Discussion Papers in Economics

Walk or Wait? An Empirical Analysis of street Crossing Decisions

Sanghamitra Das

Charles F. Manski

Mark D. Manuszak

November 2003

Discussion Paper 03-09

Indian Statistical Institute, Delhi
Planning Unit
7 S.J.S. Sansanwal Marg, New Delhi 110 016, India

Walk or Wait? An Empirical Analysis of Street Crossing Decisions

Sanghamitra Das
Planning Unit, Indian Statistical Institute at Delhi

Charles F. Manski
Department of Economics and Institute for Policy Research, Northwestern University

and

Mark D. Manuszak
Graduate School of Industrial Administration, Carnegie Mellon University

November 2003

The research of Charles F. Manski was supported in part by NSF grant SES-0001436 and 0314312. The data collection was funded by a grant from the Society of Indian Automobile Manufacturers to the Transportation Research and Injury Prevention Program of the Indian Institute of Technology, New Delhi. We are grateful to two anonymous reviewers for their comments.

Corresponding author: Mark Manuszak; address: Graduate School of Industrial Administration Rm 330, 5000 Forbes Ave., Pittsburgh PA 15213; email: manuszak@andrew.cmu.edu; phone: 412-268-5993.

Abstract

We examine the behavior of pedestrians wishing to cross a stream of traffic at signalized intersections. We model each pedestrian as making a discrete crossing choice by comparing the gaps between vehicles in traffic to an individual-specific "critical gap" that characterizes the individual's minimal acceptable gap. We propose both parametric and nonparametric approaches to estimate the distribution of critical gaps in the population of pedestrians. To estimate the model, we gather field data on crossing decisions and vehicle flows at three intersections in New Delhi. The estimates provide information about heterogeneity in critical gaps across pedestrians and intersections and permit simulation of the effect of changes in traffic light sequence (flows) on pedestrian crossing behavior and waiting times.

1. Introduction

Juxtaposed against the major life course decisions that social scientists commonly study, street crossing may seem, well, a pedestrian topic for empirical research. Yet serious policy concerns, scientific objectives, and practical considerations motivate empirical analysis of this routine aspect of human behavior.

The policy concern is that the design and operation of road systems poses a complex social planning problem. Society ostensibly wishes to maximize vehicular and pedestrian mobility, maximize traffic safety, and minimize the cost of road systems. Effective resolution of the tension between these objectives requires some level of understanding of driver and pedestrian behavior.

The scientific motivation is that street crossing is an intriguing decision problem whose analysis may shed light on how humans value their time and their lives, how they perceive their environments, and how they interact with one another. A pedestrian who observes oncoming road traffic faces an optimal stopping problem: Walk or wait? Walking carries a risk of accident, with possible injury or loss of life; where jaywalking ordinances are enforced, walking also may carry a risk of fine. Waiting entails a loss of time until a suitable future opportunity to walk should arise. How do pedestrians balance the savings of time achieved by walking against the associated risks of accident and fine? How do they judge the risks of accident and fines in different crossing environments? How do pedestrians interact with one another as they make their crossing decisions? How do pedestrians and drivers interact as they compete for use of road space?

The practical consideration is that, relative to many other aspects of human behavior, street crossing is unusually amenable to empirical study. Videotaping makes it possible to independently and unobtrusively observe the traffic conditions that pedestrians face and the crossing decisions that they make. In contrast, empirical research on other aspects of human behavior typically has to make do with data collected at some remove from actual decision settings; for example, self-reports of behavior by respondents to surveys or observations of behavior in experimental settings.

This paper analyzes videotape data that we have collected in New Delhi at three signalized intersections with substantial vehicular and pedestrian traffic. Study of behavior at signalized intersections is advantageous for policy, analytical, and practical reasons. Traffic planners can realistically contemplate modifying traffic flows by altering signal sequences or by use of police to enforce vehicle and pedestrian traffic ordinances. Signalized intersections have analytical appeal because the regular alteration of signals between red and green produces substantial exogenous time series variation in the traffic conditions that pedestrians face. The practical appeal is that a carefully placed videotape machine can inexpensively record extensive data on crossings.

Our specific focus on pedestrian behavior in Delhi reflects logistical, policy, and analytical considerations. The fact that one of us is located in Delhi facilitated data collection there. Delhi is an environment with a high rate of pedestrian accidents and minimal enforcement of jaywalking ordinances; hence pedestrian safety should be an important policy issue. The minimal enforcement of jaywalking ordinances also simplifies analysis of pedestrian behavior. With little or no risk of being fined for jaywalking, pedestrians in Delhi presumably make their crossing decisions purely by trading off time and safety.

Section 2 describes the three Delhi intersections and explains how we videotaped traffic conditions and pedestrian crossings. We detail the manner in which we coded the videotape data for analysis. We then characterize the observed pedestrians and their distribution of waiting time at each intersection. The waiting time distribution at an intersection, which is determined by pedestrian crossing decisions under the prevailing road traffic conditions, summarizes how road traffic affects pedestrian mobility. Traffic planning policy affects the welfare of pedestrians by influencing this distribution.

Sections 3 and 4 present our analysis of pedestrian crossing decisions. We assume throughout that pedestrians judge the safety of crossing by the time available until a vehicle will next intersect the perpendicular path from their current position to the opposing sidewalk. At each decision instant, each pedestrian has a *critical gap* in mind; the pedestrian chooses to walk if the time until the next vehicle passes is greater than the critical gap and chooses to continue waiting otherwise. Pedestrians and intersections are heterogeneous, so the critical gap may vary across pedestrians and intersections. The critical gap assumption maintained here is akin to the *reservation price* property commonly assumed, or derived from more primitive assumptions, in studies of job search, auctions, and elsewhere (McCall 1970, Wolpin 1987, Rust 1987).

In Section 3, we further assume that the critical gap for each pedestrian at a given intersection does not vary during the period that this pedestrian waits to cross. Under this *time-invariance* assumption, our videotape data makes it possible to estimate the distribution of critical gaps among the pedestrians at each intersection. Estimation of the gap distribution at an intersection enables prediction of crossing behavior as a function of road traffic conditions; for example, we can predict how the distribution of pedestrian waiting times would change if the volume of vehicular traffic were to change or if planners were to alter signal sequences. Comparison of gap distributions across demographic groups sheds light on how crossing rules vary with the gender and age of the pedestrian. Comparison of gap distributions at different intersections shows how crossing behavior varies with the configuration of the road to be crossed. Simulation of pedestrian behavior under hypothetical changes in the traffic light sequence illustrates how our empirical analysis may be useful to traffic planners.

The analysis of Section 3 has partial precedents in the transportation literature (Tanner 1951, Daganzo 1981, Mahmassani and Sheffi 1981, Palamarthy et al. 1994). These studies also attempt to estimate the distribution of critical gaps in the population of pedestrians. From a methodological perspective, our work differs in that we initially estimate this distribution nonparametrically rather than begin with a tightly parametrized model, as has been the practice. From a substantive perspective, our findings are of interest because we examine crossing situations in which large numbers of pedestrians and vehicles interact with frequent adverse consequences.

Our estimates of critical gap distributions suggest that few pedestrians would be willing to cross at gaps of less than 2 seconds while most would be willing to accept a gap of at least 8 seconds. We find significant heterogeneity in crossing behavior across intersections; pedestrians are more inclined to accept small gaps at the smallest intersection that we study. We also find that behavior varies depending on whether the pedestrian is initiating a crossing from the sidewalk or continuing a crossing from the median. When we use the estimated critical gap distributions to simulate crossing behavior under alternative traffic signal sequences, we find that traffic planning can noticeably affect the propensity of pedestrians to wait for a safe signal before crossing.

In Section 4, we examine a variety of extensions of the basic model. First, we investigate how the composition of the traffic flow influences pedestrian crossing behavior. We find that pedestrians behave more cautiously when facing larger vehicles. Second, we exploit the panel nature of the data to examine the degree to which individual pedestrians behave similarly when crossing from the sidewalk and from the median of an intersection. We find that the decisions of a given pedestrian are not highly correlated across these decision points. We also discuss, but do not examine empirically, how pedestrians may interact with one another when crossing.

2. Data Collection and Descriptive Statistics

2.1. Configuration and Operation of the Three Intersections

The three intersections studied are located at Motibagh, Africa Avenue, and All India Institute of Medical Sciences (AIIMS). Each intersection is a four-armed junction on one or the other of the two Ring Roads (beltways) of Delhi. All roads have sidewalks where pedestrians may safely wait prior to crossing, as well as medians where they may safely wait in mid-crossing. All intersections have police booths at which police were present during the periods that we videotaped.

The configurations of the intersections are presented in Figure 1 through Figure 3. The heading of each figure gives the name of the intersection and the date (month/day/year) on which videotaping was performed.

Arrows show the directions in which vehicles are permitted to travel. The figures show that vehicles in India follow left-hand side driving. They also show the existence of exit lanes which permit vehicles to turn left before reaching an intersection.

In each figure, the arms of the intersection are denoted clockwise as arms 1 through 4. In each case, the placement of the camera is such that it views the zebra crossing at arm 2. The *intersection measurements* below each figure describe this zebra crossing, specifying the distance from the near sidewalk (point C) to the median (point B), and from the median to the far sidewalk (point A). Africa Ave. is a significantly smaller intersection than the other two, while AIIMS is slightly larger than Motibagh. For a pedestrian traveling at approximately 1.5 meters per second (about 3.4 miles per hour), the travel time between the sidewalk and median would be roughly 7, 9, and 10 seconds for Africa Ave., Motibagh and AIIMS respectively.

The *phase diagram* to the right of each figure describes the sequence of traffic lights during one complete signal cycle. In general, two nonconflicting streams of traffic move simultaneously. An origin and destination lane characterizes each traffic stream. These lanes correspond to the numbered arteries on the intersection diagrams. At the Motibagh intersection, for example, vehicles in traffic stream 24 are travelling across the intersection from Vasantvihar (arm 2) towards Chankyapuri (arm 4). At any time, vehicles travelling in a given direction face one of three signals: green (1), yellow (2), or red (3). For each of the four-armed intersections, a 5-digit number describes the signal pattern. The first two digits respectively denote the arm of origin and the arm of destination of the flow of vehicles in one direction, the third and fourth denote the same for vehicles from another direction, and the last digit gives the color of the light facing those vehicles. At Motibagh, the first signal pattern is 31-13-1 which has a duration of one minute. This pattern means that vehicles travelling from 3 to 1 (31) and vehicles in the opposite traffic direction (13) face a green signal (1). Since only two vehicle streams are (legally) permitted to move at any time, vehicles travelling in a direction not noted in a given signal code face a red signal. Thus, a single signal description suffices to know the signals faced by all vehicles attempting to travel through an intersection in every direction at a given time.

For some intersections, the signal phase changes periodically. In these instances, there are two phase diagrams. Figures 2 and 3 for AIIMS and Africa Avenue respectively note the timing of such changes as well as the timing of the new cycle. Beyond these periodic and known changes in the signal pattern, however, the cycles are stable. These intersections do not contain pedestrian push-buttons or traffic sensors which could alter the signal phase from one instant to another.

The figures also provide some additional details about the intersections and the filming. Included in these details is the time at which the intersection was filmed. AIIMS was recorded in the morning while the others were filmed during afternoon hours. These recording episodes overlap with peak traffic intervals which are generally viewed as 9-11 a.m and 5-8 p.m. Finally, except for AIIMS which was filmed in cloudy conditions, all intersections were filmed during normal weather conditions.

2.2. Videotaping Road Traffic and Pedestrian Crossings

Initially, we considered videotaping pedestrian and vehicular traffic crossing all arms of each intersection. However, integrated analysis of all arms requires coordinated data collection using multiple cameras, a daunting task. Hence, we eventually decided to focus on one arm of each intersection, for which a single camera suffices. At all intersections except for AIIMS, we were able to position the camera to provide an unimpeded view of all pedestrians crossing the relevant arm. The position of the police booth and bill boards at AIIMS somewhat narrowed the effective field of vision. The result was that we were able to record only a subset of the pedestrians at this intersection.

Although a single camera sufficed to record the relevant pedestrian and vehicular traffic, the camera was not able to record the contemporaneous traffic light sequence. To solve this problem, we used a stopwatch to determine the light sequence and began taping precisely at the beginning of a new light cycle. As discussed earlier, this light sequence did not change except at known times for AIIMS and Africa Avenue. This regularity allowed us to deduce the light pattern at every instant of the subsequent taping period from the timers on the tapes.

Choosing a camera required considerable care. Ordinary camcorders are only capable of timing events to the nearest second at best. We required much finer time resolution in order to adequately establish the status of the traffic lights and traffic conditions as pedestrians made their crossing decisions. Hence, we engaged the services of a professional videotaping company which used a high-quality Umatic digital camera equipped with a frame-by-frame timer, a frame being taken every 1/25 of a second. The resulting films can be stopped at each frame. Thus, we were able to effectively obtain vehicle and pedestrian information at each instant.

At each intersection, we taped for periods of roughly one to two hours. The cassette of the Umatic camera holds about twenty-two minutes of tape, which meant using up to six cassettes per intersection. Taping had to be momentarily interrupted to change cassettes and, on occasion, to change batteries. In order to maintain knowledge of the status of traffic signals, it was very important to record the precise time periods of these interruptions in taping.

2.3 Coding and Interpreting the Videotape Data

The videotape data were coded at the laboratory of the Traffic Research and Injury Prevention Program (TRIPP) at the Indian Institute of Technology, Delhi. Each pedestrian was viewed in slow motion, the tape progressing one frame (1/25 seconds) at a time. The frames were displayed on a 29” screen and two research assistants together coded the data. One viewed the video and announced relevant data values, while the other wrote the values on paper. The coded data were later transcribed for computer processing.

Coding was a very laborious task. We videotaped for five hours and twenty minutes at the three intersections, yielding 480,000 frames. Moreover, the tape has to be viewed many times to code all of the relevant information for pedestrians and vehicles. To reduce errors to the degree feasible, we performed twice the entire process of tape viewing, data coding, and computer transcription. In all, we coded data on 1,275 pedestrians and 6,980 vehicles.

Two sets of variables were coded for each pedestrian. The first set describes the pedestrian's attributes and movements. The coded attributes include gender, age group, and situation (with or without heavy baggage, with or without children, handicapped). The movement information includes the direction of the crossing, the time of arrival at the intersection, the time that crossing begins, the times of arrival at and departure from the median, and the time at which crossing is completed. Moreover, we established the status of the traffic lights at each of these times by matching the decision times to the known light sequences in the manner described earlier.

The second set of variables describes the flow of potentially conflicting vehicles during the period that a pedestrian waits and crosses. For each vehicle, we coded the location of the pedestrian during the period of potential conflict (origin sidewalk, near road lane, median, far road lane, or destination sidewalk), the position of the vehicle relative to the pedestrian (whether it passes in front of or behind the pedestrian), and the type of vehicle. (Indian roads witness a much greater diversity in the types of vehicles than in the West. Besides the usual motorized vehicles, nonmotorized vehicles such as carts and rickshaws are also present.) Moreover, we coded the direction of travel, lane of travel, and the precise period during which the vehicle physically conflicts with crossing.

To measure the period of the conflict with crossing, we placed a reference line on the video screen along the perpendicular path from the near sidewalk to the far sidewalk (i.e., from point A to point C on Figures 1. For each vehicle taped, we coded the times in frames in which the front and rear ends of the vehicle crossed this line. The interval between these two times constitutes the period of physical conflict.

2.4. Characteristics of the Observed Pedestrians and Their Crossings

Table 1 describes demographic characteristics of the observed pedestrians. Recording gender was straightforward, but we could only categorize pedestrians into four broad age groups: child, young adult, middle age and old. The table shows that we largely observe pedestrians who are male and young adults. This reflects the fact that pedestrians at the intersections and times we filmed tend to be persons traveling to or from work, who are disproportionately male and young in Delhi.

Table 2 describes the types of pedestrian crossings that we observe. The table categorizes pedestrians by the point of departure for their crossings and by the state of the traffic signal when they arrive at the departure point. The point of departure can be a sidewalk (origin) or median (intermediate point). Consider, for example, the Motibagh diagram in Figure 1. Points C and A are sidewalk departure points, while point B is the median between these two points. We consider sidewalks and medians to be distinct points of departure because, as noted earlier, the medians at these intersections are wide enough for pedestrians to wait safely at these points. Thus, our analysis of crossing behavior will consider each pedestrian to make two crossing decisions – first when to cross from a sidewalk to the median, and then when to cross from the median to the opposing sidewalk.[1] The panels of Table 2 provide information about pedestrians at each of these points.

We categorize the traffic signal initially facing a pedestrian as "safe" or "unsafe" according to whether the pedestrian should be able to cross the upcoming leg of her journey with no interference from vehicles. For instance, a pedestrian at point C at Motibagh faces a "safe" signal upon arrival when the signal is 32-14-1 or 31-13-1. The former implies that vehicles are permitted to turn across the intersection only into the lane furthest from the pedestrian, thereby leaving the immediate lane open. The latter implies that vehicles are only

[1] Occasionally a pedestrian enters or leaves the camera's field of vision in the middle of the intersection. In these cases, we only observe one of the pedestrian's two crossing decisions. These pedestrians are noted in table 1 as observations with partial crossings rather than complete crossings.

permitted to pass through the intersection on a course that is parallel to the pedestrian's path.

This characterization of safe signals is accurate, of course, only to the extent that vehicles adhere to traffic laws; vehicular traffic violations could render the signals ineffective in practice. Table 2 indicates that vehicles generally do obey the traffic signals. Around 98.1% of all pedestrians who face a safe signal on arrival at the intersection cross immediately, without waiting at all.[2] Thus, pedestrians who encounter a safe signal on arrival appear to generally face a straightforward crossing decision, with no tension between waiting time and safety. As a result, we focus our analysis on the more interesting cases of pedestrians who arrive at an unsafe signal. In all, we observe 697 such pedestrian observations. In some of these cases, the same pedestrian is observed at two unsafe signals, one on arrival at the intersection and the other at the median. In other cases, a pedestrian is observed at one unsafe signal and one safe signal. In yet other cases, we observe pedestrians at only a single decision point, as noted earlier. Hence, the 697 observations comprise an unbalanced panel.

Pedestrians who arrive at a decision point facing an unsafe signal experience a variety of waiting times. Table 2 provides some descriptive statistics on waiting times for these pedestrians.[3] Some pedestrians do not wait at all upon reaching the intersection. However, 63.7% wait at least some amount of time before crossing. The lengths of these waits vary widely, with the average pedestrian waiting 20.8 seconds conditional on waiting at all.

3. The Time-Invariant Critical Gap Model

In this section and the next, we present econometric models of street-crossing behavior. We use our

[2] Of the 1.2% of cases in which a pedestrians does wait at a safe signal, most wait either a very short time or a very long time without facing any conflicting vehicles, suggesting that these do not represent common situations.

[3] In table 2, waiting time is the time until a pedestrian accepts a gap in the traffic and begins to cross, as defined in Section 3.

data on pedestrian choices at the three intersections in Delhi to estimate a priori unconstrained features of the models. We also give illustrative applications.

In both sections, we assume that the pedestrians observed to cross at each intersection are a random sample of the pedestrians who potentially cross at the intersection. We assume that the flow of vehicle traffic is exogenous with respect to pedestrian crossing; that is, the presence of pedestrians does not affect driver behavior. The traffic flow produces a sequence of exogenous time *gaps* between successive vehicles. Each successive gap confronts a waiting pedestrian with a choice problem: Walk now (i. e., accept the gap) or wait longer (i. e., reject the gap)? We assume that a waiting pedestrian watches the traffic flow, observes the duration of each gap as it appears, and accepts the first gap whose duration exceeds his *critical gap*, the minimal gap that the pedestrian is willing to accept. Thus, we assume that street crossing decisions have the reservation-offer property familiar in the analysis of optimal stopping problems.[4]

In this section, we make the strong but enormously simplifying assumption that each pedestrian has a *time-invariant critical gap*.[5] Thus, we assume that the minimal gap a pedestrian crossing at a given

---

[4] We impose the critical-gap property directly, rather than take the space to derive it formally from more primitive considerations. It is reasonable to assume that (1) a pedestrian contemplating crossing at a given gap places a subjective probability on occurrence of a crossing accident; (2) this probability is a decreasing function of the gap size; and (3) the pedestrian suffers a loss if an accident does occur. These assumptions plus standard regularity conditions in the literature on optimal stopping may be used to derive the critical gap property.

[5] Continuing the discussion of footnote 4, the time-invariance property could be derived from a model of optimal stopping with sufficiently strong assumptions. In particular, the property would emerge if pedestrians have infinite horizons, gap sizes are statistically independent over time, and other state variables are time-invariant.

There are various reasons why critical gaps might be time or state dependent. However, we think it difficult to predict the direction of departures from time-invariance. One possibility is that pedestrians become "impatient" as they wait longer, so critical gaps decline with waiting time. Recall, however, that the traffic light follows a fixed time sequence. As waiting time increases, a pedestrian knows that the light status will soon change, making a safe crossing possible. Hence, someone who has been waiting a long time may reasonably become more patient with time, rather than less patient.

Another possibility is that critical gaps vary with the observed traffic flow, with pedestrians forming expectations about future gaps based on their observations of past gaps. Before deciding to work with our model of time-invariant gaps, we gave considerable thought to formulation of a dynamic model with such state-intersection is willing to accept remains the same regardless of the phase of the signal cycle, the volume of vehicular traffic, and the behavior of other pedestrians. Under this assumption, the distribution of decision rules in the population of pedestrians who cross at a given intersection is fully characterized by the distribution of critical gaps in this population. Hence the objective of our empirical analysis is to estimate the distribution of critical gaps for the pedestrians at each intersection.

Section 3.1 describes how we measure gaps and define gap acceptance. Sections 3.2 and 3.3 respectively present nonparametric and parametric estimates of the distribution of critical gaps among the pedestrians who cross at each intersection. Section 3.4 applies the estimates to predict the safety and waiting times of crossings under alternative traffic light sequences. Section 4 begins with a general discussion of reasons why the assumption of time-invariant critical gaps may not be realistic, and then examines a set of specific departures from this assumption.

3.1. Traffic Gaps and Gap Acceptance

In general terms, a gap in the traffic flow is a period of time during which a pedestrian may cross a road without threat of conflict with a passing vehicle. To perform empirical analysis, however, this general notion of a gap does not suffice. We need to specify how gaps are to be measured, and we need to establish a criterion for coding gap acceptance.

Measurement of Gaps

To begin, we define a gap as the time between a reference event and the arrival of the next conflicting vehicle. The first reference event is a pedestrian's appearance at a decision point, which may be the origin

---

dependent critical gaps. However, we found it difficult to formulate a tractable model of expectations formation that we could accept with high credibility.

sidewalk or the median. The *initial gap* faced by a pedestrian is the time interval from his appearance until the arrival of the first vehicle that poses a crossing conflict. The second reference event is the departure of this first conflicting vehicle, and the second gap is the time interval from this event until the arrival of the second conflicting vehicle. Subsequent reference events and gaps are defined in the same manner.

The next step is to define the arrival and departure times of conflicting vehicles. As described in Section 2, we recorded the times at which the front and rear end of each conflicting vehicle crossed a reference line indicating the perpendicular path that pedestrians would ordinarily use to cross the road. We define the arrival and departure of a vehicle to be the times at which its front and rear end respectively cross this line.

In some cases, no conflicting vehicle appears within the twenty-second period following a reference event. The reason may be either a lull in the traffic or a change in the traffic signal that halts all vehicle movement in a certain direction. Rather than expend the effort to code the exact duration of these large gaps, we simply recorded them as *censored gaps* of duration 20+ seconds. This coding is innocuous because every pedestrian facing such a gap chooses to cross.

In other cases, pedestrians face *platoons* of conflicting vehicles moving in the same direction either back-to-back in the same road lane or overlapping but in different road lanes. We coded each platoon as a single virtual vehicle, which arrives when the front end of the first vehicle crosses the reference line and which departs when the rear end of the last vehicle crosses the line. We could alternatively have coded a platoon as a sequence of single vehicles separated by gaps of zero seconds. This more detailed coding was unnecessary because pedestrians never cross while a platoon is passing.

Coding Gap Acceptance

We code a pedestrian as accepting a gap if she crosses in front of the next conflicting vehicle. We code the pedestrian as rejecting all previous gaps; that is, if the oncoming vehicle passes in front of her.

To illustrate, consider a scenario in which a pedestrian appears at the intersection at $t = 0$. Suppose that the first conflicting vehicle arrives at $t = 3$ and departs at $t = 4$, the second arrives at $t = 9$ and departs at $t = 11$, and the third arrives some time after $t = 31$. Suppose that the pedestrian crosses after the second vehicle departs and before the third arrives. Then we code the pedestrian as facing crossing decisions at times {0, 4,11}, with associated gaps of size {3, 5, 20+} where the last gap reflects the censoring described above. The pedestrian rejects the first two gaps and accepts the third. Her waiting time before crossing is 11 seconds.

Alternatively, suppose that the pedestrian crosses between the first and second vehicles. Then we code the pedestrian as facing crossing decisions at times {0, 4}, with gap sizes {3,5}. The pedestrian rejects the first gap, accepts the second, and realizes a waiting time of 4 seconds. In this case, we do not observe the third gap because coding ceased once the pedestrian chose to cross.

Observe that our criteria for coding gap acceptance and for measuring the size of gaps are not dependent on when a pedestrian chooses to move from the sidewalk or median onto the roadway. In the second version of the above illustration, a pedestrian who moves onto the roadway at $t = 2$ but allows the first vehicle to pass in front of her is coded as rejecting the first gap. A pedestrian who chooses to remain on the sidewalk until $t = 5$ and then crosses in front of the second vehicle is coded as accepting a gap of size 5 seconds and as realizing a waiting time of 4 seconds.

3.2. Nonparametric Estimation of the Distribution of Critical Gaps

We now formalize the decision model and present simple nonparametric estimates of the distribution of critical gaps within the population crossing at a given intersection.

The Probability of Gap Acceptance

Let J denote this population and let each pedestrian $j \in J$ have a time-invariant critical gap $\delta_j > 0$.

Suppose that, when pedestrian j arrives at the intersection, she faces an infinite gap sequence $G_j \equiv (g_{jk}, k = 1,\ldots,\infty)$ drawn from a population of potential such sequences. Let $y_{jk} = 1$ if pedestrian j chooses to accept gap k and $y_{jk} = 0$ otherwise. Note that the event $\{y_{jk} = 0\}$ may occur either because the pedestrian rejects this gap or because she accepts an earlier gap.

We assume that δ and G are statistically independent. Under this assumption of *exogenous gap sequences*, the basic construct needed to predict pedestrian behavior is the probability that a pedestrian who is drawn at random from the population and who faces a given gap sequence G chooses to cross at a particular gap, say gap K. This probability is

$$
\begin{aligned}
(1) \quad \text{Prob}(y_K = 1 \mid G) &= \text{Prob}(y_k = 0, k = 1, \ldots, K - 1; y_K = 1 \mid G) \\
&= \text{Prob}(\delta > g_k\ k = 1, \ldots, K - 1; \delta \leq g_K \mid G) \\
&\equiv \text{Prob}(m_K < \delta \leq g_K \mid G) = \max[F(g_K) - F(m_K), 0],
\end{aligned}
$$

where $m_K \equiv \max(g_k, k = 1, \ldots, K - 1)$ is the largest gap faced prior to gap K and where F(.) denotes the population cumulative distribution of critical gaps. The first equality holds because a pedestrian crosses only once. The second follows from the assumption of time-invariant critical gaps and the final one from the statistical independence of δ and G.

It follows from (1) that the hazard rate at gap K, giving the probability that a pedestrian accepts gap K conditional on not crossing earlier, is

$$
(2) \quad \text{Prob}(y_K = 1 \mid y_k = 0, k = 1, \ldots, K - 1; G) = \frac{\max[F(g_K) - F(m_K), 0]}{1 - F(m_K)}.
$$

Observe that the crossing probability and hazard rate at gap K depend only on the size m of the largest previous gap size and the size g of the current gap. Other features of the gap sequence G do not matter, nor does the value of K.

Using Data on Initial Gap Acceptance to Estimate the Distribution of Critical Gaps

The above shows that knowledge of the cumulative gap distribution F(.) suffices to predict the behavior of a randomly drawn pedestrian who faces any given gap sequence. The empirical problem is to use the available data to estimate F(.).

The data are observations of N randomly drawn pedestrians, the sequences of gaps that they reject, and the sizes of the gap that they accept. It is reasonable to think that statistically efficient nonparametric estimates could be obtained by using all of the available data to estimate the hazard rates, subject to the restrictions on F(.) contained in equation (2). However, we are unaware of previous research developing such estimators and do not undertake to do so in this paper. Instead, we present findings using a simple nonparametric estimator that uses only observations of pedestrian behavior at the *initial gaps* they face.

All pedestrians have $m_1 = 0$, so the crossing probability at gap 1 is

$$(3) \quad \text{Prob}(y_1 = 1 \mid G) \;=\; F(g_1).$$

This crossing probability is the conditional expectation $E(y_1 \mid g_1)$. Hence any consistent estimator of $E(y_1 \mid g_1)$ provides a consistent estimate of F(.). The fact that F(.) is a distribution function implies that $E(y_1 \mid g_1)$ is an increasing function of g_1. We shall assume that critical gaps are continuously distributed in the population, which implies that $E(y_1 \mid g_1)$ is a continuous function of g_1. To exploit both the monotonicity and the continuity properties, we use the smoothed monotone regression method of Mukerjee (1988) to estimate F(.). This two-step method first applies the Brunk (1958) isotonic regression estimator and then uses a kernel estimator to smooth the isotonic estimates. This nonparametric approach provides a parsimonious, yet flexible way to estimate the distribution of critical gaps.

Estimates for Young Adults

We have used the smoothed monotone regression method to estimate the distribution of critical gaps among pedestrians crossing at each decision point (i. e., origin or median) and intersection. Disaggregating pedestrians by gender and age, we found that men and women in the same age group have quite similar distributions of critical gaps, but pedestrians appear to differ somewhat by age in their crossing behavior. As shown in Table 1, young adults form by far the largest age group crossing at the three intersections; hence our estimates for this group are much more precise than for the other age groups. The discussion below focuses on this age group.

The main features of the nonparametric findings are presented in the six panels of Figure 4, which show the estimates for young adults (aggregating males and females) across decision points and intersections. In addition to showing the regression estimates and bootstrapped confidence intervals, the figure gives the implied estimates and confidence intervals for the median critical gaps.

These estimates reveal a number of interesting features of crossing behavior. First, almost no pedestrians have critical gaps smaller than 2 seconds, and almost all have critical gaps smaller than 11 seconds.[6] Recall that we code a gap as accepted if a pedestrian crosses the road in front of a vehicle. A gap of two seconds is too small for a pedestrian to plausibly cross the entire roadway at any of the three intersections, but may suffice for someone walking quickly (say 3 meters per second) to cross the nearest travel lane and therefore avoid a conflicting vehicle. A gap of 11 seconds is large enough for a young adult pedestrian to

[6] Critical gaps larger than 11 seconds are indicated only for some crossings from the origin at the Motibagh intersection. However, the sample of pedestrians who make this crossing at an unsafe signal has only 33 members, and the estimated gap distribution is correspondingly imprecise.

complete the longest crossing (i.e., AIIMS) walking at the leisurely pace of 1.5 meters per second.

Second, we find that the critical gap distributions vary across intersections in a sensible manner. Recall that the road at the Africa Avenue intersection is significantly smaller in width than those at Motibagh and AIIMS. We find that at both the origin and median decision points, the critical gap distribution for Africa Ave. lies to the left of those for Motibagh and AIIMS, indicating that pedestrians at Africa Ave tend to be willing to cross at smaller gaps. The median critical gaps at Africa Ave. are estimated to be 4.2 seconds for crossings from the origin and 3.2 seconds for crossings from the median. The corresponding median critical gaps at AIIMS, the largest intersection, are estimated to be 9 seconds and 3.9 seconds respectively.

Third, at each intersection, the critical gap distribution for crossings from the origin sidewalk to the median lies substantially to the right of and shows much more dispersion than the distribution for crossings from the median to the destination sidewalk. Thus, pedestrians appear to be much more cautious and heterogeneous when crossing from the origin to the median than when crossing from the median to the destination. We conjecture that pedestrians view standing at the median to be less safe than standing at the origin and, hence, tend to be willing to accept smaller gaps at the median.

3.3. Parametric Estimation of the Distribution of Critical Gaps

The nonparametric estimates presented in Section 3.2 are appealing in principle because they avoid distributional assumptions, and are attractive in practice because they reveal a good deal about crossing behavior. However, because these estimates are nonparametric and use only the data on acceptance of initial gaps, they are not always as precise as one may like. This section presents complementary estimates that achieve greater precision through parametric restriction of the distribution of critical gaps.

An Ordered Probit Model of Gap Acceptance

We now aggregate the populations of pedestrians crossing at the six observed intersection and decision point locations. We suppose that, for pedestrian j crossing at location i,

$$(4) \quad \log(\delta_{ji}) = \alpha + I_j\gamma + X_j\beta + \lambda_{ji}.$$

Here I_j is a vector of dummy variables indicating the intersection and decision point, X_j is a vector of dummy variables indicating pedestrian age and gender, λ_{ji} is an unobserved variable, and (γ, β, α) are parameters. We suppose further that realizations of λ are statistically independent across pedestrians, decision points and intersections with $N(0, \sigma^2)$ distribution. Then the probability of gap acceptance given in equation (1) has the ordered probit form

$$(5) \qquad \begin{aligned} \text{Prob}(y_K = 1 \mid G, I, X) &= \text{Prob}[\log(m_K) < \log(\delta) \le \log(g_K) \mid G, I, X) \\ &= \max\{\Phi[(\log(g_K) - \alpha - I\gamma - X\beta)/\sigma] - \Phi[(\log(m_K) - \alpha - I\gamma - X\beta)/\sigma],\ 0\}, \end{aligned}$$

where $\Phi(.)$ is the cumulative distribution function of a standard normal random variable.

Maximum Likelihood Parameter Estimates

We estimate the parameters $(\gamma, \beta, \alpha, \sigma)$ by maximum likelihood. Table 3 presents two sets of estimates, one using only the data on acceptance of initial gaps and the other using all of the available data.[7]

[7] The number of observations using all of the data is slightly smaller than the number using only the data on acceptance of initial gaps (N = 678 versus N = 697). This small reduction in sample size is due to the fact that 19 pedestrians were observed to accept a smaller gap than a gap they had previously rejected. This behavior is inconsistent with the maintained assumption that critical gaps are time-invariant. One could construe these

Comparison of the two sets of estimates shows that exploitation of all the data substantially improves precision, with reported asymptotic standard errors typically being 2/3 the magnitude of those obtained when only data on initial gaps are used. The discussion below focuses on the estimates using all of the data.

Table 3 corroborates the nonparametric findings reported earlier. The small magnitude of the coefficient on the gender variable indicates that males and females have similar central tendencies in their distributions of critical gaps. The positive and nearly equal coefficients on the AIIMS and Motibagh intersection dummies indicate that pedestrians crossing at these intersections tend to have larger critical gaps than do those crossing at Africa Ave. The negative coefficient on the median dummy indicates that pedestrians tend to accept smaller gaps when crossing from the median than from the origin sidewalk.

The one new finding present in the parametric estimates is a statistically imprecise but suggestive pattern of coefficients across age groups. It appears that the distribution of critical gaps moves rightward as a function of age, with children tending to accept gaps that old pedestrians would reject.

Table 4 translates the parametric estimates into several percentiles of the critical gap distribution for young adults at each of the decision points and intersections. The table also presents these percentiles for the nonparametric estimates in Figure 4. Not surprisingly, the parametric estimates are statistically more precise than the nonparametric ones. However, the parametric percentile estimates have the same qualitative properties as the nonparametric ones. Here, as before, the distribution of critical gaps at the median decision point for each intersection lies to the right of the distribution at the origin. The distributions of critical gaps for pedestrians crossing at Motibagh and AIIMS lie to the right of the distribution at Africa Ave.

---

observations to be empirical evidence rejecting the time-invariant critical gap model. However, we view the small percentage of such observations (2.73%) to be evidence that the model is generally consistent with observed behavior.

3.4 Simulating the Impact of Changes in the Light Sequence

In this section, we use our nonparametric estimates of the critical gap distribution to perform illustrative policy simulations. In principle, we can simulate at least two kinds of changes to an intersection. First, we can alter the physical structure of an intersection by, for example, changing its dimensions. The differences in our estimated critical gap distributions across intersections suggest that the structure of an intersection substantively affects pedestrian crossing behavior. However, we observe too few intersections to credibly estimate how specific intersection characteristics affect the critical gap distribution. Hence, we do not perform such simulations here.

Second, we can hold the structure of the intersection fixed and alter the traffic flow by, for example, modifying the traffic light sequence. This type of design change would not alter the critical gap distribution. Instead, the light sequence affects pedestrian behavior because it determines the frequency with which pedestrians face oncoming traffic and the length of the unsafe crossing periods. Modification of the traffic light sequence is particularly interesting to contemplate because the light sequence is an instrument that traffic planners can control.

The simulations presented here predict the effects on pedestrians of plausible changes in the light sequence at the AIIMS intersection; similar simulations could be performed at the other intersections. In what follows, we explain how we simulate pedestrian arrivals and crossings, describe the light sequences that we consider, and then report predictions of pedestrian behavior under these light sequences.

Simulating Pedestrian Arrivals and Crossings

To populate the AIIMS intersection, we assume that the number of pedestrians who arrive at the intersection each second is the realization of a Poisson process. For the sake of realism, we use the actual arrival rate per second at AIIMS as the Poisson parameter. For simplicity, we assume that arrivals are statistically independent across seconds. As the number of simulated seconds grows large, this simulation procedure becomes equivalent to assuming that pedestrians arrive uniformly over the light sequence as it repeats over time.

We simulate pedestrian arrivals at the origin decision point for crossings from both A to C and C to A. We generate two critical gaps for each pedestrian, one for the crossing from the origin to the median and the other for the crossing from the median to the destination. We draw these critical gaps at random from the nonparametric distributions of critical gaps that we estimated at the two decision points. Thus, a pedestrian is described by a crossing direction, an arrival time, a critical gap for the origin, and a critical gap for the median. The simulated pedestrians remain the same across the various light sequences that we consider.

For each pedestrian who arrives during an unsafe crossing period, we draw a sequence of gaps by repeated random sampling from the empirical distribution of uncensored initial gaps. It would be more realistic to draw a sequence of gaps from the empirical distribution of gap *sequences*. We are unable to do this because our observed gap sequences are endogenously censored, ending when a pedestrian chooses to cross. Because it eliminates censored gaps, which tend to be large, our procedure for drawing gaps subjects the simulated pedestrians to more challenging gaps sequences than those faced by the actual pedestrians that we observed at AIIMS.

We simulate pedestrian crossings as follows. If the pedestrian arrives at the origin under a safe signal, she crosses immediately. If she faces an unsafe signal, we subject her to the simulated gap sequence. As the pedestrian rejects gaps, we keep track of her waiting time. Under an unsafe signal, the pedestrian crosses when she observes a gap that exceeds her critical gap or when the light changes to a safe signal, whichever comes first. Whether she faces an initially safe signal, crosses under an unsafe signal or waits until a safe signal arises, we suppose that she arrives at the median ten seconds after she departs the origin. As noted earlier, this crossing time implies a speed of approximately 1.5 meters per second (about 3.4 miles per hour). We then analyze the pedestrian's decision problem at the median in an analogous fashion.

The designer of an intersection must consider a variety of conflicting objectives. She ostensibly would like to facilitate vehicular movement while also allowing pedestrians to cross without waiting long periods of time or facing inordinately unsafe conditions. Changing conditions for pedestrians or vehicles moving in one direction may alter outcomes for those moving in different directions. Our simulations provide information about part of the planner's problem. Under each proposed light sequence, we examine the fraction of pedestrians who face unsafe signals at the origin or median, the waiting behavior of pedestrians under unsafe signals, and the fraction of pedestrians who accept small gaps.

Alternative Light Sequences

Table 5 presents the four light sequences that we consider. Sequence 1 is the current light sequence at AIIMS. Sequence 2 removes the two turn signals in the current sequence. Sequence 3 decreases all signal lengths proportionately to yield a total sequence length of 120 seconds. Sequence 4 imposes the light sequence from Africa Ave.

Each light sequence exhibits different characteristics. Relative to the current light sequence at AIIMS, sequence 2 increases the proportion of time that is safe for a complete crossing, but potentially changes the safety of a partial crossing. The presence of turn signals in sequence 1 implies that a pedestrian may face different safety conditions at the origin and median decision points. Sequence 2 removes such asymmetries. Sequences 2 and 3 both decrease the maximum time that a pedestrian would have to wait for a safe signal.

Researchers at TRIPP (Transportation Research and Injury Prevention Programme at the Indian Institute of Technology in Delhi) are interested in the effects of signal cycles in which no color exceeds a minute.[8] This implies that pedestrians will never have to wait more than a minute for a safe signal. While the current light sequence at AIIMS can make pedestrians wait at the origin up to 95 seconds, sequences 2 and 3 substantially reduce this maximum wait. Sequence 4 exhibits substantial asymmetries across decision points; the presence of turn signals yields very different conditions for pedestrians crossing from A to C and from C to A.

Simulation Results

We implement the different light sequences by imposing the light cycles on the simulated time sequence. Recall that we hold the pedestrian characteristics constant across the simulations, including their arrival times, gap sequences, and critical gaps. Hence, the different light sequences alter the safety of the signal facing a given pedestrian as well as the time until a signal change and the nature of that change.

Table 6 reports the results of the simulations. Due to asymmetries in the light sequences across decision points and crossing directions, we separately track pedestrians making crossings in different directions. In addition, we separately report outcomes depending on whether a pedestrian is crossing from the origin or is completing a crossing from the median. The table shows that that altering the light sequence has three primary effects: (1) changes in the percentage of pedestrians who arrive at different decision points under unsafe signals; (2) changes in the percentage of pedestrians who wait until a safe signal; and (3) changes in average waiting time.

For crossings in both directions, the percentage of pedestrians who wait at the origin until a safe signal increases under all three alternative light sequences, relative to the current one. For example, the first and third columns of panel A shows that approximately 23% of pedestrians wait until a safe signal under sequence 2

[8] We are thankful to Professor Geetam Tewari of TRIPP for many helpful discussions on this and related matters.

compared to 15% under the current sequence. These effects reflect changes in the proportion of the light sequence that is safe for departures from the origin. The time until the light changes to a safe signal decreases under the alternative light sequences, relative to the current one. A pedestrian making a crossing from A to C under the current light sequence could potentially wait up to 95 seconds before facing a signal that is safe for a crossing to the median. Under sequences 2 and 3, this time falls to 60 and 65 seconds respectively. As a result, more pedestrians wait until a safe signal.

Compression of the light sequences also leads to shorter waiting times for pedestrians. Pedestrians who cross during unsafe periods experience similar wait times under all of the light sequences. However, with a safe signal arriving sooner under the alternative sequences, pedestrians who wait for a safe signal have shorter waits than under the current light sequence. Consequently, average waiting time falls under all three alternative scenarios.

The outcomes for pedestrians at the origin have implications for conditions at the median. Under most of our alternative light sequences, there is a decrease in the percentage of pedestrians who face an unsafe signal at the median. For sequence 2, this decrease reflects the increased proportion of the light sequence devoted to the globally safe 13-31 signal. Pedestrians who wait until a safe signal at the origin will cross under signal 13-31. Upon reaching the median, they will still face signal 13-31, which is also a safe signal for a median crossing. Pedestrians who leave the origin under an unsafe signal will be more likely to face the signal 13-31 upon their arrival at the median, particularly if they wait any length of time at the origin. Similar effects arise across the other light sequences although light sequence 3, which does not alter the nature of the signals, does not exhibit this feature. In this case, pedestrians will still be as likely to face an unsafe signal at the median whether they cross under an unsafe signal or wait until a safe signal.

The propensity of our simulated pedestrians to wait until a safe signal partly reflects the challenging gap sequences that they face when we draw gaps at random from the distribution of uncensored initial gaps.

However, our simulation results remain qualitatively the same if we alter the gap distribution by, for example, scaling all gaps up by some factor. As long as the time until a safe signal falls and the probability of facing a safe signal at the median rises, we obtain the same basic effects on the propensity to wait for a safe signal, the average waiting time, and the likelihood of reaching the median under an unsafe signal. Simulations based on our estimates enable us to quantify these effects.

4. Extensions of the Basic Model

The basic critical gap model presented in section 3 is attractive due to its parsimony. However, this parsimony is achieved by making a variety of simplifying assumptions about pedestrian behavior. In this section, we consider a number of extensions to the basic model. In Section 4.1 we permit the vehicular composition of the traffic flow to impact crossing decisions. In Section 4.2 we jointly model a pedestrian's crossing behavior at the origin and median decision points. Section 4.3 conjectures social interactions among pedestrians, as well as between pedestrians and drivers.

4.1. Accounting for Vehicle Type

The analysis of Section 3 assumed that traffic flow can be summarized by the gap sequence that a pedestrian faces. However, crossing behavior may depend on features of the conflicting vehicles. For example, the expected cost of an accident may depend on whether the oncoming vehicle is a rickshaw or a bus.

Ideally, we would allow each pedestrian to associate a distinct critical gap with each vehicle type. Thus, pedestrian j crossing at location i would have critical gap δ_{kji} for vehicles of type k. However, it is complex to estimate a model permitting complete heterogeneity in the response of pedestrians to different types of vehicles. Hence we make the simplifying assumption that vehicle type shifts the critical gap for each pedestrian by the same amount; that is, we assume that critical gaps have the form $\delta_{kji} = \delta_{ji} + c(k)$, where c(.) shifts the gap as a function of vehicle type. Going further, we suppose that c(.) is a function of vehicle length, measured in meters.[9] With these assumptions, it is a simple matter to extend the parametric analysis of Section 3.3 to permit critical gaps to vary with vehicle type.

Columns (1) and (2) of Table 7 present two maximum likelihood parametric estimates. The columns differ only in that column (1) supposes that vehicle length has a linear effect on the critical gap, whereas column (2) permits a quadratic effect. Both estimates use data on crossing behavior at uncensored initial gaps. Data on decision making after an initial gap is rejected are excluded because the simple ordered probit model of equation (5) does not hold when critical gaps vary by vehicle type. Censored gaps are excluded because vehicle types were not recorded in these cases.

We find that the vehicular composition of the traffic flow does have an impact on the crossing behavior of pedestrians. Critical gaps clearly increase with vehicle length; the linear and quadratic specifications yield much the same results. The point estimate in the former case indicates that a one-meter increase in vehicle length implies a 0.1445 second increase in the critical gap. Inclusion of the vehicle length variable essentially does not affect the parameter estimates shown earlier in Table 3.

4.2. Bivariate Critical Gap Distributions for Crossing at Origins and Medians

The analysis of Section 3 abstracted from the fact that some pedestrians in our dataset are observed to make two crossings, one from the origin of an intersection and the other from the median. In principle, it is

[9] We also estimated models that use type-specific dummy variables (e.g., car, truck, bus) to differentiate vehicles. We found that vehicle length, which is more parsimonious, performed as well or better empirically.

straightforward to pose a decision model in which each pedestrian has two critical gaps, one for crossing from the origin and the other for crossing from the median. We could then use our data on crossings at these two decision points to estimate the bivariate distribution of critical gaps across the pedestrians at each intersection. While the earlier model also has two critical gaps for pedestrians observed at two decision points, it implicitly assumes that each pedestrian experiences draws from the critical gap distribution that are independent across decision points.

In practice, we have only small samples on which to base estimation of these bivariate distributions. A pedestrian provides bivariate crossing information only if she faces an unsafe signal and uncensored gap at both decision points. Only 66 of the pedestrians whom we observe meet this criterion. Most pedestrians face an unsafe signal and uncensored gap at no more than one decision point and so are uninformative about the bivariate critical gap distribution.

The relevant sample of 66 pedestrians is too small to contemplate either nonparametric estimation or parametric estimation using only initial gap data. However, this sample is large enough to estimate a bivariate probit model using all of the gap data.[10] The third column of Table 7 presents the findings for such a model. The estimated model permits critical gaps to be correlated across decision points and also permits the variances of the marginal critical gap distributions to vary across decision points. In other respects, the model is the same as the one estimated in Section 3.3.

We find that both the mean and the variance of the critical gap distribution for crossings at the origin are larger than for crossings at the median. This reinforces our earlier nonparametric finding that critical gap distributions for crossings from the origin are shifted to the right and have greater dispersion than for crossings from the median. However, we find no evidence of correlation in the critical gaps that pedestrians use at the

[10] Pedestrians facing an unsafe signal at only a single decision point will provide information about the marginal distributions of critical gaps across decision points. Hence, we continue to include these observations in the estimation although they contain no information about correlation in critical gaps across decision points.

two decision points; the estimated correlation is essentially zero and its standard error is large.

4.3. Social Interactions

The analysis of Section 3 viewed pedestrians at making decisions in isolation from one another. We also assumed that drivers do not react to the decisions of pedestrians. Neither assumption is entirely realistic.

Interactions among the pedestrians waiting at a given decision point certainly arise if the pedestrians arrive as a group and wish to continue together. Even if pedestrians are not traveling together, the crossing decisions of one person may influence those of another by changing the likelihood of an accident (i.e. safety in numbers) or by conveying information about the safety of crossing. Interactions among pedestrians could be studied using the formal apparatus of game theory or, less ambitiously, by comparing the behavior of pedestrians who wait alone at an intersection with that of those who wait in groups.

Interactions between pedestrians and automobile drivers arise if drivers adjust speed in response to their observation of pedestrian crossing decisions. Few drivers would willingly hit a pedestrian if such an accident were avoidable. If drivers adjust speed by slowing down when pedestrians are in the road, then the gap sizes that pedestrians face are, to some degree, endogenous rather than exogenous as has been assumed throughout our analysis. In particular, the gaps that we observe when pedestrians cross may be larger than they would have been if pedestrians had chosen to wait. If so, our analysis makes pedestrians appear more cautious than they actually are.

Empirical analysis of social interactions among pedestrians and between pedestrians and drivers would be of great interest. However, we must leave these matters for future research because our data are not sufficiently rich to study them. We do not observe enough cases in which multiple persons wait at an intersection to enable meaningful empirical analysis of interactions among pedestrians. Study of interactions between pedestrians and drivers would require placement of video cameras in positions that show not only what pedestrians see as they wait to cross, but also what drivers see as they look down the road ahead.

5. Conclusion

Economists and other social scientists use statistical methods to examine many aspects of human behavior. In this paper, we consider one of the more mundane problems facing nearly every individual in her day-to-day life, namely the choice of when to cross a road. The need to balance the possibility of an accident with the cost of waiting yields a non-trivial decision problem.

Our analysis rests on the intuitive premise that a pedestrian will only cross an intersection if the time until the arrival of the next vehicle, and hence the time until an accident could occur, exceeds some pedestrian-specific threshold. We use a unique dataset to estimate the distribution of these thresholds in the population of pedestrians crossing at several signalized intersections in New Delhi.

Beginning with a flexible nonparametric model and moving to more restrictive parametric models, our estimates yield a number of interesting and robust conclusions. For example, few pedestrians would be willing to cross at gaps of less than 2 seconds while most would be willing to accept a gap of at least 8 seconds. Behavior varies depending on whether the pedestrian is initiating a crossing from the sidewalk or is continuing a crossing from the median. Pedestrians behave more cautiously when facing larger vehicles than when facing smaller ones.

As discussed in section 4.3, it would be of great interest to be able to extend the analysis to examine interactions among pedestrians, as well as between pedestrians and drivers. It would also be of interest to study more fully how crossing behavior varies with the configuration of the intersection and with attributes of the pedestrian, such as age. Finally, situations in which jay-walking laws are more strictly enforced than in New Delhi could provide insight into the impact of law enforcement on crossing behavior. Information about any of these issues could provide policymakers with invaluable assistance in the construction and operation of safe and efficient traffic systems.

References

Brunk H. 1958. On the Estimation of Parameters Restricted by Inequalities. *Annals of Mathematical Statistics* 29: 437-454.

Daganzo C. 1981. Estimation of Gap Acceptance Parameters Within and Across the Population from Direct Roadside Observation. *Transportation Research* 15B: 1-15.

Mahmassani H, Sheffi Y. 1981. Using Gap Sequences to Estimate Gap Acceptance Functions. *Transportation Research* 15B: 143-148.

McCall J. 1970. Economics of Information and Job Search. *Quarterly Journal of Economics* 84: 113-126.

Mukerjee H. 1988. Monotone Nonparametric Regression. *Annals of Statistics* 16: 741-750.

Palamarthy S, Mahmassani H, Machemehl R. 1994. Models of Pedestrian Crossing Behavior at Signalized Intersections. Research Report 1296-1, Center for Transportation Research, University of Texas at Austin.

Tanner J. 1951. The Delay to Pedestrians Crossing a Road. *Biometrika* 38: 383-392.

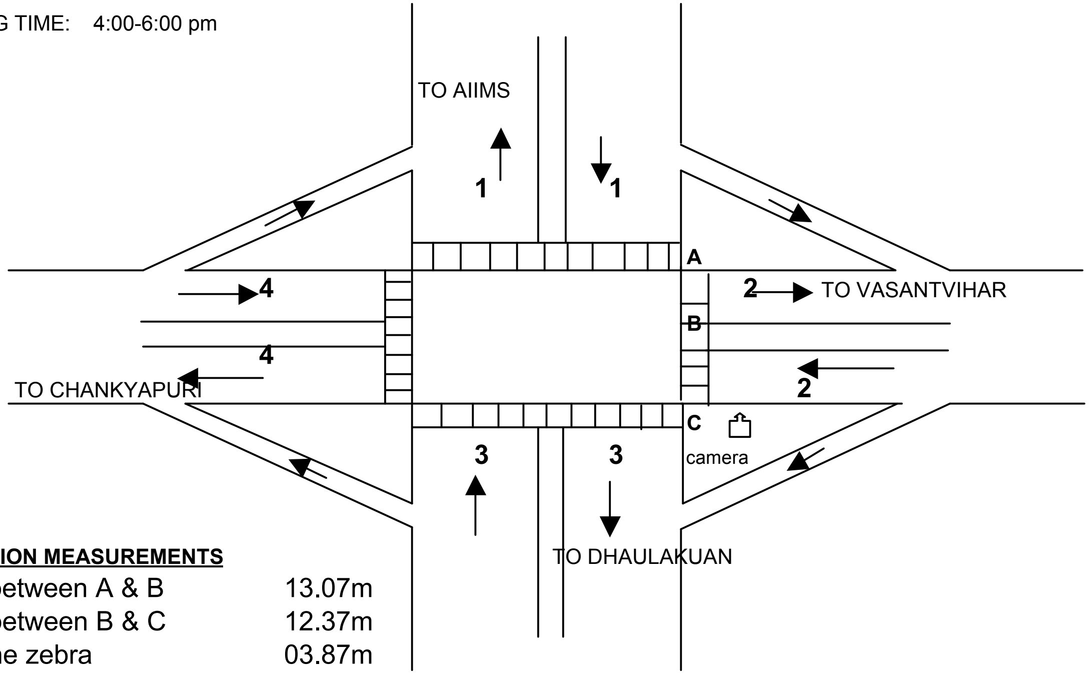
Rust J. 1987. Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica* 55: 999-1033.

Wolpin K. 1987. Estimating a Structural Search Model: The Transition from School to Work. *Econometrica* 55: 801-817.

FIGURE 1: MOTIBAGH RING ROAD CROSSING

RECORDING DATE: 03/03/99

RECORDING TIME: 4:00-6:00 pm



INTERSECTION MEASUREMENTS

distance between A & B	13.07m
distance between B & C	12.37m
width of the zebra	03.87m
width of the median	03.15m

PHASE DIAGRAM

Signal	Duration
31-13-1	01 min
31-13-2	04 sec
31-13-3	01 sec
32-14-1	23 sec
32-14-2	01 sec
32-14-3	01 sec
24-42-1	45 sec
24-42-2	04 sec
24-42-3	01 sec
21-43-1	30 sec
21-43-2	01 sec
21-43-3	01 sec

TIME FOR ONE CYCLE: 2 min 52 sec

Note: Description for signal (a-b-s)

a = direction of traffic flow 1

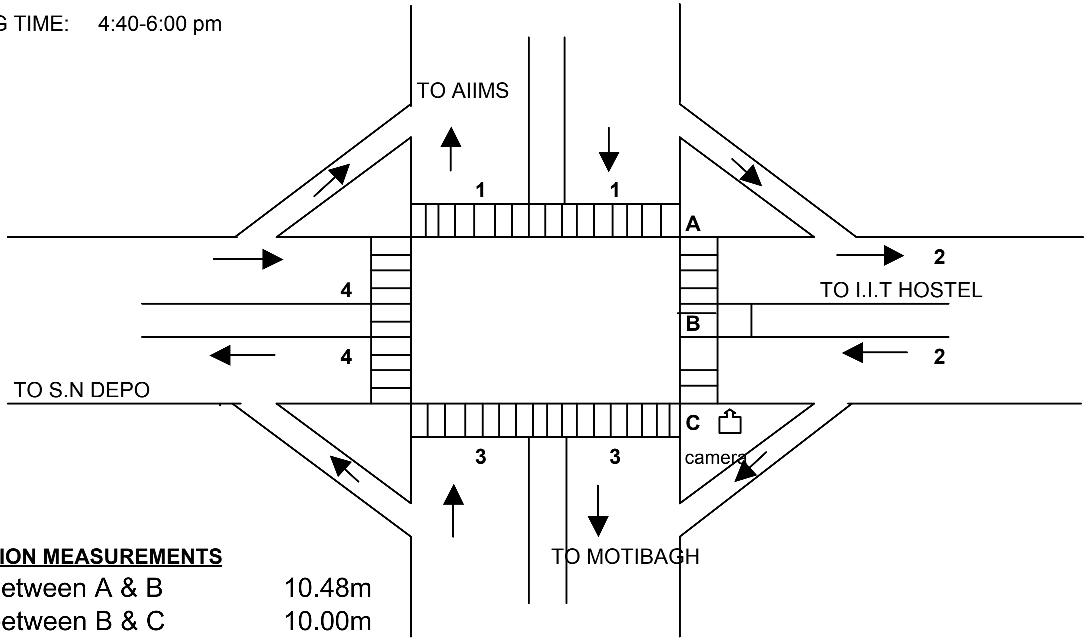
b = direction of traffic flow 2

s = signal color with 1=green,2=yellow,3-red

FIGURE 2: AFRICA AVENUE RING ROAD CROSSING

RECORDING DATE: 08/04/99

RECORDING TIME: 4:40-6:00 pm



INTERSECTION MEASUREMENTS

distance between A & B	10.48m
distance between B & C	10.00m
width of the zebra	03.56m
width of the median	01.52m

PHASE DIAGRAM 1

Signal	Duration
14-13-1	28 sec
13-31-1	48 sec
31-32-1	34 sec
42-43-1	45 sec
42-43-2	04 sec
42-43-3	01 sec
24-21-1	45 sec
24-21-2	04 sec
24-21-3	01 sec

TIME FOR ONE CYCLE: 3 min 30 sec

APPLICABLE TIME: 4:40-5:25:38

PHASE DIAGRAM 2

Signal	Duration
14-13-1	40 sec
13-31-1	39 sec
31-32-1	39 sec
42-43-1	55 sec
42-43-2	04 sec
42-43-3	01 sec
24-21-1	47 sec
24-21-2	04 sec
24-21-3	01 sec

TIME FOR ONE CYCLE: 3 min 50 sec

APPLICABLE TIME: 5:25:38-6:00

Note: Description for signal (a-b-s)

a = direction of traffic 1

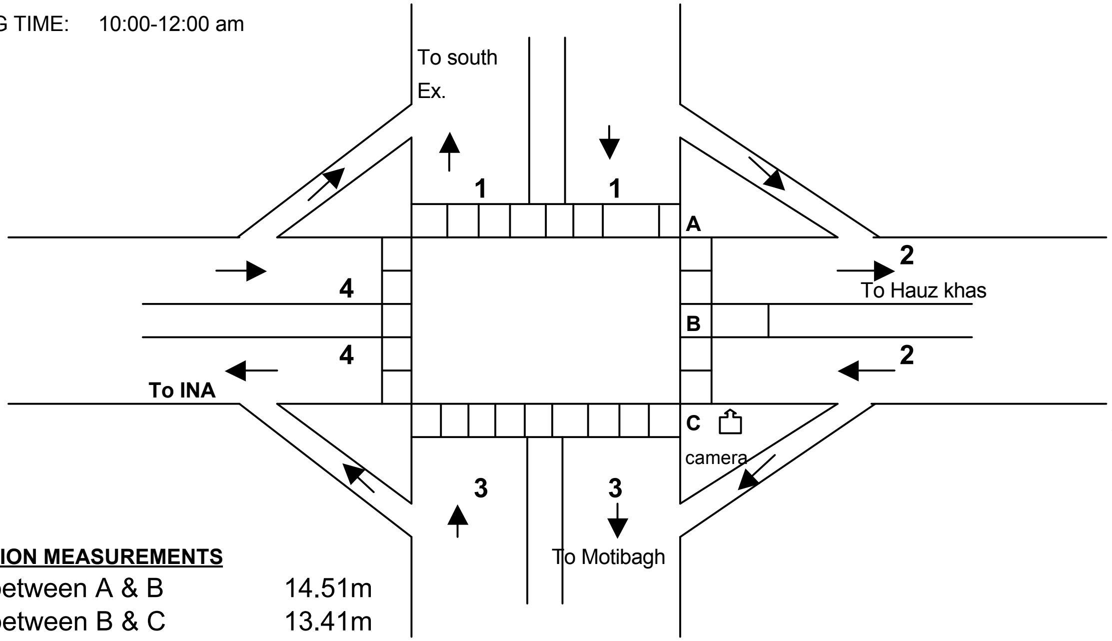
b = direction of traffic 2

s = signal color with 1=green,2=yellow,3-red

FIGURE 3: AIIMS RING ROAD CROSSING

RECORDING DATE: 09/09/99

RECORDING TIME: 10:00-12:00 am



INTERSECTION MEASUREMENTS

distance between A & B	14.51m
distance between B & C	13.41m
width of the zebra	02.37m
width of the median	01.49m

PHASE DIAGRAM 1

Signal	Duration
13-31-1	60 sec
13-31-2	04 sec
13-31-3	01 sec
14-32-1	34 sec
14-32-3	01 sec
24-42-1	60 sec
24-42-2	04 sec
24-42-3	01 sec
21-43-1	24 sec
21-43-3	01 sec

TIME FOR ONE CYCLE: 3 min 10 sec

APPLICABLE TIME: 10:00-10:22:01

PHASE DIAGRAM 2

Signal	Duration
13-31-1	55 sec
13-31-2	04 sec
13-31-3	01 sec
14-32-1	34 sec
14-32-3	01 sec
24-42-1	55 sec
24-42-2	04 sec
24-42-3	01 sec
21-43-1	24 sec
21-43-3	01 sec

TIME FOR ONE CYCLE: 3 min

APPLICABLE TIME: 10:22:01-12:00

Note: Description for signal (a-b-s)

a = direction of traffic flow 1

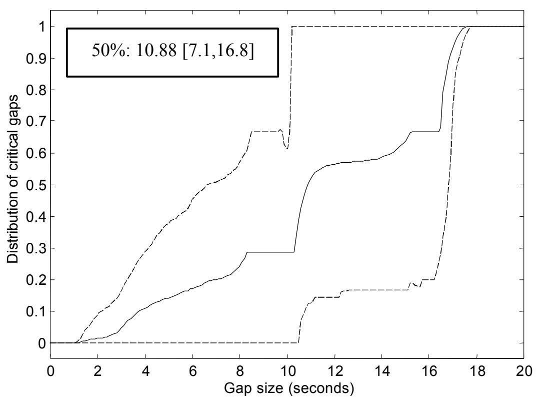
b = direction of traffic flow 2

s = signal color with 1=green,2=yellow,3-red

Figure 4: Nonparametric Estimates of Critical Gap Distributions for Young Adults by Intersection and Decision Point

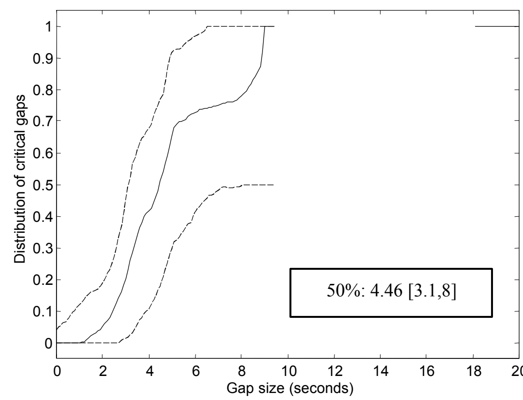
A. Motibagh

Origin Decision Point



N = 33, bandwidth = 3.5076

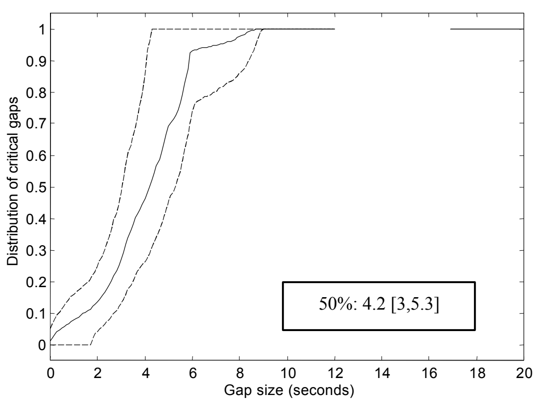
Median Decision Point



N = 47, bandwidth = 1.9306

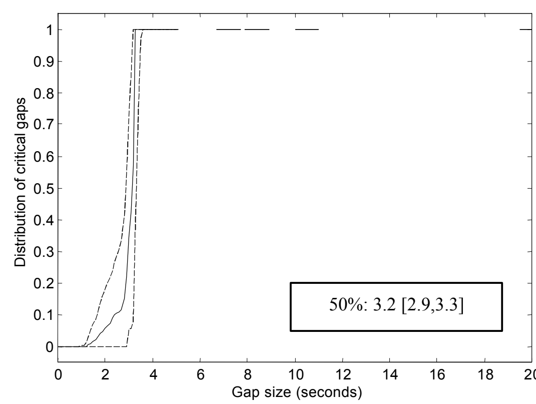
B. Africa Ave

Origin Decision Point



N = 73, bandwidth = 3.1483

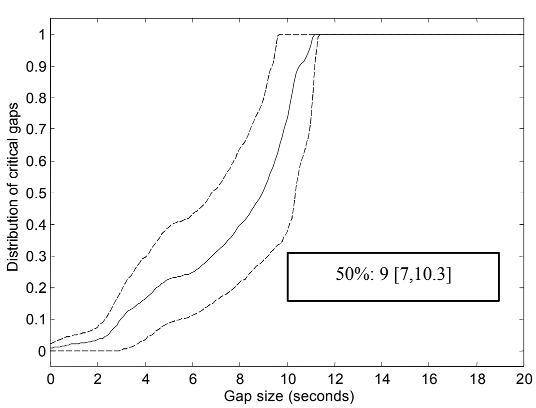
Median Decision Point



N = 127, bandwidth = 0.55705

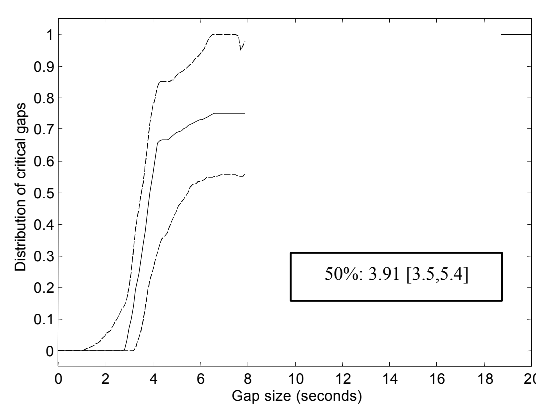
C. AIIMS

Origin Decision Point



N = 109, bandwidth = 2.7788

Median Decision Point



N = 133, bandwidth = 1.3347

Note: Mukerjee's (1988) smoothed monotonic estimates using Epanechnikov kernel and Silverman's rule of thumb bandwidth. Dashed lines are bootstrapped 95% confidence intervals. 50% refers to estimated median critial gap (95% confidence interval in brackets).

Table 1: Pedestrian Characteristics by Intersection

		Intersection			
		Motibagh	Africa Avenue	AIIMS	All
All observations	**Complete crossing**	125	379	429	989
	Partial crossing	32	99	211	350
	Total	157	478	640	1275
Gender	**Male**	144	406	509	1059
	Female	13	72	131	216
Age	**Child**	2	6	25	33
	Young Adult	104	343	475	922
	Middle Age	43	124	98	265
	Old	8	5	42	55
Handicapped		2	0	3	5
With Children		2	4	40	46
With Heavy Packs		2	1	13	16

Table 2: Pedestrians Arrivals and Waiting Times by Intersection and Decision Point

			Intersection and Decision Point					
			Motibagh		Africa Ave		AIIMS	
Type of Arrival			**Origin**	**Median**	**Origin**	**Median**	**Origin**	**Median**
All arrivals	**Number**		129	148	391	466	491	621
Arrivals at "safe" signal	**Total number**		82	77	299	287	357	447
	Number waiting		1	6	5	2	8	7
Arrivals at "unsafe" signal	**Total number**		47	71	92	179	134	174
	Number waiting		24	39	48	143	76	115
	Waiting time	**mean**	11.08	7.77	15.93	19.89	21.04	30.34
		std dev	7.43	5.9	14.08	16.79	24.03	30.06
		max	26	30.04	64.96	71.4	92.36	97.48
		90th	22.68	12.04	39.76	41.34	62.27	79.36
		75th	15.18	8.68	22.64	32.39	24.56	55.54
		median	9.8	7.12	10.52	14.72	11.4	15.92
		25th	6.2	5.62	5.28	5.57	6.14	6.32
		10th	0.48	1.1	3.16	2.19	2.45	3.2
		min	0.08	0.04	1.8	0.24	0.52	0.36
	Number facing unobserved initial gap		17	18	26	14	37	35

Note: Waiting time summary statistics are conditional on waiting a positive length of time before accepting a gap. Unobserved initial gap is initial gap larger than 20 seconds. All unobserved gaps are accepted.

Table 3: Probit Estimates of the Critical Gap Distribution, All Intersections and Demographic Groups

Parameter	Gap Data: Initial gaps	Gap Data: All gaps
a	1.3879	1.4757
	(0.0969)	(0.0658)
b_{female}	0.0369	0.0819
	(0.133)	(0.0842)
b_{child}	-0.0355	-0.1825
	(0.4168)	(0.2789)
$b_{mid\ age}$	-0.0221	0.0334
	(0.1096)	(0.0777)
b_{old}	-0.181	0.3237
	(0.2402)	(0.1587)
g_{median}	-0.2958	-0.1764
	(0.0878)	(0.0628)
g_{AIIMS}	0.4698	0.2631
	(0.1009)	(0.0691)
$g_{Motibagh}$	0.5749	0.2739
	(0.1223)	(0.0874)
s	0.459	0.4779
	(0.0445)	(0.0261)
Loglik	-118.07	-292.66
N	697	678

Note: Asymptotic standard errors in parentheses.

Table 4: Estimated Percentiles of Critical Gap Distribution for Young Adults by Intersection and Decision Point

		Intersection and Decision Point					
		Motibagh		AIIMS		Africa Ave	
Specification	**Percentile**	**Origin**	**Median**	**Origin**	**Median**	**Origin**	**Median**
Parametric (Initial Gaps)	**25th**	5.23	3.89	4.71	3.51	2.95	2.19
		[4.07,6.39]	[3.07,4.72]	[3.93,5.5]	[2.9,4.12]	[2.39,3.51]	[1.79,2.59]
	50th	7.12	5.30	6.41	4.77	4.01	2.98
		[5.62,8.62]	[4.2,6.39]	[5.41,7.41]	[3.96,5.58]	[3.25,4.77]	[2.42,3.54]
	75th	9.68	7.20	8.72	6.48	5.45	4.05
		[7.59,11.78]	[5.64,8.77]	[7.26,10.17]	[5.29,7.68]	[4.32,6.58]	[3.2,4.91]
Parametric (All Gaps)	**25th**	4.18	3.50	4.13	3.46	3.18	2.66
		[3.47,4.88]	[2.95,4.05]	[3.62,4.64]	[3.03,3.89]	[2.76,3.59]	[2.36,2.96]
	50th	5.75	4.82	5.69	4.77	4.37	3.67
		[4.82,6.69]	[4.08,5.56]	[5.02,6.36]	[4.19,5.35]	[3.81,4.94]	[3.25,4.09]
	75th	7.92	6.64	7.84	6.57	6.02	5.05
		[6.63,9.22]	[5.6,7.68]	[6.89,8.78]	[5.73,7.41]	[5.2,6.85]	[4.41,5.69]
Nonparametric	**25th**	8.09	3.17	6.08	3.41	2.9	2.93
		[3.73,16.53]	[2.5,4.71]	[3.44,8.66]	[3.09,3.95]	[2.05,3.94]	[2.43,3.27]
	50th	10.88	4.46	8.97	3.91	4.19	3.16
		[7.15,16.81]	[3.13,8.02]	[6.99,10.25]	[3.47,5.44]	[3.01,5.27]	[2.9,3.32]
	75th	16.55	6.99	10.03	6.7	5.43	3.25
		[10.14,17.01]	[4.4,NaN]	[8.77,11.02]	[3.98,NaN]	[3.79,6.06]	[3.06,3.41]

Note: 95% confidence intervals in parentheses.

Table 5: Simulated Light Sequences

Signal	Safe Directions for Crossing A-B/B-A	Safe Directions for Crossing C-B/B-C	Signal Lengths (seconds) Sequence 1	Sequence 2	Sequence 3	Sequence 4
14-13	Y	Y	-	-	-	28
13-31	Y	Y	60	60	40	48
14-32	N	Y	35	-	25	-
24-42	N	N	60	60	40	-
21-43	Y	N	25	-	15	-
31-32	N	Y	-	-	-	34
42-43	N	Y	-	-	-	50
24-21	Y	N	-	-	-	50
Total sequence length			180	120	120	210

Table 6: Simulated Crossing Outcomes under Different Light Sequences

A. Crossing from A to C

	Light sequence 1		Light sequence 2		Light sequence 3		Light sequence 4	
	Origin	**Median**	**Origin**	**Median**	**Origin**	**Median**	**Origin**	**Median**
All Arrivals								
Percentage arriving under unsafe signal	52.74%	55.95%	49.68%	28.45%	54.03%	54.06%	40.35%	39.30%
Unsafe Arrivals								
Percentage waiting	87.72%	73.98%	87.51%	74.49%	87.89%	74.08%	87.21%	74.09%
Percentage waiting until safe signal	38.22%	0.41%	47.06%	0.37%	45.49%	0.73%	40.39%	0.32%
Average waiting time conditional on waiting	26.01	11.85	19.13	11.71	20.11	11.66	24.06	11.95
Percentage crossing at gap <2 seconds	1.34%	0.31%	1.20%	0.29%	1.18%	0.29%	1.22%	0.28%
Percentage crossing at gap <4 seconds	2.86%	1.56%	2.61%	1.55%	2.59%	1.54%	2.68%	1.65%

B. Crossing from C to A

	Light sequence 1		Light sequence 2		Light sequence 3		Light sequence 4	
	Origin	**Median**	**Origin**	**Median**	**Origin**	**Median**	**Origin**	**Median**
All Arrivals								
Percentage arriving under unsafe signal	47.22%	38.62%	49.71%	28.58%	45.64%	39.15%	23.77%	40.19%
Unsafe Arrivals								
Percentage waiting	87.43%	74.55%	87.73%	73.48%	87.79%	73.55%	87.53%	75.06%
Percentage waiting until safe signal	40.54%	0.20%	46.99%	0.36%	48.53%	0.26%	50.14%	0.21%
Average waiting time conditional on waiting	24.23	11.47	18.99	11.37	17.86	11.21	16.69	11.25
Percentage crossing at gap <2 seconds	1.17%	0.26%	0.95%	0.28%	0.95%	0.24%	1.14%	0.21%
Percentage crossing at gap <4 seconds	2.73%	1.50%	2.40%	1.61%	2.42%	1.60%	2.54%	1.45%

Table 7: Additional Probit Estimates of the Critical Gap Distribution, All Intersections and Demographic Groups

Parameter	(1)	(2)	(3)
a	0.9483	0.8916	1.4802
	(0.1581)	(0.1641)	(0.0703)
b_{female}	0.1197	0.1237	0.0654
	(0.1457)	(0.1458)	(0.0783)
b_{child}	-0.0145	-0.0173	-0.2147
	(0.453)	(0.4516)	(0.2328)
$b_{mid\ age}$	-0.0731	-0.0699	0.068
	(0.122)	(0.122)	(0.0715)
b_{old}	-0.0804	-0.0776	0.3488
	(0.2611)	(0.2608)	(0.1547)
g_{median}	-0.2964	-0.2976	-0.1948
	(0.094)	(0.0939)	(0.0663)
g_{AIIMS}	0.4357	0.4409	0.2275
	(0.1104)	(0.1102)	(0.0644)
$g_{Motibagh}$	0.5062	0.508	0.2465
	(0.1317)	(0.1315)	(0.0796)
$g_{Veh\ Length}$	0.1445	0.1726	
	(0.041)	(0.0461)	
$g_{Veh\ Length\ Sqd}$		-0.0028	
		(0.001)	
s	0.4594	0.4582	
	(0.0534)	(0.0534)	
s_{origin}			0.5839
			(0.0441)
s_{median}			0.369
			(0.0305)
r			-0.0234
			(0.2347)
Loglik	-102.31	-102.07	-284.30
N	550	550	571

Note: Asymptotic standard errors in parentheses. Columns 1 and 2 use only initial gap data, exclude observations with censored initial gaps, and use the sum of vehicles in a platoon for platoon length. Column 3 allows correlation in critical gaps and different variances across decision points. Column 3 uses all gap data.