalkwyk et al. in 2006 used Traffic Analysis Zones (TAZ) as the unit of analysis, and not jurisdiction. With that said, however, the explanatory power of the Van Schalkwyk model is 75% compared to about 15% for the 'best' model developed in this effort (Model 3).

Fatal Crash Models

Variable	*Existing Models*		*New Models*		
	De Guevara et al. 2004	Van Schalkwyk et al. 2006	Model 1	Model 2	Model 3
Constant		0.652	0.017	7.7549	6.49361
POPDEN	0.050782		-0.0006		-0.0004
POPMINOR	-5.18194	0.319			
INTDEN	-4.81647	-0.0924			
PERCHAN			0.005736	0.004494	0.006366
HUDEN			0.003984	0.002577	0.003154
POPELDER				-0.16036	-0.13637
PPHH				-2.30763	-1.96688
PNF_0111*		1.762			
PNF_0512**		1.389			
POP00_15***		0.000263			
Log-likelihood	-394.882	N/A	-149.847	-145.369	-143.801
Pseudo R2	N/A	0.75	0.1147	0.1412	0.1504

*PNF_0111 = Proportion (of total road mileage) of urban and rural interstates in jurisdiction; **PNF_0512 = Proportion of freeways and expressways in jurisdiction; ***POP00_15 = Population aged 0 to 15 years old in jurisdiction.

Table 10: Three New Fatal Crash Models and Comparison to Previously Estimated Models

This result is partially explained by the fact that variability is lost with increasing levels of aggregation, and thus the ability to explain variability is also lost. Jurisdictions are certainly more aggregated than Traffic Analysis Zones. The second important reason is that the questionnaire administered as part of this research aimed to capture many of the important variables that were thought to help explain variation in safety risk across jurisdictions. Because of the poor overall response rate, and incomplete surveys among those who did respond, the opportunity to capture additional explanatory variables was missed. As is the case for most surveys, response rates are directly proportional to the ease of providing the information, the motivation of the respondent, the frequency of follow-ups, and the ability for the survey team to assist when possible. Improving response rates for future surveys of this type would require additional resources to enable these critical elements.

Procedure to Apply the Safety Incentive or Safety Target Forecasting Model

Model 3 shown in Table 10 can be used to predict fatal crashes; however, the precision of this model is quite low. In other words, for any given set of predictors, there is a large amount of unexplained variability in the number of fatal crashes occurring within a jurisdiction. Regardless of the questionable precision of the model, the following procedure reveals how safety forecasting is performed with this or a similar model.

1. ***Collect variables needed to run models***: All model variables need to be collected for the jurisdiction being analyzed. The five variables needed to forecast fatal crashes include POPDEN, PERCHAN, HUDEN, POPELDER, and PPHH (see Table 1 for a description and measurement units of these predictor variables).

2. ***Generate the expected crash counts in a spreadsheet program (such as Microsoft Excel) or database management software program (such as Microsoft Access):*** The simple equation derived from the logarithmic negative binomial regression model estimation results presented in the previous section is used to calculate the expected fatal crash count (e.g., pedestrian, total, fatal, etc.) by jurisdiction in the base year and the incentive year (e.g., an incentive is given if the jurisdiction meets safety targets in two years). The model predicts the expected (mean for all jurisdictions with these predictors) count of fatal crashes expected per year.
A spreadsheet model can be simply set up in Excel that uses the negative binomial prediction equation, $E(y_i) = \exp \sum \beta_j x_{ij}$, and the estimated coefficients shown for Model 3 in Table 10.

Table 11 shows a spreadsheet developed in Microsoft Excel for Model 3 using a hypothetical jurisdiction in the Base year and after years 1 and 2. Shown in the table are the values of the predictor variables, reflecting growth and change in the predictor variables expected during a two year time horizon. For example, the 10-year population growth rate is expected to increase from 55% to 59% by year 2. Population density is expected to increase from the base year of 150 to 160 persons per square mile.

		Predictor Values		
	Coefficient	Jurisdiction A: Base	Jurisdiction A: Year 1	Jurisdiction A: Year 2
constant	6.49361	n/a	n/a	n/a
POPDEN	-0.0004	150	155	160
POPCHANGE	0.006366	55	57	59
HUDEN	0.003154	150	150	150
POPELDER	-0.13637	10	12	14
PPHH	-1.96688	2	2.05	2.1
Fatal Crash Prediction		7.09	4.95	3.45

Table 11: Example 1: Excel Prediction Spreadsheet for Fatal Crash Prediction Model

In example 1 (Table 11), the predicted number of fatal crashes is expected to decrease from 7.09 to 3.45 crashes in year 2. Thus, without any safety investments, fatal crashes in Jurisdiction A will be reduced by approximately 3.5 fatal crashes. Safety targets for this jurisdiction should be set appropriately to account for the natural reduction in fatal crashes.

Example 2, shown in Table 12, shows Jurisdiction B growing rapidly. Its elderly population is expected to decrease by 2% over the next 2 years, while the 10-year growth rate is expected to increase by 25%. In the base year 14.46 fatal crashes are expected, while in year 2, 22.27 fatal crashes are expected. Thus, approximately 7 additional fatal crashes are expected in Jurisdiction B based on growth trends alone. In this case a safety target to maintain 14 crashes in year 2 would be extremely aggressive and would require a significant safety investment.

		Predictor Values		
	Coefficient	Jurisdiction B: Base	Jurisdiction B: Year 1	Jurisdiction B: Year 2
constant	6.49361	n/a	n/a	n/a
POPDEN	-0.0004	75	75	75
POPCHANGE	0.006366	55	65	80
HUDEN	0.003154	150	150	150
POPELDER	-0.13637	5	4	3
PPHH	-1.96688	2	2	2
Fatal Crash Prediction		14.46	17.66	22.27

Table 12: Example 2: Excel Prediction Spreadsheet for Fatal Crash Prediction Model

3. ***Incorporate modeling results into incentive program:*** The modeling results in Table 11 predict that fatal crashes will decrease as a result of projected growth in the jurisdiction and without any safety investments. Thus, safety investments should be expected to improve safety above and beyond that expected from growth alone and be the result of effective safety investments. Thus, Jurisdiction A might be expected to show a reduction in fatal crashes to less than expected, say one or two crashes, while Jurisdiction B (Table 12) might be allowed to show an increase of three crashes over the base year. The structuring of any incentive program, of course, would be devised in the state and administered by the appropriate agencies and/or stakeholders.

4. ***Calibration to Local Conditions:*** With an improved model (not one with low explanatory power), a calibration procedure will be conducted to translate the expected reduction/increase in crashes to the observed counts of crashes in a jurisdiction. This calibration procedure will be developed along with improved models in the future.

Conclusions and Recommendations

The data used in this research were obtained from ADOT and the U.S. Census Bureau, and were compiled for the year 2000. A significant effort was undertaken to collect a host of additional and important explanatory variables to improve the safety prediction model (see Appendix A), by mailing a survey to all jurisdiction representatives in Arizona. Many important exposure-related variables, such as road mileage by functional class, number and type of schools, number of intersections (signalized or stop-controlled), and weather-related variables were sought via the survey and were ultimately unavailable. This set of additional variables represents important predictors of safety, and fatal crashes in particular. Unfortunately, the lack of a complete set of predictors undermined the modeling effort.

The statistical modeling results reflect the exclusion of important variables and will lead to imprecise predictions of fatal crash frequencies. As shown in Table 10, the 'best' model in this effort produced a model with an R-Square value of approximately 15% – suggesting that 15% of the variation in fatal crashes across jurisdictions in Arizona is explained by the set of predictors in the model. Initial efforts in NCHRP 8-44 (by the same authors of this report) have produced similar models with explanatory power above 75%. These models have included many of the predictors sought in the questionnaire shown in Appendix A and also shown in Table 13.

While this report provides an analytical procedure for predicting fatal crash frequencies using a predictive model (see Table 11 and Table 12), its use is not recommended due to lack of explanatory power and precision of the model. These model deficiencies lead to the following problems:

1. The variability across jurisdictions with a similar set of predictor values will be large. Thus, the predicted fatal crash frequency will represent the mean value of a highly dispersed distribution of values.
2. Problem 1 above leads to a dilemma of dealing with the dispersion not explained by the model. A Bayesian correction of sorts, for example, could be used if known important variables were not missing; however, a Bayesian correction is problematic in the absence of known important explanatory variables.
3. When an observed fatal crash count in a jurisdiction is not close to the predicted fatal crash count in the base year a remedy is not known. The difference could be due to safety deficiencies, due to omitted important predictor variables, or due to mostly random and unknown effects. The missing important predictor variables need to be minimized.
4. Problems of omitted variables—well known to the econometrics community—will bias model predictions (see Kim et al. 2006a and 2006b and Washington et al. 2003).

The recommendation is to supplement this report with NCHRP 8-44-2 findings when they become available in April of 2009. The objectives of this NCHRP project—led by Dr. Simon Washington—are nearly the same as those of this research effort except that the NCHRP project is for a national audience. The objectives of this NCRHP research, and the similarity to this ADOT project, can be seen at http://www.trb.org/TRBNet/ProjectDisplay.asp?ProjectID=919.

It is anticipated that a more complete set of predictors, currently being developed for NCHRP 8-44-2 for Pima and Maricopa counties in Arizona, will yield a more precise and reliable model for the purpose of devising safety incentives—at least for all jurisdictions in the MAG and PAG regions. Table 13 shows a sample of some of the predictor variables that will be available in project 8-44-2. Compared to the data available in this effort, the current effort does not have variables related to weather, to high-risk non-motorized populations, to speeds and design standards, or to conflicts. Again, the intent was to capture many of these variables through the questionnaire administered as part of this project, but poor response rates resulted in a genuine lack of useful data which lead to the inability to estimate precise forecasting models.

Major Contributing Factor	Potential Aggregate (TAZ level) Variables that may capture effect of Major Factor (assumes time scale is year)
Weather	Proportion of wet pavement days per year Proportion of icy pavement days per year Proportion of snow days per year Proportion of fog/reduced visibility days per year Proportion of sunny days per year
High risk driving populations	Population/number of licensed drivers Proportion of population between 16 and 24 Proportion of population over 60 Number of DUI arrests Employed/unemployed workers
High risk non motorized populations	Number of crosswalks Number of schools (elementary, middle, high, college) Percentage/mileage of sidewalks (of street mileage) Percentage/mileage of bicycle facilities
Speed, design standards of facilities, and access control	Total street mileage Proportion of local road mileage Proportion of collector road mileage Proportion of arterial road mileage Proportion of rural highway mileage (urban/rural) Proportion of interstate (urban and rural)
Conflicts	Number/proportion of signalized intersections Number/proportion of stop-controlled intersections Intersection density Total area

Table 13: Description of Important Predictor Variables for Safety Forecasting Model

It is anticipated that NCHRP 8-44-2 will produce statistical models with superior predictive ability and for numerous safety outcomes in addition to fatal crashes. For example, in addition to fatal crash models, it is anticipated that 8-44-2 will produce statistical models able to forecast total accidents, property damage accidents, incapacitating injury and fatal crashes, night-time crashes, pedestrian crashes, injury crashes, and bicycle-related crashes. Thus, an incentive or target program that sets targets across a broad range of crash outcomes, not just fatal crashes, could be supported with such models.

Finally, while the goals and objectives are quite complementary, it is important to note that the NCHRP 8-44-2 effort will not result in the ability to forecast crash outcomes outside of the MAG and PAG regions. Thus, a safety incentive or target program cannot be supported statewide, but instead only in these two major regions.

To support a statewide safety incentive or safety target program, the data needed to estimate these models would need to be made available and/or collected. Data collection efforts would be needed in jurisdictions throughout the state including tribes, counties, and townships.

Appendix A: Questionnaire Sent to Jurisdictions

Thank you for taking the time to complete the following survey. The information you provide will be used to help us develop fatality crash models to predict a fair and reasonable fatality rate for each jurisdiction. Please note that the information you provide should be based on the year of 2000.

1. Your city or town name: ______________________________

2. Road mileage information by road functional classification:

   - Total mileage of all functional classes of roads: ___________________
   - Total mileage of principal arterial interstate: ___________________
   - Total mileage of principal arterial expressway: ___________________
   - Total mileage of principal arterial: ___________________
   - Total mileage of minor arterial: ___________________
   - Total mileage of major collector rural: ___________________
   - Total mileage of minor collector rural: ___________________
   - Total mileage of urban collector: ___________________
   - Total mileage of local streets: ___________________

3. School information

   - Number of elementary schools: ___________________
   - Number of middle schools: ___________________
   - Number of high schools: ___________________

4. How many intersections and bus stops were there within your city (or town) in 2000?

   - Intersections: ___________________
   - Bus stops: ___________________

5. How many people had driver licenses in 2000? ___________________

6. How many tickets were issued in 2000? ___________________

7. How many tickets related to speeding were issued in 2000? ___________________

8. How many DUI related accidents occurred in 2000? ___________________

9. Weather information:

   - What is the annual average precipitation in 2000? ___________________ inches
   - What is the annual average snowfall in 2000? ___________________ inches

Please return this survey by November 10th by fax on 480-965-0557 (Do-Gyeong Kim) or via email at dokkang@u.arizona.edu.

Appendix B: Arizona Population and Population Change Statistics by Jurisdiction (2000)

	Jurisdictions	Population in 1990	Population in 2000	Change (%)
Mohave County	Bullhead City	21,951	33,769	53.8
	Colorado City	2,426	3,334	37.4
	Kingman	12,722	20,069	57.8
	Lake Havasu City	24,363	41,938	72.1
La Paz County	Parker	2,897	3,140	8.4
	Quartzsite	1,876	3,354	78.8
Yuma County	San Luis	4,212	15,322	263.8
	Somerton	5,282	7,266	37.6
	Wellton	1,066	1,829	71.6
	Yuma	56,966	77,515	36.1
Pima County	Marana	2,187	13,556	519.8
	Oro Valley	6,670	29,700	345.3
	Sahuarita	1,629	3,242	99.0
	South Tucson	5,171	5,490	6.2
	Tucson	405,371	486,699	20.1
Apache County	Eagar	4,025	4,033	0.2
	Saint Johns	3,294	3,269	-0.8
	Springerville	1,802	1,972	9.4
Coconino County	Flagstaff	45,857	52,894	15.3
	Fredonia	1,207	1,036	-14.2
	Page	6,598	6,809	3.2
	Williams	2,532	2,842	12.2
Navajo County	Holbrook	4,686	4,917	4.9
	Pinetop-Lakeside	2,422	3,582	47.9
	Show Low	5,020	7,695	53.3
	Snowflake	3,679	4,460	21.2
	Taylor	2,418	3,176	31.3
	Winslow	9,279	9,520	2.6
Yavapai County	Camp Verde	6,243	9,451	51.4
	Chino Valley	4,837	7,835	62.0
	Clarkdale	2,144	3,422	59.6
	Cottonwood	5,918	9,179	55.1
	Jerome	403	329	-18.4
	Prescott	26,592	33,938	27.6
	Prescott Valley	8,904	23,535	164.3
	Sedona	7,720	10,192	32.0

	Jurisdictions	Population in 1990	Population in 2000	Change (%)
Maricopa County	Avondale	16,169	35,883	121.9
	Buckeye	4,436	6,537	47.4
	Carefree	1,657	2,927	76.6
	Cave Creek	2,925	3,728	27.5
	Chandler	89,862	176,581	96.5
	El Mirage	5,001	7,609	52.1
	Fountain Hills	10,030	20,235	101.7
	Gila Bend	1,747	1,980	13.3
	Gilbert	29,122	109,697	276.7
	Glendale	147,864	218,812	48.0
	Goodyear	6,258	18,911	202.2
	Guadalupe	5,458	5,228	-4.2
	Litchfield Park	3,303	3,810	15.3
	Mesa	288,104	396,375	37.6
	Paradise Valley	11,773	13,664	16.1
	Peoria	50,675	108,364	113.8
	Phoenix	983,392	1,321,045	34.3
	Queen Creek	2,667	4,316	61.8
	Scottsdale	130,075	202,705	55.8
	Surprise	7,122	30,848	333.1
	Tempe	141,993	158,625	11.7
	Tolleson	4,434	4,974	12.2
	Wickenburg	4,515	5,082	12.6
	Youngtown	2,542	3,010	18.4
Gila County	Globe	6,062	7,486	23.5
	Hayden	909	892	-1.9
	Miami	2,018	1,936	-4.1
	Payson	8,377	13,620	62.6
	Winkelman	676	443	-34.5
Pinal County	Apache Junction	18,092	31,814	75.8
	Casa Grande	19,076	25,224	32.2
	Coolidge	6,934	7,786	12.3
	Eloy	7,211	10,375	43.9
	Florence	7,321	17,054	132.9
	Kearny	2,262	2,249	-0.6
	Mammoth	1,845	1,762	-4.5
	Superior	3,468	3,254	-6.2

	Jurisdictions	Population in 1990	Population in 2000	Change (%)
Cochise County	Benson	3,824	4,711	23.2
	Bisbee	6,288	6,090	-3.1
	Douglas	13,137	14,312	8.9
	Huachuca City	1,782	1,751	-1.7
	Sierra Vista	32,983	37,775	14.5
	Tombstone	1,220	1,504	23.3
	Willcox	3,122	3,733	19.6
Graham County	Pima	1,725	1,989	15.3
	Safford	7,359	9,232	25.5
	Thatcher	3,763	4,022	6.9
Greenlee County	Clifton	2,840	2,596	-8.6
	Duncan	662	812	22.7
Santa Cruz County	Nogales	19,489	20,878	7.1
	Patagonia	888	881	-0.8

Appendix C: Arizona Population Density by Jurisdiction (2000)

	Jurisdiction	Land Area	Population	Population Density
Mohave County	Bullhead City	45.2	33,769	746.6
	Colorado City	10.5	3,334	317.3
	Kingman	30.0	20,069	669.7
	Lake Havasu City	43.0	41,938	974.4
La Paz County	Parker	22.0	3,140	142.8
	Quartzsite	36.3	3,354	92.4
Yuma County	San Luis	26.4	15,322	579.5
	Somerton	1.3	7,266	5,483.2
	Wellton	2.5	1,829	727.3
	Yuma	106.7	77,515	726.8
Pima County	Marana	72.7	13,556	186.6
	Oro Valley	31.8	29,700	933.1
	Sahuarita	15.2	3,242	213.2
	South Tucson	1.0	5,490	5,446.6
	Tucson	194.7	486,699	2,500.1
Apache County	Eagar	11.3	4,033	355.6
	Saint Johns	6.6	3,269	494.8
	Springerville	11.5	1,972	170.8
Coconino County	Flagstaff	63.6	52,894	831.9
	Fredonia	7.4	1,036	139.7
	Page	16.6	6,809	410.5
	Williams	43.5	2,842	65.3
Navajo County	Holbrook	15.4	4,917	318.4
	Pinetop-Lakeside	11.3	3,582	318.1
	Show Low	27.9	7,695	276.2
	Snowflake	30.8	4,460	144.8
	Taylor	24.6	3,176	129.1
	Winslow	12.3	9,520	773.1
Yavapai County	Camp Verde	42.6	9,451	222.0
	Chino Valley	18.6	7,835	421.6
	Clarkdale	7.3	3,422	466.9
	Cottonwood	10.7	9,179	860.3
	Jerome	0.7	329	462.1
	Prescott	37.1	33,938	915.6
	Prescott Valley	31.7	23,535	742.0
	Sedona	18.6	10,192	548.0

	Jurisdiction	Land Area	Population	Population Density
Maricopa County	Avondale	41.3	35,883	869.7
	Buckeye	145.8	6,537	44.8
	Carefree	8.8	2,927	330.8
	Cave Creek	28.2	3,728	132.0
	Chandler	57.9	176,581	3,050.5
	El Mirage	9.7	7,609	786.8
	Fountain Hills	18.2	20,235	1,113.8
	Gila Bend	22.8	1,980	86.7
	Gilbert	43.0	109,697	2,553.7
	Glendale	55.7	218,812	3,929.5
	Goodyear	116.5	18,911	162.4
	Guadalupe	0.8	5,228	6,813.9
	Litchfield Park	3.1	3,810	1,216.6
	Mesa	125.0	396,375	3,171.3
	Paradise Valley	15.5	13,664	881.7
	Peoria	138.0	108,364	2,781.9
	Phoenix	474.9	1,321,045	167.3
	Queen Creek	25.8	4,316	1,100.4
	Scottsdale	184.2	202,705	443.9
	Surprise	69.5	30,848	3,959.4
	Tempe	40.1	158,625	894.1
	Tolleson	5.6	4,974	441.7
	Wickenburg	11.5	5,082	2,296.1
	Youngtown	1.3	3,010	726.8
Gila County	Globe	18.0	7,486	415.5
	Hayden	1.3	892	707.1
	Miami	1.0	1,936	2,008.0
	Payson	19.5	13,620	699.6
	Winkelman	0.7	443	612.3
Pinal County	Apache Junction	34.2	31,814	929.3
	Casa Grande	48.2	25,224	523.7
	Coolidge	5.0	7,786	1,549.1
	Eloy	71.7	10,375	144.8
	Florence	8.3	17,054	2,056.2
	Kearny	2.8	2,249	805.4
	Mammoth	1.1	1,762	1,626.5
	Superior	1.9	3,254	1,684.6

	Jurisdiction	Land Area	Population	Population Density
Cochise County	Benson	35.7	4,711	131.9
	Bisbee	4.8	6,090	1,266.3
	Douglas	7.7	14,312	1,852.7
	Huachuca City	2.8	1,751	626.5
	Sierra Vista	153.5	37,775	246.1
	Tombstone	5.6	1,504	894.1
	Willcox	6.0	3,733	622.3
Graham County	Pima	2.5	1,989	787.0
	Safford	7.9	9,232	1,166.1
	Thatcher	4.4	4,022	919.4
Greenlee County	Clifton	14.9	2,596	174.8
	Duncan	2.6	812	317.6
Santa Cruz County	Nogales	20.8	20,878	1,002.1
	Patagonia	1.2	881	738.7

References

Abdel-Aty, M. A. and A. E. Radwan. 2000. Modeling traffic accident occurrence and involvement. *Accident Analysis and Prevention* 32: 633 – 642.

Arizona Department of Transportation. *Arizona Motor Vehicle Crash Facts 2000*. Prepared by the Motor Vehicle Division in cooperation with the Traffic Records Section. Phoenix, 2001.

Chin, H. C. and M. A. Quddus. 2003. Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersections. *Accident Analysis and Prevention* 35: 253–259.

Greene, W. 2000. *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall.

Harwood, D.W., F. M. Council, E. Hauer, W. E. Hughes, and A.Vogt. 2000. *Prediction of the Expected Safety Performance of Rural Two-Lane Highways*. FHWA-RD-99-207. Washington D.C.: Federal Highway Administration.

Hauer, E., J. C. N. Ng, and J. Lovell. 1988. Estimation of safety at signalized intersections (with discussions and closure). *Transportation Research Record* 1185: 48-61.

Joshua, S. and N. Garber. 1990. Estimating truck accident rate and involvement using linear and Poisson regression models. *Transportation Planning and Technology* 15: 41-58.

Jovanis, P. and H. Chang. 1986. Modeling the relationship of accidents to miles traveled. *Transportation Research Record* 1068: 42-51.

Kim, D.-G., J. Oh, and S. Washington. 2006a. Modeling crash types: new insights into the effects of covariates on crashes at rural intersections. ASCE *Journal of Transportation Engineering* 132 (4):282-292.

Kim, D.-G., Y. Lee, S. Washington, and K. Choi. 2006b. Modeling crash outcome probabilities at rural intersections: application of hierarchical binomial logistic models. *Analysis and Prevention* 39 (1): 125-134.

Ladron de Guevara, F. and S. Washington. 2004. Forecasting crashes at the planning level: a simultaneous negative binomial crash model applied in Tucson, Arizona. *Transportation Research Record* 1897: 191-199.

Lord, D., S. Washington, and J. Ivan. 2004. Poisson, Poisson-Gamma, and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. *Accident Analysis and Prevention* 37 (1): 35-46.

Lyon, C., J. Oh, B. Persaud, S. Washington, and J. Bared. 2003. Empirical investigation of interactive highway safety design model accident prediction algorithm: rural intersections. *Transportation Research Record* 1840: 78–86.

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

Mitra, S., H. C. Chin, and M. A. Quddus. 2002. Study of intersection accidents by maneuver type. *Transportation Research Record* 1784: 43–50.

Oh, J., C. Lyon, S. Washington, B. Persaud, and J. Bared. 2003. Validation of the FHWA crash models for rural intersections: lessons learned. *Transportation Research Record* 1840: 41–49.

Oh, J., S. Washington, and K. Choi. 2004. Development of accident prediction models for rural highway intersections. *Transportation Research Record* 1897: 18-27.

Persaud, B. and T. Nguyen. 1998. Disaggregate safety performance models for signalized intersection on Ontario Provincial roads. *Transportation Research Record* 1635: 113–120.

Shankar, V., F. Mannering, and W. Barfield. 1995. Effect of roadway geometric and environmental factors on rural freeway accident frequencies. *Accident Analysis and Prevention* 27 (3): 371–389.

Stutts, J., W. Hunter, and W. Pein. 1996. Pedestrian-vehicle crash types: an update. *Transportation Research Record* 1538: 68–74.

U.S. Bureau of the Census. *Census Data for State of Arizona, 2000*. Washington D.C., 2003.

Van Schalkwyk, I., M. Sudeshna, and S. Washington. 2006. Incorporating weather into region-wide safety planning prediction models. *Proceedings of the 85th Transportation Research Board Annual Meeting.*

Vogt, A. and J. Bared. 1998. *Accident Prediction Models for Two-Lane Rural Roads: Segments and Intersections.* FHWA-RD-98-133. Washington D.C.: Federal Highway Administration.

Vogt, A. 1999. *Crash Models for Rural Intersections: Four-Lane by Two-Lane Stop-Controlled and Two-Lane by Two-Lane Signalized.* FHWA-RD-99-128. Washington D.C.: Federal Highway Administration.

Washington, S., J. Metarko, I. Fomunung, R. Ross, F. Julian, and E. Moran. 1999. An inter-regional comparison: fatal crashes in the southeastern and non-southeastern United States: preliminary findings. *Accident Analysis & Prevention* 31 (102): 135-146.

Washington, S., M. Karlaftis, and F. Mannering. 2003. *Statistical and Econometric Methods for Transportation Data Analysis.* Boca Raton, FL: Chapman and Hall.

Washington, S., M. Meyer, I. van Schalkwyk, E. Dunbaugh, S. Mitra, and M. Zoll. 2006. Incorporating Safety into Long Range Transportation Planning. NCHRP 8-44. Washington D.C.: Transportation Research Board.

White, D. and S. Washington. 2001. Safety restraint use rate as a function of law enforcement and other factors: preliminary analysis. *Transportation Research Record* 1779: 109-115.

Zellner, A. 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association* 57 (298): 348–368.