Submitted for Presentation at the 2004 Transportation Research Board Annual Meeting

Testing Behavioral Hypotheses on Street Crossing

Xuehao Chu
Center for Urban Transportation Research
University of South Florida
4202 East Fowler Avenue, CUT 100
Tampa, Florida 33620
xchu@cutr.usf.edu
Tel.: (813) 974-9831
Fax: (813) 974-5168

July 2003

Words: 215 (Abstract) + 3,154 (Body) + 2*250 (Tables & Figures) = 3,853

This paper tests several hypotheses on the street crossing behavior of pedestrians. These hypotheses relate to pedestrians' tradeoff between direct attributes such as time and safety, the role of the street environment, the role of pedestrian law, and pedestrians' false sense of security for crossing at a marked crosswalk. These hypotheses are tested with two nested logit models. One is based direct attributes, and the other is based on indirect attributes as represented by the street environment. Both control for personal attributes. These models were estimated with data from a reality-based stated-preference survey in the Tampa Bay area of Florida. Respondents of the survey were placed in real traffic conditions at the curb side and asked to state their crossing choices without actually crossing the street. While indirect attributes are engineering measures, direct attributes are perceived values. Pedestrians do respond to safety improvements, and they are responsive to engineering changes in the individual elements of a street environment. Furthermore, the evidence is consistent with the hypothesis that pedestrians have a false sense of security in a marked crosswalk at uncontrolled locations. On the other hand, there is no evidence that knowledge of street-crossing law affects how pedestrians cross streets. These results have direct implications to all three areas of public policies—engineering, education, and enforcement.

INTRODUCTION

Policy proposals have been put forward over recent years in the United States for improving the pedestrian environment. These proposals fall into three general categories: engineering, enforcement, and education. Engineering proposals tend to help improve the street environment, while enforcement and education proposals help improve the institutional environment. The street environment consists of trip characteristics (e.g., origins and destinations of crossings), traffic conditions (e.g., traffic volume and speed), roadway characteristics (e.g., crosswalks and sidewalks), and control characteristics (e.g., pedestrian signalization and signage). The institutional environment consists of design standards and guidelines, pedestrian laws, and pedestrians' knowledge of these laws. The street and institutional environments combined form the pedestrian environment.

Recent emergence of these policy proposals has been motivated by the perception that investments in the pedestrian environment have many benefits. Travel by automobiles may be reduced if more travel is done through walking or through transit as access and egress may become friendlier. It is argued that less travel by automobiles means a better environment. People are also physically active while walking as a mode of travel. It is argued that being more physically active means better health, reduced health costs, and higher quality of life. In addition to the benefits of more travel by walking, improving the pedestrian environment is also argued to reduce pedestrian injuries from crashes with motor vehicles.

Many of these policy proposals have never been subject to empirical testing as to whether pedestrians would respond as expected. Without pedestrian responses to policy proposals, it would be difficult to realize their expected benefits. This paper tests several hypotheses on the spatial aspect of pedestrian street-crossing behavior that have implications to these policy proposals. This is accomplished through estimating nested logit models of crossing location choices by pedestrians. The estimation uses data collected from a reality-based stated-choice survey in which respondents were placed in real traffic conditions at the curb side and asked to state their crossing choices without actually crossing the street. The rest of the paper has four sections that formulates the hypotheses, describes the data, estimates the models, and concludes the paper, respectively.

HYPOTHESES

The first hypothesis relates to whether pedestrians consciously tradeoff the direct attributes of street crossing options, such as safety and time. During spring 2001, I followed a number of bus runs along several busy routes in Tampa, Florida as part of another research project. Around 40 bus riders were observed crossing streets after getting off their buses. These riders varied in terms of age, gender, physical condition, and whether being with young children. So did the street environment. Almost without exception, however, these pedestrians crossed the street at where they got off their buses. These observations would suggest that pedestrians do little tradeoff between the direct attributes of available crossing options. If pedestrians do not do tradeoffs between safety and time, it would be difficult to influence pedestrian behavior by increasing crossing safety or reducing crossing time through improvements in the street environment for pedestrians.

The second hypothesis relates to the role of the street environment itself in the spatial aspect of pedestrian street-crossing behavior. There are two alternative behavioral foundations for a model in which the street environment directly explains pedestrian crossing behavior. It is possible that the indirect attributes that characterize the street environment determine the direct attributes (i.e., safety, time, etc.) and the model is just a reduced form. It is also possible that it is the indirect attributes that pedestrians respond to rather than the direct attributes of the street environment. An acceptance of the second hypothesis would provide a behavioral foundation for the engineering proposals even if the first hypothesis were rejected. This hypothesis goes beyond Chu et al. (2003) by controlling for pedestrians' personal attributes. It is possible that it is who the pedestrian is rather than what is in the street environment that determines how the pedestrian behaves. Their exclusive focus on the street environment is motivated by the desire to exclude variables for which data would not be readily available for model applications.

The third hypothesis relates to the role of pedestrians' knowledge of pedestrian street-crossing law. Two elements of the law are tested: 1) For street blocks with traffic signals at both ends, pedestrians must cross in a marked crosswalk; and 2) For other street blocks, pedestrians crossing outside a marked crosswalk must yield the right of way to motor vehicle traffic. Educational policy proposals assume that pedestrians do not know pedestrian laws and that they will behave appropriately once they are educated. The test here considers whether pedestrians who know these elements of the law would behave differently. If knowing the law does not influence how pedestrians behave, educational proposals would do little to influence behavior. By focusing on crash involvement by previous efforts in evaluating educational efforts, it was impossible to know why educational efforts may be ineffective: 1) People did not get the message; 2) People did get the message but did not behaviorally respond; and 3) People got the message and behaviorally responded.

The fourth hypothesis relates to how pedestrians may behave with or without a marked crosswalk at locations without traffic signalization. The argument against marking crosswalks at uncontrolled locations is the perception that pedestrians would have a false sense of security while crossing at marked crosswalks. This perception has been formed with the help of several studies of pedestrian involvement in motor vehicle crashes at uncontrolled locations with and without marked crosswalks (Herms, 1972; Gibby et al., 1994; Koepsell et al., 2002). While the degree varies across these studies, they all suffer a lack of control for other factors that would have contributed to the differential crash involvement. The paper tests this hypothesis by determining how the presence of a marked crosswalk at uncontrolled locations changes the role of perceived safety in pedestrian crossing behavior.

DATA

The data were collected from a sample of 86 respondents and 48 street blocks in the Tampa Bay area in Florida in spring of 2002. The respondents were placed in real traffic conditions at the curbside of these street blocks. After observing the street environment at each street block for three minutes, they were asked to state preferred location choices for crossing the street. In addition, they were also asked to provide perceived direct attributes (safety, time, and predictability of time) of each crossing option available at the street block. The perceived attributes were on a scale from 1 to 10 with 10 being the most desirable. Each respondent provided these data for two crossing scenarios at each of six street blocks. Once all six street blocks were finished, the respondents were asked to provide background information, including their age, gender, and household income ranges.

A crossing scenario consisted of these elements: the street environment, the start and end points of the crossing, and the options available. The street environment for a particular scenario is determined through pre-selecting a large set of indirect attributes describing the street environment and through selecting street blocks that vary in wide ranges in these indirect attributes.

To select a start and end combination for scenario, five potential locations for either the start or end point were considered with equal distance between them. For either the start or end point, two potential locations were at the intersections. These potential locations allowed a total of 25 different start-end combinations. Two combinations of start and end points were randomly selected for each street block. The nearside for a scenario is where the start point is.

For a given start-end combination, a set of up to six discrete options was defined that can approximate most of the potentially infinite number of crossing options. These options are labeled as A through F for ease of reference and defined as follows (left and right are relative to the nearside):

- A = Crossing at the left intersection
- B = Crossing at a mid-block start point at a right angle
- C = Crossing with a jaywalk between the start and end points
- D = Walking on the nearside to the opposite side of a mid-block end point and crossing there at a right angle
- E = Crossing at the right intersection
- F = Crossing at a mid-block crosswalk that is away from a start or end point

The exact options depend on the particular start-end combination. In general, there are a total of five possible sets of options from the 25 possible start-end combinations discussed earlier. These are: A-E, A-C-E, A-B-E, A-C-D-E, A-B-C-E, and A-B-C-D-E. On the other hand, option F is available only when a mid-block crosswalk is present and located away from a start or end point.

Figure 1 shows an example of an instrument. In this case, all options are available. The reader is referred to Chu (2003) for details on the design of the reality-based stated-choice survey and, the collection of the data, and a description of the participants and sites.

ESTIMATION

It was hypothesized that the most appropriate econometric model is the nested logit (Ben-Akiva and Lerman, 1985). It is natural to view the six potential options for street crossing as two groups: those related to crossing at intersections and those related to crossing at mid-block locations. That is, the nested logit has a two-level structure. The top level has two branches: intersection and mid-block. The bottom level has two options in the intersection branch (A and E) and up to four options in the mid-block branch (B, C, D, F).

To test the hypotheses, two models are estimated. Both include personal attributes as controls. More important, one focuses on the direct attributes of the street environment as explanatory variables and the other on the indirect attributes of the street environment. Furthermore, the model with the direct attributes is used to test whether pedestrians do tradeoffs in choosing street-crossing locations and whether pedestrians have a false sense of security when crossing in a marked crosswalk at uncontrolled locations. On the other hand, the model with the indirect attributes is used to test the role of the street environment and whether knowledge of pedestrian street-crossing law influences how pedestrians choose street-crossing locations. Both models are estimated with full information maximum likelihood.

Table 1 presents the results with six columns. The first column lists all variables that appear in either model. The second column shows how each variable is specified in relation to the individual crossing options. The last four columns show the results for the indirect-attribute and direct-attribute models, respectively. The common portions of the models are reported at the bottom of the table, including personal attributes, alternative-specific constants, trip attributes in branch utility functions, inclusive values, and overall statistics.

Both models are well behaved. All basic variables have the expected signs. These include direct attributes, indirect attributes, personal attributes, and trip attributes in the branch utility functions. Both models fit the data well. The ρ^2 adjusted for the number of variables is 0.48 for the indirect-attribute model and 0.56 for the direct-attribute model. The models are consistent with utility maximization (Hensher and Green, 2002). Finally, the estimated coefficients of the inclusive values are significantly different from 1, indicating that the nested logit is preferred over the regular logit. The rest of the discussion focuses on the hypotheses.

The evidence from the direct-attribute model is consistent with the hypothesis that pedestrians consciously tradeoff safety, time, and predictability of time in choosing crossing locations. They value safety slightly higher than time and they value predictability less than half of what they value safety and time. It is not surprising that predictability is valued less here. Before stating their preferred location choices for street crossing, the respondents were asked to ignore time constraints such as catching a coming bus on the other side of the street.

The evidence from the indirect-attribute model is consistent with the hypothesis that pedestrian street-crossing behavior is responsive to the street environment. These include crossing distance and roadside walking distance as trip attributes, traffic volume as traffic conditions, crosswalk marking, medians, left-turn lanes, shoulders, and sidewalks as roadway characteristics, and signalization as control characteristics. It cannot be determined within this framework, however, whether the indirect-attribute model is just a reduced form between the direct-attribute model and a relationship between direct and indirect attributes or whether it is the indirect attributes that pedestrians respond to rather than the direct attributes of the street environment.

Testing the other two hypotheses is somewhat more involved. The role of knowing pedestrian street-crossing law is tested in the indirect-attribute model. The two elements of the law are repeated here: 1) For street blocks with traffic signals at both ends, pedestrians must cross in a marked crosswalk; and 2) For other street blocks, pedestrians crossing outside a marked crosswalk must yield the right of way to motor vehicle traffic. For this discussion, the first element is referred to as full signalization and the second as partial signalization. In addition to the dummy variable for marked crosswalk as part of the street environment, two variables are added. One is for testing partial signalization and is a product of three dummy variables: a dummy for no crosswalk marking, a dummy for partial signalization, and a dummy for knowing the law. The other is for testing full signalization and is also a product of three dummy variables: a dummy for a marked crosswalk, a dummy for full signalization, and a dummy for knowing the law. Neither variable is significant. The variable for full signalization has the expected positive sign but the variable for partial signalization has the wrong positive sign.

To test the hypothesis on the role of crosswalk marking on perceived safety in the direct-attribute model, a product of perceived safety, a dummy for a marked crosswalk, and a dummy for no traffic signalization was added. Separate coefficients are estimated for intersection and mid-block options. The coefficient for intersection options is significant while that for mid-block options is marginally significant. Both coefficients are positive. That is, the presence of a marked crosswalk at an uncontrolled location increases the role of safety in a pedestrian choosing the option with this marked crosswalk. In addition, this increased role of safety is larger with intersections than mid-block locations. That is, the presence of the marking adds the perceived level of safety, and this added safety is higher at intersections than at other places. This is consistent with the hypothesis that pedestrians have a false sense of security in a marked crosswalk at uncontrolled locations.

CONCLUSION

The paper has tested four hypotheses on the spatial aspect of pedestrian street-crossing behavior with two nested logit models estimated using data from a reality-based stated-choice survey. It is critical to recognize the shortcomings of the paper in understanding its contributions. Some of these shortcomings are behavioral while others are data related.

One behavioral shortcoming is that the modeling does not account for the dynamics of traffic conditions and pedestrian's street crossing behavior. The models relate the average traffic conditions during a three-minute period with how a pedestrian may have chosen to cross a street block under such average conditions. Whether safe traffic gaps are available can change quickly over time and across locations along a street block. Such temporal and spatial dynamics in traffic conditions lead to dynamics in the street crossing behavior of pedestrians as well.

Another behavioral shortcoming is that the modeling ignores the role of time constraints. Relative to other direct attributes, time predictability would become far more important to a pedestrian when he has a tight time constraint. As a result, he may take crossing options with relatively lower safety levels. By excluding time constraints, the usefulness of the model is reduced in understanding the behavior of transit users in trying to catch a coming bus on the other side of the road.

One data shortcoming is that the sample was recruited through a temporary staffing agency. This approach to selecting participants gave greater certainty in the number of recruited participants who actually would show up. Given the fact that completing the field surveys for any given participant took about five hours, recruiting volunteers through random sampling of residents in the study area would not have worked as well. On the other hand, the sample does include participants with a wide range of gender, age, income, and other personal and household attributes.

Despite of these shortcomings, the paper makes several contributions to the literature on policy debates over improving the pedestrian environment and on modeling pedestrian street-crossing behavior. The hypothesis testing indicates that many policy proposals to improve the street environment have a sound behavioral foundation because pedestrians appear to directly respond to changes in the street environment. The hypothesis testing also indicates that pedestrians appear to indirectly respond to policy proposals that aim at improving the safety, crossing time, and time predictability for street crossing. In addition, knowledge of pedestrian street-crossing law does not appear to influence where pedestrians would choose to cross a street. Finally, the evidence is consistent with the perception that pedestrians get a false sense of security while crossing in a marked crosswalk at locations without traffic signalization.

In addition to hypothesis testing, the paper contributes to the literature in a number of other ways. It represents the first effort to model pedestrian location choices for street crossing. It uses data from a reality-based stated-choice approach under which respondents face real traffic conditions rather than hypothetical ones under the standard approach to stated-preference surveys. The model can be potentially used to evaluate the benefit of policy proposals for improving the pedestrian environment. Such evaluation has become increasingly important because improvements in the pedestrian environment often negatively affect motor-vehicle traffic.

ACKNOWLEGMENT

The Florida Department of Transportation provided funding through the National Center for Transit Research. I like to thank District Seven of the Department for performing the traffic counts and the HARTline for providing bus transportation. A previous version was presented at the 49th Annual North American Meetings of the Regional Science Association International in San Juan, Puerto Rico, November 14-16, 2002. I thank Steve Polzin for comments on an earlier version. However, I am solely responsible for all statements and opinions expressed.

REFERENCES

Ben-Akiva, Moshe, and Steven R. Lerman (1985), Discrete Choice Analysis: Theory and Application to Travel Demand, the MIT Press, Cambridge, Massachusetts.

Chu, Xuehao (2003), A Reality-Based Approach to Stated-Preference Surveys, paper submitted to the 2004 Transportation Research Board Annual Meeting.

Chu, Xuehao, Martin Guttenplan, and Michael R. Baltes (2003), Why People Cross Where They Do: The Role of the Street Environment, paper submitted to the 2004 Transportation Research Board Annual Meeting.

Gibby, A. Reed, Janice Stites, Glen S. Thurgood, and Thomas C. Ferrara (1994), Evaluation of Marked and Unmarked Crosswalks at Intersections in California: Final Report, California Department of Transportation, Sacramento, California.

Hensher, David A., and William H. Greene (2002), Specification and Estimation of the Nested Logit Model: Alternative Normalizations, Transportation Research B 36: 1-17.

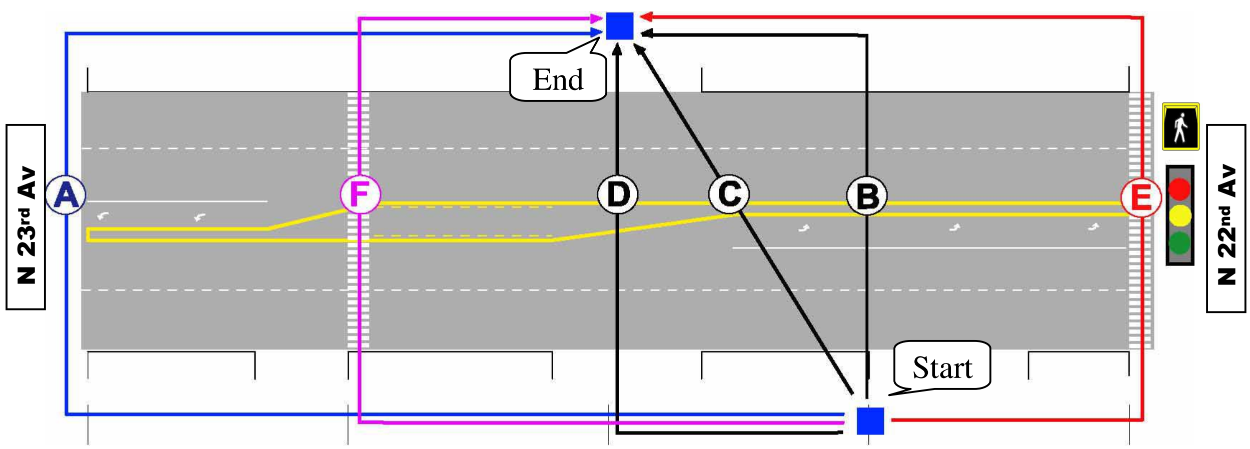
Herms, BF (1972), Pedestrian Crosswalk Study: Accidents in Painted and Unpainded Crosswalks, Highway Research Record 406: 1-13.

Koepsell, Thomas, Lon McCloskey, Marsha Wolf, Anne Vernez Moudon, David Buchner, Jess Kraus, and Matthew Patterson (2002), Crosswalk Markings and the Risk of Pedestrian-Motor Vehicle Collisions in Older Pedestrians, Journal of American Medical Association 17: 2136-2143.

Figure 1. Sample Survey Instrument for Stated Preferences

Please enter your PIN here: _______________

The diagram below shows your start point, your end point, and your location options for crossing the street within this block.



Please stand at your start point and observe the block characteristics and traffic conditions for 3 minutes. Based on your observation of the block and evaluation of the options during these 3 minutes,

1. Please tell us your choice for crossing this street by selecting **one** from below:

A **F** **D** **C** **B** **E**

2. Please tell us how you feel about these crossing options in terms of four attributes: safety, crossing time, and predictability of crossing time. State the levels of these crossing attributes in the table below **on a scale from 1** (least desirable) **to 10** (most desirable).

Crossing Attributes	**Crossing Options**					
	A	**F**	**D**	**C**	**B**	**E**
Safety						
Crossing Time						
Predictability of Crossing Time						

Thank you for your valuable inputs.

Table 1. Full Information Maximum Likelihood Estimation

Variables	Specification	Model Using Indirect Attributes		Model Using Direct Attributes	
		Coeff.	t-ratio	Coeff.	t-ratio
Safety (on a 1-10 scale)	Generic			0.3658	10.31
Time (on a 1-10 scale)	Generic			0.3497	9.65
Time predictability (on a 1-10 scale)	Generic			0.1501	3.99
Safety * Marking * No signal	Intersections			0.1022	3.32
Safety * Marking * No signal	Mid-block Locations			0.0557	1.30
Crossing distance (feet)	Generic	-0.0032	-2.66		
Roadside walking (feet)	Generic	-0.0035	-11.85		
Traffic volume (vehicles per hour)	Mid-block Locations	-0.0002	-1.00		
Crosswalk marking (1 if marked)	Intersection	1.3058	4.62		
Crosswalk marking (1 if marked)	Mid-block Locations	0.8301	3.12		
Painted median (feet)	Mid-block Locations	0.0467	1.99		
Restrictive median (feet)	Mid-block Locations	0.0213	1.26		
Left-turn lanes (feet)	Generic	0.0487	3.21		
Nearside shoulder (1 if present)	B, C	-0.0858	-1.20		
Farside shoulder (1 if present)	C, D	-0.1732	-2.07		
Nearside sidewalk (1 if present)	B, C	-0.2554	-0.90		
Farside sidewalk (1 if present)	C, D	-0.4695	-1.51		
Traffic signal (1 if present)	Intersections	0.4947	2.28		
Pedestrian signal (1 if present)	Intersections	1.0806	4.05		
No marking * Partial signalization * Know law	Generic	0.0751	0.34		
Marking * Full signalization * Know law	Generic	0.1846	0.64		
Income (1 if ≥ $60,000)	Mid-block Locations	0.7514	2.35	0.6061	1.79
Age (1 if 65 years or older)	Mid-block Locations	-1.6032	-3.72	-0.6447	-1.51
Gender (1 if male)	Mid-block Locations	0.7252	3.07	0.7105	2.99
Alternative specific constant	B	1.7584	4.40	1.3037	5.82
Alternative specific constant	C	0.7116	1.26	-0.2211	-0.58
Alternative specific constant	D	1.4657	3.78	0.6631	2.75
Alternative specific constant	Intersection Branch	1.7078	3.17	1.0762	1.90
Start and end at mid-block locations	Mid-block Branch	1.8847	3.06	1.9887	5.47
Start at mid-block & end at intersection	Mid-block Branch	1.0906	2.58	0.6642	2.03
Inclusive Value	Intersection Branch	0.9542	7.43	0.9453	8.87
Inclusive Value	Mid-block Branch	0.8657	5.37	0.8835	9.52
Number of observations			1,005		981
Log likelihood function					
Zero			-1,730.8		-1690.6
Convergence			-899.3		-739.4
Adjusted ρ^2			0.48		0.56

NLOGIT 3.0 of Econometric Software, Inc. was used to estimate this model with full information maximum likelihood. The RU1 normalization was used for the scale parameters. The nested logit model has two levels with variable options across observations. Options B and C involve no nearside walking and C and D involve in farside walking. The side of a block with the start point was called the nearside and the other the farside. The reported t-ratios do not correct for potential underestimation in standard errors due to the repeated observations from individual respondents.