How Do Police Use Race in Traffic Stops and Searches?

Tests Based on Observability of Race

Joseph A. Ritter[a]

February 2017

Forthcoming: *Journal of Economic Behavior Organization*

DOI: http://dx.doi.org/10.1016/j.jebo.2017.02.005

[a]Department of Applied Economics, University of Minnesota, 231 Ruttan Hall, St. Paul, MN 55108, United States. jritter@umn.edu. The Center for Urban and Regional Affairs at the University of Minnesota provided partial support for this research. I am especially grateful to Samuel Myers, Jr. and Aaron Sojourner who carefully read multiple drafts. Tim Beatty, Marc Bellemare, Juan Chaparro, Janna Johnson, seminar participants at Iowa State University, and reviewers all provided helpful discussions and feedback.

Abstract

When a police officer decides whether to initiate a traffic stop, the driver's race is less likely to be known during darkness, but always observed after the stop takes place. If officers use information optimally, this flow of information about race leads to specific empirical predictions, which are tested using data on traffic stops in Minneapolis. The prediction about stops is supported, but those concerning searches are not. This pattern of results implies that police choices were inconsistent, which is evidence against both statistical discrimination and optimizing with a taste for discrimination. The results may reflect cognitive biases present in the time-sensitive decision to initiate a stop.

Keywords: racial bias, traffic stop, search, discrimination, police

JEL: J15, K14, K42

1 Introduction

There are large racial disparities in traffic stops, both in the rate at which drivers are stopped and in the probability of a search subsequent to a stop. The use of race in policing is contrary to public policy in the United States and illegal in many jurisdictions.[1] Attributing these disparities to race per se—establishing a racial bias—is difficult, however. Does race affect decisions about people who are—to a police officer—otherwise observationally identical? This question can be asked about either traffic-stop decision point, but in economics they have been disconnected. Most analyses of search rates have assumed that police use race in their decisions and have developed methods to distinguish between reasons for bias, specifically, statistical discrimination and prejudice (Knowles, Persico and Todd, 2001; Hernández-Murillo and Knowles, 2004; Anwar and Fang, 2006; Close and Mason, 2006; Antonovics and Knight, 2009). The sparser literature about stops has been concerned with testing for racial bias without regard to causes (Jeffrey Grogger and Greg Ridgeway, 2006—GR henceforth—and Horrace and Rohlin, 2016).[2]

This paper brings together the analysis of stops and searches by developing and empirically testing a model of optimizing police officers—police officers as *homo economicus*—that makes separate empirical predictions about stops and searches based on the information available at the two decision points. The analysis relies on two simple behavioral assumptions. First, police officers maximize expected net benefits from stops and searches. Second, race is payoff-relevant information for police officers, either because of law enforcement benefits or taste for discrimination or both.

The empirical predictions of the model are driven by two features of the flow of information about race. First, during darkness race is less likely than during daylight to be known when the officer decides whether to initiate a stop. This is the fact that GR leverage to identify racial bias in stops. The second, which I use

[1]As of 2014, 30 states had some form of racial profiling statute (NAACP, 2014).

[2]Part of the reason for the disconnect is that any study of racial bias in traffic stops and searches faces a nearly intractable statistical challenge: only a tiny fraction of what a police officer observes is recorded in the data. Thus all of the cited papers rely on strong identifying assumptions of various sorts.

here to extend GR's methodology to searches, is that race can be easily observed after the stop takes place, so that during darkness stops are more likely to reveal information about race, which can then be used in the search decision.

Statistical identification also relies on GR's insight that in urban areas driving patterns are more tied to clock time than to whether it is light or dark. Thus driving and policing patterns at, say, 6:30 P.M. are essentially the same whether it is June (still light) or December (already dark). By restricting the sample to stops that occur between the times of the earliest and latest sunset of the year, which they term the intertwilight period, GR argue that they isolate exogenous variation in the rate at which police officers can discern the race of drivers. Since this assumption is pivotal, it is discussed extensively in section 3.1 and robustness checks are reported in section 6.1, including restricting the analysis to stops near switches to and from daylight saving time.

Section 3 shows that these behavioral and identification assumptions lead to GR's prediction (the proportion of stopped drivers who are black is lower during darkness) and two additional predictions about discretionary search rates during the intertwilight period: (i) Search rates for black drivers will be higher during darkness than during daylight. (ii) Search rates for white drivers will be lower during darkness. The search-rate predictions are generated by the fact that a higher fraction of drivers of unknown race are stopped during darkness, and the implied selection changes in opposite directions for white and black drivers. The predictions about search rates are a direct consequence of optimizing police officers using race in the decision to stop a vehicle.

There may at first seem to be little point in looking for racial bias in two separate decisions made by the same person, but consistency in decision making is the hallmark of optimization in economics. The model highlights the fact that the stop and search decisions are not separable, so testing both predictions can, on the one hand, strengthen conclusions about presence or absence of racial bias or, on the other hand, reveal behavioral inconsistencies that a single test cannot. In fact, applied to data collected during 2002 by the Minneapolis Police Department, the tests generate results that imply inconsistent decision making. There is strong evidence that police use race in stops—the fraction of stopped white or black drivers who are black is more than 5 percentage points lower during darkness than daylight.[3] If police used race consistently, the search rate predictions should be supported as well, but there is evidence against them. Since statistical discrimination is an outcome of optimization, the inconsistency is evidence against statistical discrimination. It is also evidence against optimal exercise of a taste for discrimination.

There are three possible reasons for finding this pattern of results. First, the evidence of racial bias in stops could be a false positive because of failure of the identification assumptions; section 6 argues that this is unlikely. The second is that the additional assumptions about information in search decisions are invalid. These assumptions are weak, however, so the most plausible interpretation is that police officers' behavior was, in fact, inconsistent (section 7). The possibility that police officers do not behave as *homo economicus* is likely to have important implications, not only for understanding their choices, but also for designing interventions to mitigate the pernicious effects of racial bias.

One explanation for inconsistent decisions, which adds structure to the deviation from optimality, is that cognitive processing of information involved in a police officer's decision to stop a motorist is different than that employed in the decision to search a motorist who has already been stopped. Although this hypothesis is necessarily speculative in the present context, research on decision processes in other contexts reveals that outcomes often differ sharply between decisions made rapidly on the basis of intuition and those to which conscious thought has been applied (Kahneman, 2011). Specifically, the stop decision is taken under considerable time pressure and is based on much less specific information (beyond an observed traffic violation). In the metaphor used by Kahneman (2011), "System 1" is active in the stop decision, while the search decision can be processed by "System 2." System 1 employs heuristics that can lead to bad judgements in some contexts.

In the context of processing information about race, this bifurcation of cogni-

[3]The traffic stop data used in this paper were collected by the Minneapolis Police Department during 2002 because of contemporary concern about racial profiling in Minnesota. The Department subsequently instituted efforts to avoid profiling, some of which were outlined in a federal mediation agreement in December 2003. There was also a change in Department leadership at the beginning of 2004. More recent data would be required to evaluate the impacts of these changes.

tive processes suggests that the circumstances of the stop decision are more likely to trigger implicit discrimination (Bertrand, Chugh, and Mullainathan, 2005). This hypothesis would imply that subconscious racial bias can be offset in the deliberative process following the stop so that evidence of racial bias would tend not to be evident in searches, even if present in stops, as found here. Second, it would suggest that training with a narrow focus on avoiding the use of heuristics could be more effective than training with a general goal of making officers aware of bias in their decisions.

The paper proceeds as follows. Section 2 discusses GR's key identifying assumption in more detail. Section 3 analyzes how darkness affects stop and search decisions when race enters into officers' objective function. Section 4 describes the data in and discusses issues about how to define a discretionary search. Section 5 tests the predictions about stops and searches, and section 6 considers robustness to violations of assumptions. Section 7 considers implications and interpretations of the empirical results.

2 Grogger and Ridgeway's empirical strategy

Grogger and Ridgeway's approach to detecting racial bias in stops (which I adopt) starts with the observation that recognizing the race of a driver is more difficult in darkness than in daylight.[4] A simple day-night comparison is insufficient, however, because driving and policing patterns differ throughout the day. There is no reason to think these variations are independent of race. Moreover, data on these patterns are not generally available and expensive to collect.

GR refine the day-night comparison by noting that there are times of day that are before sunset during part of the year and after sunset during part of the year (also between earliest and latest sunrise, but there are many fewer traffic stops early in the morning). They argue that in urban areas driving and policing patterns are more tied to the clock than to whether it is dark, so that the populations of drivers at risk of being stopped as well as how they

[4] GR cite evidence that rates of recognizing race are rather high during daylight and very low during darkness.

are observed by police officers are approximately the same at a given time of day whether it is dark or not. If this identification assumption is correct, the darkness-daylight contrast generates variation in the ability of police officers to know the race of a driver before initiating a traffic stop that is not correlated with unobservable determinants of stops. Thus they propose discarding stops that did not happen between the time of earliest sunset and latest sunset; they refer to this as the "intertwilight period." (In Minneapolis the evening intertwilight period is roughly between 5 and 9 P.M.)

Given this source of variation, the empirical question is: How does the the experience of black drivers change relative to white drivers when it becomes dark? GR assume that, if there is a bias against black drivers, the fraction of drivers stopped who are black should be higher when race is observed than when race is not observed (that is, the prediction is not based on an explicit model of police behavior). They show that this prediction carries over to an imperfect indicator of whether race is observed, namely whether it is dark, though the measured effect is biased downward. Thus, the method identifies only the sign of bias.

For essentially the same reason, the test can have low power; an observation carries no information if the officer could not identify the driver's race ex ante.[5] Also, as GR point out, if police are able to infer the driver's race from characteristics of the vehicle that are visible in darkness, the power of the test will be lower (because the race recognition gap between daylight and darkness is reduced).

A related concern is that police may use "vehicle profiling"—statistical discrimination based on a correlation between characteristics of vehicles (which may be less visible in darkness) and characteristics of drivers. To the extent that vehicle profiling is an indirect way of identifying black drivers, it strengthens the identification of racial bias. But if, for example, it is used to identify drivers with low socioeconomic status, which is correlated with race, GR's test for racial bias could create false positives. This concern is discussed further in section 6.

[5]Horrace and Rohlin (2016) use detailed information about the locations of streetlights relative to stops to increase the power of the test. Precise location information is not available in the data used here.

3 Predictions about stops and searches

If police officers systematically use information about race, they have two opportunities to do so: in deciding whether to stop a vehicle and in deciding whether to initiate a search. This section develops predictions about how decisions of an optimizing officer are affected by how darkness alters the flow of information about race. It provides an explicit decision-theoretic foundation for GR's stop rate test and extends the logic to testable predictions about searches.

The overall decision structure and information flow are intended to mimic that which occurs in an encounter between a driver and a police officer. The officer first observes some characteristics and behaviors of a driver and, sometimes, race. The officer then decides whether or not to stop the vehicle. If a stop takes place, the officer observes the driver's race as well as additional characteristics and behaviors, then decides whether to search.

Prior to a stop, the officer observes a set of non-race characteristics and behaviors x (shortened to "characteristics" in the sequel). This generally includes a traffic violation. The officer also observes $R \in \{w, b, u\}$, an indicator of whether the driver's race is white (w), black (b), or unknown (u). For simplicity I assume that officers do not make errors (e.g., identifying a w driver as b) or probabilistic judgements of race (for example, "The probability that driver is white is 0.7.").[6] The latter is inconsequential, as explained below.

Let $\ell = 1$ indicate daylight and $\ell = 0$ darkness. Race is observed with probability δ during daylight and η during darkness, i.e., $P(R = u|\ell = 1) = 1-\delta$ and $P(R = u|\ell = 0) = 1 - \eta$. I assume that, conditional on ℓ, whether race is observed and x are independent. Until stated otherwise, time of day (t) is held constant, and I suppress this time notation until needed.

After a stop takes place, the officer observes additional non-race characteristics and race. Denote all characteristics known after the stop by x^* (to be clear,

[6]The statement that officers do not make probabilistic judgements refers to direct observation of race. Non-race characteristics may contain information correlated with race and, in any case, police officers would have priors based on population proportions.

$x \subset x^*$), and denote the post-stop observation of race by $R^* \in \{b, w\}$. I will also use R^* to denote the actual race of an individual.

The police receive benefits from traffic enforcement and from criminal law enforcement, but the process of stopping and searching a motorist whom they have observed also has psychological costs or benefits, potential legal costs, and opportunity costs (such as foregone traffic enforcement). For analytical convenience I separate these into three parts. The first part is the benefit stemming directly from an observed traffic violation (e.g., issuing a speeding ticket), which I denote by A. A is a random variable in the sense that officers encounter motorists randomly, but for simplicity, I assume these benefits are known prior to the stop and are independent of whether race is observed (though not necessarily of race). A does not play a large role in the sequel, but highlights the role that traffic enforcement plays in the law enforcement aspects of police decisions.

The second component, S, is the net of all other benefits and costs of stopping an observed motorist, including the cost of the stop, but excluding expected costs and benefits of a potential search. For instance, S could represent the realized benefit of stopping motorist suspected of being intoxicated or, if there is a taste for discrimination, the psychological benefit of stopping a black driver. The marginal value of foregoing the stop is normalized to zero.

The third component, S^*, is the net benefit of a search. The marginal value of foregoing a search (returning to patrol after issuing a citation or warning) is also normalized to zero. It is not necessary to assume that A, S or S^* is independent of race, time of day, or location. It is also not necessary to reduce S^* to a binary success measure, as papers focusing on hit rates implicitly do.

Letting $s = 1$ denote that a stop takes place with $s = 0$ otherwise, the stop decision is

$$\max_{s\in\{0,1\}} \big\{0, s\big(A + \mathrm{E}[S|x, R] + \max\{0, \mathrm{E}[S^*|x, R]\big)\big\}$$

for $R \in \{b, w, u\}$. After a stop takes place, benefits (or costs) A and S are received, and the officer observes race (R^*) and x^*. Letting h be an indicator for a search, the officer solves

$$\max_{y\in\{0,1\}} \big\{(1-h)(A+S) + h\big(A + S + \mathrm{E}[S^*|x^*, R^*]\big)\big\}\,.$$

I now turn to assumptions that add more structure to the decision problems, giving them empirical content. Let $\mathrm{E}[V|x,R] = \mathrm{E}[S|x,R] + \max\{0, \mathrm{E}[S^*|x,R]\}$; ; this is the expected net benefit of a prospective stop given x and pre-stop information about race.

Assumption 1. *Officers order characteristics such that if x ranks higher than y, $\mathrm{E}[V|x,R] > \mathrm{E}[V|y,R]$ for any R and if x^* ranks higher than y^* then $\mathrm{E}[S^*|x^*,R^*] > \mathrm{E}[S^*|y^*,R^*]$. The orderings meet standard conditions on preference orderings: completeness, transitivity, and reflexivity.*

This assumption means, first, that officers coherently evaluate the non-race characteristics and, second, that these characteristics affect the perceived return in the same direction, regardless of race. For notational convenience, represent the ordering of characteristics by a scalar function $J(x)$ or $J^*(x^*)$.[7] By construction $\mathrm{E}[V|J,R]$ is increasing in J and $\mathrm{E}[S^*|J^*,R^*]$ is increasing in J^*.

The stop decision can be rewritten as

$$\max_{s\in\{0,1\}} \big\{0, s\big(A + \mathrm{E}[V|J,R]\big)\big\}, \tag{P-1}$$

where I assume that $\mathrm{E}[V|J,R] \neq \mathrm{E}[V|R]$ (officers observe relevant information other than race). The search decision is

$$\max_{y\in\{0,1\}} \big\{(1-y)(A+S), y\big(A+S+\mathrm{E}[S^*|J^*,R^*]\big)\big\}. \tag{P-2}$$

A lot of detail is swept into the rather inclusive definitions of S and S^*. However, the only specific assumption about S and S^* that is required for the empirical predictions is that race enters officers' choice problem in a specific way:

Assumption 2. *Race changes the expected return from stops and searches:*

(i) $\mathrm{E}[V|J_1,b] = \mathrm{E}[V|J_2,w]$ *implies* $J_1 < J_2$, *and*

[7] One could instead directly assume "suspicion indexes" J and J^*. However, it is helpful later to recognize that J and J^* summarize an array of characteristics.

(ii) $\mathrm{E}[S^*|J_1^*, b] = \mathrm{E}[S^*|J_2^*, w]$ *implies* $J_1^* < J_2^*$.

Note that J affects the return from a stop only through its influence on the return from a possible search, which cannot take place unless the stop does. Assumption 2 does not say *why* expected returns differ by race, only that when officers observe a driver's race, their (discrete) choices may be affected. The results in the remainder of this section elucidate how this assumption interacts with darkness to affect the outcomes of the stop and search decisions.

3.1 Decision to stop

Let $F(J, A, R, \ell)$ denote the joint distribution of observed characteristics, traffic enforcement benefits, race observation, and light level. The following assumption is sufficient (but not necessary) to rule out the possibility that the distribution of observed characteristics reverses the effect of Assumption 2 in stops.

ASSUMPTION 3. *(i) For any fixed* A *and* ℓ, $F(J|A, w, \ell) \geq F(J|A, b, \ell)$. *(ii) For any fixed* J *and* ℓ, $F(A|J, w, \ell) \geq F(A|J, b, \ell)$.[8]

To understand the role of this assumption, suppose that the characteristics of black drivers led police officers to think them much less likely on average to be carrying contraband than white drivers. In terms of the net expected benefit of making a stop, that would offset to some extent whatever effect race has through Assumption 2. In principle, the offset could be large enough to reverse any effect of racial bias; black drivers would be treated better than white drivers despite racial bias.

The next assumption is that the distribution of J does not depend on ℓ after conditioning on whether race is observed.

ASSUMPTION 4. $F(J|A, R, \ell = 1) = F(J|A, R, \ell = 0)$ *and* $F(A|J, R, \ell = 1) = F(A|J, R, \ell = 0)$.

[8]These are the inequalities required for first-order stochastic dominance, but I do not assume strict inequality for some J or A, respectively.

The first part of this assumption greatly simplifies notation and results in clearer empirical predictions, but it is difficult to judge whether it is a good approximation to reality. In particular, problems could arise if non-race characteristics used by police are correlated with race and their visibility changes with darkness. Potentially that could either hide evidence of bias or generate findings of bias where none exists. These possibilities are therefore assessed in detail in section 6. Substantively, the assumption is a stand-in for the idea that changes in visibility of non-race characteristics do not overwhelm or, more importantly, mimic the effect of changing visibility of race. The assumption implies that ℓ enters the analysis in the remainder of this section only via its effect on the visibility of race.

The second part, about the distribution of observed traffic violations is easier to evaluate. Driving violations, such as speeding or running red lights are readily visible at night. Equipment and registration violations are less visible in darkness, but can be excluded as a robustness check (section 6).

Consider a police officer's observation of a single motorist at a particular time t. Since $\mathrm{E}[V|J, R]$ is increasing in J, the equation $A + \mathrm{E}[V|J_r(A), R] = 0$ defines $J_b(A)$, $J_w(A)$, and $J_u(A)$, which are the levels of J that trigger a stop when race is observed to be b, w, or u, respectively. Assumption 2 implies that $J_b(A) < J_w(A)$. Conversely, if race does not affect the expected return from stops, $J_b(A) = J_w(A)$. Let $W(t)$ and $B(t)$ be the rates at which officers encounter vehicles driven by white and black drivers, respectively. Then

$$\mathrm{E}[V|J, u] = \frac{W(t)}{W(t) + B(t)}\mathrm{E}[V|J, w] + \frac{B(t)}{W(t) + B(t)}\mathrm{E}[V|J, b]. \tag{1}$$

Thus $J_u(A)$ lies between $J_b(A)$ and $J_w(A)$.[9]

Denote by $P(\text{stop}|R^*, R, A)$ be the probability an R^* driver ($R^* \in \{b, w\}$), observed as $R \in \{b, w, u\}$, is stopped, conditional on A. (Recall that time of day is being held fixed.) The stopping probabilities for drivers of known race

[9]Relaxing the assumption that officers do not make probabilistic judgements about race would change the weights in equation (1), but would not change the conclusion $J_b(A) < J_u(A) < J_w(A)$, which is the key point here.

$(R = R^*)$ are

$$\begin{aligned}
P(\text{stop}|b, b, A) &= \int_{J_b(A)}^{\infty} dF(J|A, b) \\
P(\text{stop}|w, w, A) &= \int_{J_w(A)}^{\infty} dF(J|A, w).
\end{aligned} \tag{2}$$

Assumption 3 and the fact that $J_b(A) < J_w(A)$ imply that[10]

$$P(\text{stop}|b, b, A) > P(\text{stop}|w, w, A). \tag{3}$$

The stopping probabilities for drivers of unknown race $(r = u)$ are

$$\begin{aligned}
P(\text{stop}|w, u, A) &= \int_{J_u(A)}^{\infty} dF(J|A, w) \\
P(\text{stop}|b, u, A) &= \int_{J_u(A)}^{\infty} dF(J|A, b).
\end{aligned} \tag{4}$$

Assumption 3 and the fact that $J_b(A) < J_u(A)$ imply[11]

$$P(\text{stop}|b, b, A) > P(\text{stop}|b, u, A) \tag{5}$$

and similarly, since $J_u(A) < J_w(A)$,

$$\begin{gathered}
P(\text{stop}|w, w, A) < P(\text{stop}|w, u, A) \text{ and} \\
P(\text{stop}|b, u, A) \geq P(\text{stop}|w, u, A).
\end{gathered} \tag{6}$$

Integrating A out of inequalities (3)–(6), we have the following key inequalities:

$$P(\text{stop}|b, b) > P(\text{stop}|b, u) \geq P(\text{stop}|w, u) > P(\text{stop}|w, w). \tag{7}$$

[10]Assumption 3 makes it possible to compare the integrands in (2) to show inequality (3).

[11]This is where Assumption 4 is used in this section. Otherwise, the distributions in (2) and (4) would also depend on whether race is observed as well as actual race. Effectively, the assumption rules out the possibility that the distribution of J is so much worse when race is unobserved that it offsets the disadvantage of being known to be black and/or it is so much better that it offsets the advantage of being known to be white.

(Assumption 3 guarantees that the directions of the inequalities remain the same when A is integrated out.)

Two assumptions about the effect of darkness are used to statistically identify the effect of race on stops.

Assumption 5. *Darkness impedes recognition of race:* $\eta < \delta$.

GR's insight that darkness can be used as exogenous variation in visibility of race during the intertwilight period is embedded in the following assumption.

Assumption 6. *During the intertwilight period, the following do not depend on whether it is dark or light: (i) the distribution of actual race and non-race characteristics of drivers (as distinct from characteristics observed by police, x); (ii) the rates at which police observe vehicles driven by b motorists, $B(t)$, and w motorists, $W(t)$; and (iii) the unconditional distribution of costs of stopping or searching motorists.*

As discussed in section 2, this assumption asserts that comparing outcomes between daylight and darkness adequately controls for the mix of drivers on the road, their vehicles, their driving, and their criminal behavior. Similarly, policing might be more intensive in certain neighborhoods or certain times of day, but this assumption requires only that the pattern is stable during the sample period. It does not require absence of geographic or intra-day variation.

For example, in 1998 the Minneapolis Police Department introduced a crime-reduction strategy called CODEFOR, which uses the spatial distribution of 911 calls and crime reports to allocate police resources (Myers, 2002). Consequently, several areas received high-intensity law enforcement, and more police patrols increased the number of traffic stops. The residents of most of these areas were disproportionately racial minorities, and therefore stopped drivers were as well. However, the existence of the CODEFOR program does not violate Assumption 6 because the program did not appear or disappear during the period under study.

Another aspect of Assumption 6 deserves examination. Taken at face value, it assumes that those with outstanding warrants or those carrying contraband or stolen property do not respond to changes in police actions that result from the difference between daylight and darkness. Other things equal, if black offenders realize their race is more likely to trigger a stop during daylight, they will avoid driving then. Conversely, if white offenders realize they are more likely to be stopped during darkness, they will avoid darkness. It seems unlikely that offenders are completely unresponsive to policing strategies.

Knowles, Persico, and Todd (2001) and others have made the opposite assumption, that offenders respond optimally. This is also unlikely. Although some offenders probably behave strategically, the discussions in Engel and Johnson (2006) and Johnson (2007) of cues used by police suggest that many offenders do not, and that some of their attempts to do so create incongruities that draw police attention. Dominitz (2003) argues that offenders are unlikely to have sufficient information to efficiently neutralize police strategies.

Assumption 6 has the advantage that strategic responses by offenders would attenuate the effects of darkness predicted below, but would not eliminate them entirely if some offenders do not respond optimally. In other words, strategic responses would not create false positive findings of racial bias. It is also important to bear in mind that offenders are a small fraction of the population, and the attenuation would be proportional to that fraction.

Recall that δ and η are the probabilities that race is recognized prior to the stop during daylight and darkness, respectively. Applying the inequalities in (7), the key insight of GR's test is this:

$$\begin{aligned} P(\text{stop}|b, \ell = 1) &= \delta P(\text{stop}|b, b) + (1 - \delta) P(\text{stop}|b, u) \\ &> \eta P(\text{stop}|b, b) + (1 - \eta) P(\text{stop}|b, u) \\ &= P(\text{stop}|b, \ell = 0), \end{aligned} \tag{8}$$

and, similarly,

$$P(\text{stop}|w, \ell = 1) < P(\text{stop}|w, \ell = 0). \tag{9}$$

It follows that at a given time of day the fraction of stopped drivers who are black will be larger during daylight ($\ell = 1$) than darkness.

Inequalities (8) and (9) capture the main idea behind the test, but they hold time of day fixed, so using them empirically requires aggregating them to an interval, T. Define $\tilde{B}(t) = B(t)/\int_{s\in T} B(s)\,ds$ and $\tilde{W}(t) = W(t)/\int_{s\in T} W(s)\,ds$ and recall that $B(t)$ and $W(t)$ do not depend on ℓ. Then

$$\begin{aligned}\frac{P(\text{stop}|b,\ell=1,t\in T)}{P(\text{stop}|w,\ell=1,t\in T)} &= \frac{\int_{t\in T} P(\text{stop}|b,\ell=1,t)\tilde{B}(t)\,dt}{\int_{t\in T} P(\text{stop}|w,\ell=1,t)\tilde{W}(t)\,dt} \\ &> \frac{\int_{t\in T} P(\text{stop}|b,\ell=0,t)\tilde{B}(t)\,dt}{\int_{t\in T} P(\text{stop}|w,\ell=0,t)\tilde{W}(t)\,dt} = \frac{P(\text{stop}|b,\ell=0,t\in T)}{P(\text{stop}|w,\ell=0,t\in T)}.\end{aligned} \tag{10}$$

The center inequality follows from inequalities (8) and (9), which say that the numerator on the left is pointwise greater than the numerator on the right and that the denominator on the left is pointwise less than the denominator on the right. Rearranged, inequality (10) is GR's prediction about stop rates:[12]

PREDICTION 1. *The proportion of drivers stopped during any interval within the intertwilight period who are b is lower when $\ell = 0$ than when $\ell = 1$: $P(R = b|stop, \ell = 0) - P(R = b|stop, \ell = 1) < 0$.*

Before turning to the search decision it is useful to summarize the roles of assumptions in reaching this prediction. Assumptions 1 and 2 say that officers are optimizers and use race in that process. Assumption 3 states conditions that are sufficient to guarantee that the racial bias is not hidden by other factors that favor stopping w drivers. Assumption 6 says the comparison between daylight and darkness will not be confounded by other factors. Assumption 4 rules out the possibility that changes in the distribution of non-race characteristics generate a false positive finding of racial bias.

3.2 Decision to search

In stopping a vehicle, the officer has in essence purchased an option on a search. Once a stop has taken place, the officer observes a broader set of characteristics

[12] GR reach this prediction by assuming $P(\text{stop}|b,b)/P(\text{stop}|w,w) > P(\text{stop}|b,u)/P(\text{stop}|w,u)$ for the entire intertwilight interval and showing that the direction of the inequality is not changed when daylight/darkness is not perfectly correlated with visibility of race.

and behavior, x^*, and the driver's race is always observed. The officer then chooses whether to exercise the option.[13]

In many ways the decision to search is very similar to the stop decision, but there are two key differences. First, race is learned if it wasn't previously known. Second, the drivers about whom search decisions are made are the population of stopped drivers, which is selected differently during darkness than during daylight.

Let $G(J^*)$ be the distribution function of J^*. The analysis requires two assumptions about $G(J^*)$. I assume first that after the stop no relevant characteristics are hidden by darkness.

Assumption 7. *$G(J^*|R^*, \ell = 0) = G(J^*|R^*, \ell = 1)$.*

In practical terms this is a mild assumption since the officer can inspect the vehicle and driver at close range after the stop and can use artificial illumination. Thus the distribution of J^* is affected by ℓ only via the latter's influence on the racial mix of drivers who are stopped (Prediction 1).

Assumption 8. *If $J_2 > J_1$, then $G(J^*|J_2, R^*) < G(J^*|J_1, R^*)$.*

This assumption (which differs from first-order stochastic dominance only in using strict inequality for all J) asserts that the orderings underlying J and J^* are consistent with one another in the following sense: If, say, illegally tinted windows raise suspicion before the stop, then they also shift the distribution of post-stop suspicion to the right—a high J^* is more likely, though post-stop observations of other characteristics could still lead to low J^*. The assumption would be invalid if officers ignored non-race information in making stops. Section 7.1 argues that this conjecture is factually incorrect.

Turning now to analysis of the decision whether to search, note that the solution to problem (P-2) is determined by the sign of $\mathrm{E}[S^*|J^*, R^*]$. Since the

[13]Stricter requirements must be met for a search to be legal than a stop (Harris, 1997), but I assume the the legality of a search is factored into S^*. The option language suggests the possibility of statistical discrimination based on the variance of S, but I do not pursue that possibility.

expectation is increasing in J^*, there are two thresholds J_b^* and J_w^* such that a b driver is searched if $J^* > J_b^*$ and a w driver is searched if $J^* > J_w^*$. Assumption 2 implies that $J_b^* < J_w^*$.

For a given A, the stop decision means that the population of stopped drivers is a mixture of two truncated samples of all drivers. For b drivers it consists of drivers with $J > J_b(A)$ or $J > J_u(A)$, according to whether their race was observed or not prior to the stop. Darkness shifts the mixture toward truncation at $J_u(A)$ and therefore shifts the overall distribution of J for b drivers to the right. Assumption 8 implies that the distribution of J^* shifts right as well, so a higher percentage are searched during darkness.

Although the process just described occurs at the stopping stage, it is driven by concerns about searches because J affects the value of the stop only via a possible search; J enters problem (P-1) only because it affects the expected return from a search. Put differently, police officers use J in decisions about stops only because by initiating a stop they acquire an option to search, which they might exercise after receiving additional information.

Parallel logic for w drivers works in the opposite direction. Darkness again drives the mix toward the distribution truncated at $J_u(A)$, but for w drivers this means a leftward shift of the overall distribution of J. Their distribution of J^* also shifts to the left so that fewer searches are conducted during darkness.

Thus Assumptions 1–8 imply the following predictions about the use of race in search decisions. A more formal derivation can be found in the appendix.

PREDICTION 2. *In any interval during the intertwilight period:*

(i) A higher fraction of stopped b drivers are searched during darkness than during daylight: $P(\text{search}|\text{stop}, R^* = b, \ell = 0) > P(\text{search}|\text{stop}, R^* = b, \ell = 1)$.

(ii) A lower fraction of stopped w drivers are searched during darkness than during daylight: $P(\text{search}|\text{stop}, R^* = w, \ell = 0) < P(\text{search}|\text{stop}, R^* = w, \ell = 1)$.

To put this in terms of the public discourse about race and traffic stops, the politically charged term "pretextual stop" could be interpreted as acquiring an option to search when the observed A and J are relatively low, but the driver is observed to be black. Part (i) of the prediction is an implication of the fact that darkness interferes with this practice. Part (ii) would be the "white privilege" converse.

It may seem surprising that Bayes' Theorem was not used to reach these predictions since the model of the search decision clearly describes a process of Bayesian updating. In fact, though, Prediction 2 is a consequence only of the ordering of the thresholds for stops and searches ($J_b(A) < J_u(A) < J_w(A)$ and $J_b^* < J_w^*$). Put differently, officers do not need to be precise optimizers for the prediction to hold.

4 Data

As part of a state-sponsored study, the Minneapolis Police Department gathered data on every traffic stop that took place during 2002, a total of 53,559 stops.[14] Only stops of black and white drivers are used in the analysis. During the inter-twilight period there were 10,567 of these. Although data were collected in other jurisdictions, I use only Minneapolis to avoid complications arising from mixing jurisdictions. Also, Minneapolis recorded the most stops of any jurisdiction by a factor of four (the St. Paul Police Department did not participate). Each record includes the race of the driver as perceived by the police officer after the stop was completed (i.e., after direct observation of the driver), the time and date, the reason for the stop, and, if a search took place, details about the reason for and outcome of the search. Officer race was not recorded.

All of the data are from reports by the police officer who conducted the stop, but these data were not integrated into the normal record-keeping process of the state or city so there is no way to verify the accuracy of the data. Nor did police cars have video cameras at the time. (In fact cameras were purchased in

[14]See Myers (2002) for a discussion of the social and political context in which the data were collected.

2004 with a grant awarded for participation in the study.) Given the conflicting demands on officers, it seems likely that some stops were not recorded, but there is no reason to think that this creates any statistical bias.

As noted, race of the driver is based the officer's perception, not self-classification by the driver. Drivers' self-classifications are not available in the data, but officers' classifications would be preferred here in any case since officers are the decision makers whose choices are being examined for bias. A separate issue is whether officers accurately recorded the race they believed the driver to be, but there is no evidence of any falsification.

As illustrated by table 1, racial differences in the raw data are stark: 47.0 percent of black or white drivers stopped were black, but only 18.5 percent of the combined driving-age population in 2000 was black.[15] Discretionary searches, defined in two ways, were even more concentrated: fully three-quarters of these searches were of black drivers.

4.1 Intertwilight interval

The evening intertwilight interval is defined as clock times between the earliest and latest sunsets of the year; these are the hours for which it is light at some times of the year and dark at other times. Stops that took place before sunset are categorized as daylight stops. Stops that took place after the end of civil twilight are classified as darkness stops.[16] Stops that took place during civil twilight are excluded in order to sharpen the distinction between daylight and darkness. The morning intertwilight period (only 4.5 percent of intertwilight stops) is defined analogously. Times for sunrise, sunset, beginning of civil twilight, and end of civil twilight are from the U.S. Naval Observatory.[17]

[15] About 16 percent of the 2000 black population was foreign-born according to the Census, predominantly East Africans.

[16] Evening civil twilight is defined as the period between sunset and the time at which the center of the sun is 6 degrees below the horizon, a period which is difficult to define as either daylight or darkness.

[17] http://aa.usno.navy.mil/data/docs/RS_OneYear.php

Table 1: Stops and searches by race
(percent of total)

	White	Black
Driving age population[b]	81.5	18.5
Stops (all hours)	53.0	47.0
Stops during intertwilight interval		
Stops	52.2	47.8
Discretionary searches		
Narrow definition[c]	24.0	76.0
Wide definition[c]	24.7	75.3
Discretionary searches as percent of stops		
Narrow definition[c]	4.4	15.3
Wide definition[c]	7.8	25.8

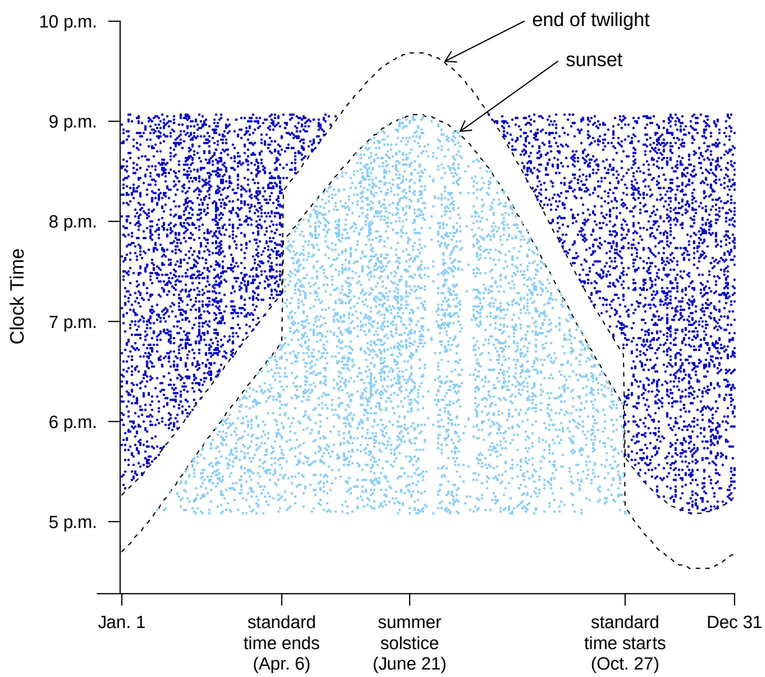
[b]Rescaled from CCJ/IRP (2003), based on 2000 Census.
[c]For definition see section 4.2.

A scatterplot of the dates and times of stops is shown in figure 1. Note that the shift to and from daylight saving time causes discontinuities in the times that separate daylight stops from darkness stops and stretches the intertwilight period. The morning intertwilight period is two hours shorter than the evening period because the discontinuities instead compress the intertwilight interval.

4.2 Discretionary searches

During the data collection, when a driver, vehicle, or passenger was searched, the officer was required to record which type of search took place and one of the following "authority to search" options: verbal or written consent, "observation of contraband," "officer safety," or "incident to arrest." The model applies only to searches the officer *decides* to undertake—discretionary searches. As Hernández-Murillo and Knowles (2004) point out, failing to make this distinction has important consequences. The Minneapolis data has the advantage of being able to separate discretionary and non-discretionary searches. However,

Figure 1: Date and Time of Intertwilight Traffic Stops



Darker points correspond to stops that occurred during darkness. The curved empty band is the civil twilight interval. Stops during civil twilight are excluded from the analysis.

because of the design of the data collection process, there is some ambiguity about whether certain searches were discretionary. I therefore use two definitions of discretionary searches that differ by the inclusion or exclusion of one ambiguous category of searches.

Consent searches and officer safety searches are clearly discretionary. There were only five contraband observed searches during the intertwilight period. None led to an arrest, so I classify them as discretionary. I exclude searches in which only the passenger is searched because the race of passengers was not recorded.

The "incident to arrest" category is where the narrow and broad definitions differ. When an arrest takes place, the police officer is required to search the arrested individual, thus the search is considered "incident to arrest." There are complications, however. If the officer had probable cause to arrest the driver before conducting the search, the search can be legally considered "incident to arrest," even if an arrest does not take place. In State v. Bauman,[18] a Minnesota case that explicated this principle, the situation was described as follows:

> At the time [the officer] asked Bauman to step out of his car so [the officer] could search for identification, [the officer] had accumulated significant information. He knew that the driver had given the name and date of birth of an individual with a valid driver's license, but the registered owner of the vehicle had a suspended license. He also knew that the driver said he did not have his license with him, could not say how old he was, was unsure of his address, was unsure of the purported car owner's address, and could not tell the officer where he was going. On these facts, he had probable cause to arrest Bauman for providing false information to an officer.

In fact, Bauman was not arrested, although the Bauman search was considered "incident to arrest." It was also clearly discretionary, which would be the case whenever an arrest did not take place and the search was undertaken under the legal authority of incident to arrest. Searches that fit this description are, therefore, considered discretionary under both definitions.

[18] State v. Bauman no. C5-98-224, November 24, 1998.

It is easy to imagine that the Bauman search could have turned up contraband or stolen property. In that scenario, an arrest would have taken place, but the search itself would have been no less discretionary. Thus the broader definition of discretionary search includes all searches classified as incident to arrest. Unfortunately, it is impossible to distinguish in the data searches that fit this scenario from searches that took place *because* the driver was arrested—for example, was identified as the suspect in a prior crime.

There is one further complication involving incident to arrest searches. When a vehicle is impounded, a search of the vehicle is mandatory, but impoundment was not an authority to search option on the form. Apparently, for impoundments, officers typically classified the search as "incident to arrest" (probably because this generally implies a mandatory search). Thus, I consider incident to arrest searches in which only a vehicle search took place to be non-discretionary under both definitions.[19]

5 Empirical implementation

5.1 Stops

This section tests Prediction 1 using 2002 data from Minneapolis. These stop rate results closely parallel the report by Ritter and Bael (2009). The test is implemented using logit regressions in which the dependent variable is an indicator of a black driver, and the key independent variable is a dummy for darkness.[20]

Table 2 reports average marginal effects of darkness from regressions in which only stops from the morning and evening intertwilight periods are included. These are estimates of the difference in the probability of a stopped driver being

[19] See CCJ/IRP (2003) for details. Except for the five "contraband observed" searches, my narrow definition of discretionary searches exactly follows that report. The broad definition includes all searches except those that appear to be impoundment searches and seven searches with incomplete data.

[20] Although I follow GR's lead in using logits, results from linear probability models are nearly identical.

Table 2: Stop rate tests

Average marginal effect of darkness	-0.052^{***} (0.010)	-0.068^{***} (0.011)	-0.058^{**} (0.021)
Time-of-day controls	No	Yes	Yes
Month controls	No	No	Yes

Significance: *** = 0.01, ** = 0.05, * = 0.1. N = 10,567 Average marginal effects from logit regressions using stops from morning and evening intertwilight periods. Bootstrap standard errors based on 500 replications. Time-of-day controls are dummies for 16 minute intervals.

black during darkness compared to daylight, $P(b|\text{stop}, \ell = 0) - P(b|\text{stop}, \ell = 1)$. The first column reports estimates that use no controls. The second columns adds dummies for 16-minute intervals of clock time (16 minutes divides both intertwilight periods almost exactly) to allow for possible intra-day changes. The third column also adds seasonal controls in the form of month dummies.

The results indicate that the share of stopped drivers who are black is between 5 and 7 percentage points lower during darkness, consistent with Prediction 1. The effects are quite robust to the inclusion of time-of-day and month controls. There seems to be little doubt that there was racial bias in traffic stops during 2002.

5.2 Searches

The additional implications of Prediction 2 are that darkness raises the probability of discretionary searches of black drivers and lowers the probability for white drivers. Table 3 reports results of regressing an indicator for a discretionary search on an indicator for darkness and controls separately for each race. The average marginal effects shown in the table are estimates of $P(\text{search}|\text{stop}, R^*, \ell = 0) - P(\text{search}|\text{stop}, R^*, \ell = 1)$ for $R^* = b$ and $R^* = w$. Only intertwilight stops not resulting in searches or leading to discretionary searches are included in the regressions. The impacts of darkness are nearly all statistically indistinguishable from zero. For black drivers the signs are opposite those predicted. The marginal effects for white drivers in columns 1 and 2 are statistically significant and their

Table 3: Search rate tests

Sample	Narrow definition			Broad definition		
Black drivers	-0.012	-0.014	-0.015	-0.019	-0.023	-0.044^{*}
	(0.012)	(0.013)	(0.023)	(0.013)	(0.014)	(0.026)
		N = 4,447			N = 4,972	
White drivers	-0.010^{*}	-0.013^{**}	-0.001	-0.007	-0.012	-0.002
	(0.006)	(0.006)	(0.013)	(0.007)	(0.008)	(0.015)
		N = 5,244			N =5,429	
Joint hypotheses (p-values):						
Prediction 2	0.210	0.199	0.189	0.065	0.061	0.034
Both equal zero	0.167	0.082	0.977	0.310	0.122	0.315
Time-of-day controls	No	Yes	Yes	No	Yes	Yes
Month controls	No	No	Yes	No	No	Yes

Significance: *** = 0.01, ** = 0.05, * = 0.1. Dependent variable is a dummy for a discretionary search; sample excludes stops resulting in non-discretionary searches. Average marginal effects of darkness from logit regressions using stops from morning and evening intertwilight periods. Bootstrap standard errors based on 500 replications. Time-of-day controls are dummies for 16 minute intervals. The joint null hypotheses are described in the text.

signs are consistent with predictions. However, these effects are robust to neither inclusion of seasonal controls nor to the definition of discretionary search.

Since the two inequalities in Prediction 2 are manifestations of the same mechanism, it makes sense to test them jointly. The bottom of the table displays results of two tests. The line labeled "both equal zero" reports tests of the null that $\beta_1 = \beta_2 = 0$ in the logit regression[21]

$$\begin{aligned} P(search) = \Lambda(\beta_0 + &\beta_1 black_i dark_i + \beta_2 white_i dark_i \\ &+ black_i controls_i \gamma + white_i controls_i \delta + \varepsilon_i) \end{aligned} \tag{11}$$

Those tests together with the effects reported in the table indicate there is not much evidence in favor of Prediction 2. But is there evidence against it?

The effects of darkness are predicted to be positive for blacks and negative for

[21] This is a convenience for testing and produces the same estimates as separate estimates for black and white drivers.

whites; Prediction 2 does not suggest a specific point null hypothesis. The row labeled "Prediction 2" reports the fraction of bootstrap replicates of equation 11 with $\widehat{\beta}_1 > 0$ and $\widehat{\beta}_2 < 0$ (the region where Prediction 2 is borne out). This can be interpreted as a simple diagnostic or as the p-value of a nonparametric (pairs) bootstrap test of the null hypothesis that $\beta_1 > 0$ and $\beta_2 < 0$.[22]

6 Robustness

Heavy reliance on strong identification assumptions is necessary in all research on race in traffic stops simply because so little information is available (e.g., Knowles, Persico and Todd's assumption that criminals respond optimally to police actions). The core assumptions used here can be summarized simply: Observed changes in the treatment of a given race between daylight and darkness are due to the differential observability of race and are neither masked by nor caused by differences in patterns of driving and policing or differences in the observability of non-race characteristics. This section assesses whether the results of the previous section could be generated by violations of the assumptions about the intertwilight period and how darkness changes the visibility of non-race characteristics.

6.1 Seasonality

Important seasonality of driving or policing would undermine the core idea behind GR's methodology (Assumption 6). Although the estimates in sections 5.1 and 5.2 include seasonality controls, two additional robustness checks were performed. First, unreported results based on excluding days outside the academic year (May 15 through Labor Day) are not materially different than those already reported.

[22]The logic behind this procedure is the same as for a test of single-parameter null $\beta_2 > 0$ in which the p-value is simply the fraction of replicates for which $\widehat{\beta}_2 > 0$, the hypothesis test equivalent of the confidence interval estimated using percentiles of the bootstrap distribution of $\widehat{\beta}_2$. Bootstrapping t-statistics rather than coefficients yields p-values that differ by no more than 0.01 from those in table 3.

Table 4: Stop rate tests near time changes

Average marginal effect of darkness	-0.198^{***} (0.051)	-0.200^{***} (0.051)
Time-of-day controls	No	Yes

Significance: *** = 0.01, ** = 0.05, * = 0.1. $N = 379$ Average marginal effects from logit regressions using stops from morning and evening intertwilight periods. Bootstrap standard errors based on 500 replications. Time-of-day controls are dummies for 16 minute intervals. Time-of-day dummies capture any difference between spring and fall because spring and fall times do not overlap (see Figure 2).

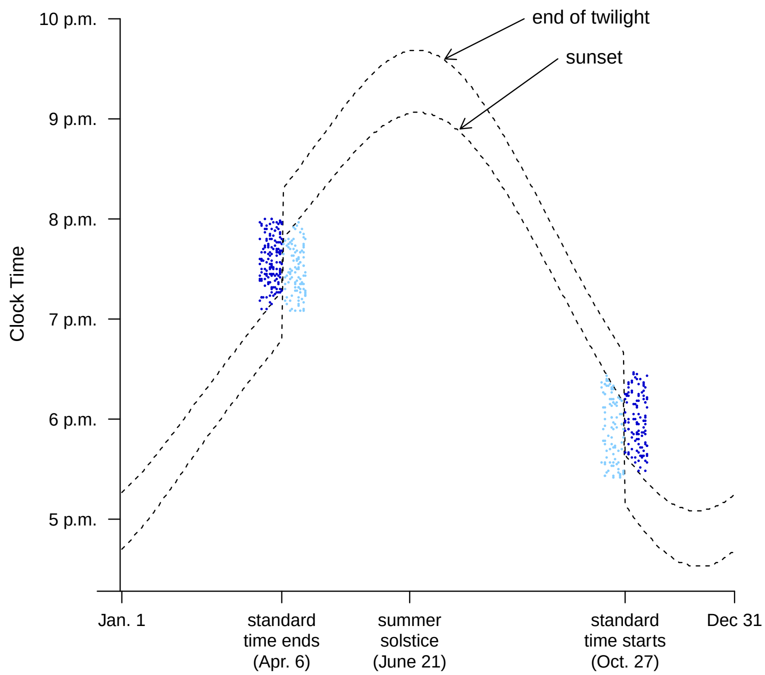
A more draconian approach to removing seasonality is to take advantage of the switches to and from daylight saving time. Near these switches there is a about an hour of clock time that switches between darkness and daylight from one day to the next, and a wider range of clock time that switches over the course of a few days. To take advantage of this situation I restricted the sample to stops that occurred within two weeks of one of the time changes, a total of 379 stops (3.6 percent of the total). The evening stops in the resulting sample are shown in figure 2.

The results for stops, reported in table 4, are generally similar to those reported above. There is clear evidence of racial bias in stops of black drivers. The point estimates are much larger, but also less precise. In table 5 the effects of darkness on searches are imprecisely estimated and mostly insignificant. In this case, all have signs opposite those predicted and, therefore, the joint test rejects Prediction 2 more decisively.

6.2 Visibility of non-race characteristics

As mentioned in section 3.1, the visibility of non-race characteristics and behavior could differ between darkness and daylight (violating Assumption 4). Clearly, the assumption is only an approximation to reality, so it is important to consider the possibility that violations might undermine Prediction 1 (and, therefore,

Figure 2: Date and Time of Traffic P.M. Stops in DST-Switches Sample



Darker points correspond to stops that occurred during darkness. The curved empty band is the civil twilight interval. Stops during civil twilight are excluded from the analysis.

Table 5: Search rate tests near time changes

Controls	Narrow definition		Broad definition	
Black drivers	-0.038	-0.041	-0.124^{**}	-0.163^{**}
	(0.059)	(0.069)	(0.063)	(0.071)
	$N = 143$		$N = 159$	
White drivers	0.010	0.038^{*}	0.062^{*}	0.068^{*}
	(0.020)	(0.022)	(0.036)	(0.037)
	$N = 194$		$N = 207$	
Joint hypotheses (p-values):				
Prediction 2	0.090	0.023	0.007	0.003
Both equal zero	0.826	0.921	0.054	0.015
Time-of-day controls	No	Yes	No	Yes

Significance: *** = 0.01, ** = 0.05, * = 0.1. Dependent variable is a dummy for a discretionary search; sample excludes stops resulting in non-discretionary searches. Average marginal effects of darkness from logit regressions using stops from morning and evening intertwilight periods. Bootstrap standard errors based on 500 replications. Time-of-day controls are dummies for 16 minute intervals. Time-of-day dummies capture any difference between spring and fall because spring and fall times do not overlap (see Figure 2). The joint null hypotheses are described in the text.

Prediction 2). If Assumption 4 does not hold exactly, the probabilities in (8) would depend on ℓ and be different on each side of the inequality. For example, $P(\text{stop}|b, b, \ell = 1) \neq P(\text{stop}(b, b, \ell = 0)$. In principle, the changes could attenuate or reverse Prediction 1 despite the presence of racial bias, or they could cause Prediction 1 to hold in the absence of bias.

In the former case, to reverse the prediction, it is evident from (8) that the changes would need to be large relative to the impact of knowing race on the probability of being stopped (i.e., $P(\text{stop}|R^*, R^*) - P(\text{stop}|R^*, u)$). The visibility of non-race characteristics is likely correlated with visibility of race, which suggests that the visibility changes would be small in absolute value. For example, if race happens to be visible during darkness, other characteristics are probably visible as well. In any case, like GR's concern about low power, this issue is germane only if the data do not support Prediction 1.

The second possibility is more important in the present application. Could changes in the visibility of non-race characteristics and behavior between darkness and daylight (a significant violation of Assumption 4) generate the stop-rate results of the previous section in the absence of racial bias? To provide a tentative answer to this question it is useful to consider characteristics and behavior in four categories: driver characteristics, vehicle characteristics, non-driving behavior, and driving behavior.

A particular behavior or characteristic would need to meet four criteria to bias the results toward a false finding of racial bias: (a) it is a variable that affects police officers' stop decisions; (b) police officers are not using it simply as an indicator of race; (c) it is correlated with race; and (d) it is differentially visible between daylight and darkness. I will discuss the possibilities with respect to these four criteria.

Although one can think of exceptions, driving behavior is generally as visible during darkness as during daylight, while most non-driving behavior is generally difficult to observe in either case.[23] Except for race and, perhaps, sex and height, driver characteristics are very difficult to observe prior to a stop, even during

[23]The data used here were collected during 2002, before widespread concern about cell phone use while driving.

Table 6: Stop rate tests, driving violations only

Average marginal effect of darkness	-0.056^{***} (0.012)	-0.072^{***} (0.013)	-0.077^{***} (0.025)
Time-of-day controls	No	Yes	Yes
Month controls	No	No	Yes

Significance: *** = 0.01, ** = 0.05, * = 0.1. $N = 6{,}334$ Average marginal effects of darkness from logit regressions using stops from morning and evening inter-twilight periods. Bootstrap standard errors based on 500 replications. Time-of-day controls are dummies for 16 minute intervals.

daylight. Sex and height differences among drivers are probably not correlated with race.[24] Thus vehicle characteristics clearly present the biggest challenge, particularly physically small or subtle ones (it is not difficult, for example, to distinguish an SUV from a subcompact after sunset).

Equipment violations and registration violations (such as expired license tags) are the leading examples. They are, in general, less visible in darkness (though non-functioning lights are more visible). They are also likely correlated with income and, therefore, race. If black drivers are more likely than white drivers to have equipment or registration violations, and these violations are less visible in darkness, then a lower proportion of stopped drivers would be black during darkness even if police officers ignore race. Therefore, I repeat the estimation with a sample that includes only stops triggered by driving violations.

These results, reported in tables 6 and 7, are very similar to those from the full sample. In particular, there is strong evidence of racial disparity in stops, but no support for the predictions about searches. The joint tests reported in the lower panel of table 7 provide slightly less evidence against Prediction 2 than those in table 3. A close comparison of the tables reveals this is largely because marginal effects for both white and black drivers are closer to zero in most cases. I conclude that the findings of table 3 are not driven by equipment and registration violations.

[24]NHANES data for 1999-2002 indicate a 0.2 inch average height difference between black and white men and no difference between black and white women (Ogden et al., 2004).

Table 7: Search rate tests, driving violations only

Controls	Narrow definition			Broad definition		
Black drivers	0.005	0.000	0.017	−0.007	−0.014	−0.053
	(0.015)	(0.016)	(0.032)	(0.016)	(0.019)	(0.033)
		N =2,431			N = 2,702	
White drivers	0.001	−0.001	0.000	0.004	−0.003	−0.002
	(0.006)	(0.006)	(0.015)	(0.008)	(0.008)	(0.017)
		N =3,460			N = 3,549	
Joint hypotheses (p-values):						
Prediction 2	0.307	0.327	0.401	0.104	0.164	0.044
Both equal zero	0.828	0.961	0.540	0.796	0.804	0.373
Time-of-day controls	No	Yes	Yes	No	Yes	Yes
Month controls	No	No	Yes	No	No	Yes

Significance: *** = 0.01, ** = 0.05, * = 0.1. Dependent variable is a dummy for a discretionary search; sample excludes stops resulting in non-discretionary searches. Average marginal effects of darkness from logit regressions using stops from morning and evening intertwilight periods. Bootstrap standard errors based on 500 replications. Time-of-day controls are dummies for 16 minute intervals. The joint null hypotheses are described in the text.

A more subtle issue than equipment violations is that police might target particular vehicle characteristics (vehicle profiling). (Note that if police use vehicle or driver characteristics to infer race, the indirect use of race will be detected by the darkness-daylight comparison; this is a strength of the method, rather than a weakness.) Engel and Johnson (2006) summarized cues officers are trained to look for, including a few exterior vehicle characteristics that could be visible prior to a stop. (Most characteristics on the list would be visible only after the stop. The list included no non-driving behaviors or driver characteristics that would be visible prior to a stop.)

The characteristics Engel and Johnson itemize seem unlikely to explain the large relative effect of darkness on stop rates found above. The first of these is vehicle type (not model).[25] However, as mentioned earlier, broad vehicle type is generally visible at night, so it would not generate false finding of racial bias using GR's approach even if police target vehicle type. Another vehicle characteristic cue mentioned by Engel and Johnson was vehicle modification consistent with carrying an atypically heavy load. The implication is that the driver is attempting to conceal the fact that he is carrying cargo by doing so in an ordinary passenger vehicle. Such a modification may well be less visible at night, but also seems unlikely to be generally favored by one race and, in any case, is extremely rare among the general driving public. Third, training frequently mentioned license plates from a drug source or distribution center state. It is not obvious whether out-of-state plates from those specific states are associated with race on Minneapolis streets, but this would seem to be a more likely threat to identification on highways.

Additionally, Johnson (2007) reports that police officers participating in focus groups mentioned illegal window tinting, missing license plates, and evidence of license plates on the wrong vehicle (dirty plates on an otherwise clean car). All of these are probably less visible in darkness, and it is not clear what association these would have with race. However, all are equipment or registration violations

[25]Ikner, Ahmad, and del Carmen (2005) found experimental evidence that officers associate vehicle *types* with racial or ethnic groups. Recall, however, that if the results reflect use of vehicle type to target black drivers, that is a strength of GR's method, not a weakness. More specific information about the brand and model was not mentioned by Engel and Johnson as something officers are trained to look for, perhaps because the sheer number of models makes this impractical.

that are excluded from the sample used in table 6 (provided these were the reported reason for the stop).

Overall, the available evidence suggests that if police do target particular vehicle characteristics, these are likely to be equipment or registration violations, prominent characteristics such as vehicle type that would generally be visible in darkness, or probably not associated with race. Therefore it seems unlikely that the results of the previous section are driven by a practice of targeting vehicles with characteristics correlated with race (but not used *because* they are correlated with race).

Two more observations also suggest that visibility of non-race characteristics does not account for the empirical findings. First, if targeting particular non-race characteristics is good policing and is the reason for apparent racial bias in table 2, one would expect to observe the same thing in other jurisdictions. However, the method has been applied in both Oakland and Cincinnati without finding similar results (GR, 2006; Schell et al., 2007).

Second, if the reason for the strong effect of darkness in table 2 comes from targeting of vehicle characteristics associated with race, one would expect the effect to carry through to search rates because, during darkness, suspicion-arousing characteristics would be more likely to be revealed after the stop (the same logic that leads to Prediction 2 when applied to race). In other words, if there are false positives in testing Prediction 1, there should be false positives in tests of Prediction 2 as well. But there is no evidence of that in Tables 3 or 7.

7 Discussion

The model developed in section 3 assigns the same motivation for racial bias in both stops and in searches through the assumption that race affects expected payoffs. The stop rate tests and the search rate tests are thus looking for evidence of the same underlying phenomenon. At the stop decision, police officers are thinking about both traffic and law enforcement. Prediction 2 is a consequence of using race in the stop decision with an eye toward an option to search.

The empirical results about stops and searches are inconsistent with one another, however. It appears that Minneapolis police were using race in deciding whom to stop, but that this does not have any connection with whom they decided to search. Their actions were, therefore, not consistent with optimization, either in the form of statistical discrimination or exercise of a taste for discrimination. The previous section addressed the possibility that this is because of failure of the identifying assumptions. This section instead evaluates possible interpretations of the inconsistency assuming the effects of darkness on search rates are close to zero.

If police pay attention to race in the stop decision, *and* they are sorting drivers on the basis of J, then average J among stopped black drivers will be higher during darkness than daylight, while the reverse will be true for white drivers.[26] Prediction 2 simply describes how this sorting matters for search rates. Therefore, if darkness affects stop rates, but not search rates, it follows either that police are not using non-race information to sort drivers (effectively violating Assumption 8 or that the sorting somehow becomes irrelevant to search decisions. Either scenario undermines the logic of Prediction 2.

7.1 Do police ignore non-race information in stop decisions?

Available evidence says that the first scenario—that police officers (optimally) ignore non-race behavior and characteristics—is factually incorrect. Some things police officers are explicitly trained to look for in determining whether to search (Engel and Johnson, 2006) are sometimes visible before the stop, and police officers in focus groups frequently mention some types of driving behavior as suspicious (Johnson, 2007). In addition, officers exercise considerable discretion about which traffic violations trigger a stop, and that fact enables (though does not necessarily imply) pretextual stops.[27]

[26] The following discussion abstracts from traffic violations, but recall that the model incorporates traffic enforcement benefits of stops.

[27] Probably the most important court case concerning race and traffic stops was Whren v. United States. The plaintiff argued that a traffic stop leading to an arrest and conviction was pretextual and, therefore, unconstitutional, but the Supreme Court held that a traffic stop is constitutional provided the officer observes *any* traffic violation, regardless its severity

Although logically possible, ignoring non-race information is an especially poor fit with statistical discrimination, as this would imply that police officers see policing value *only* in race, with no mechanism other than randomization for deciding which black drivers to stop. This seems a very likely description of a complicated statistical decision process.

For police officers with a taste for discrimination, it would be optimal to ignore non-race information only if they believe there is no connection between race and crime. They could exercise their prejudice by, for example, more strictly enforcing traffic laws for black drivers, but all drivers would be regarded as equally suspicious, so there would be no reason for darkness to affect search rates.[28]

7.2 Does pre-stop information become irrelevant?

The remaining possibility, that pre-stop, non-race information somehow becomes irrelevant or neutralized in search decisions, necessarily implies non-optimality; there is no reason to use non-race information (other than traffic violations) in the stop decision except to retain the option for further law enforcement actions.

The context of the two decisions made by police officers suggests an alternative to optimization. Research on decision processes reveals that outcomes often differ sharply between decisions made rapidly on the basis of intuition and those to which conscious thought has been applied (Kahneman, 2011). The decision to stop a vehicle must usually be made very quickly. Behavior observed prior to a stop is unlikely to be direct evidence of a crime, and much of what is observed, including race, might not be consciously observed. These circumstances are likely to promote the expression of cognitive and/or racial biases in how information is processed prior to a stop. Subsequent steps in a traffic stop are taken much more deliberately. By the time a search decision is made, the officer will have gathered much more information about the driver and vehicle, and usually several minutes

or the officer's motive (Harris, 1997; Davis, 1997). The arresting officers in the Whren case specifically cited *unusual* driving behavior as a motivation for the stop (Davis, 1997).

[28] The fact that darkness does not affect search rates would not imply there is no bias in searches, only the tests used here would not detect it.

will have elapsed. Also, officers are aware that they might be required by a court to produce an explicit explanation of probable cause for a search. Applying the metaphor used by Kahneman, the stop decision must frequently be made by System 1, while System 2 enters in the decision of whether to initiate a search.

With respect to race, cognitive processes can be influenced by "implicit discrimination", which is hypothesized to be a consequence of unconscious mental associations of race with actions or characteristics, and thus a feature of System 1 (Bertrand, Chugh, and Mullainathan, 2005). Bertrand et al. point to evidence that decisions made under time pressure or involving a high level of ambiguity encourage the exercise of implicit discrimination.[29] One would thus expect implicit discrimination to lead to evidence of bias in stop decisions, but that the conscious deliberation associated with a search would inhibit its expression.[30] In the absence of sorting on J, implicit discrimination would produce racial bias in stops, as in Prediction 1, because race would be observed, consciously or subconsciously, less frequently during darkness. If the implicit bias disappears with deliberation, though, changing observability of race would not create a difference in search rates between daylight and darkness.

Sorting on J almost certainly does take place, however, in which case the results for stops and for searches can be fully reconciled only if the sorting done by System 1 also turns out to be generally unreliable in light of System 2 processing of (more) information after the stop. For example, the sorting could rely on various cognitive biases and heuristics such as neglect of base rates or availability. As Kahneman (2011) puts it, "System 1 is designed to jump to conclusions from little evidence—and it is not designed to know the size of its jumps."

If this interpretation is correct, it is important to recognize that it provides

[29] Price and Wolfers (2010) argue that their evidence of taste-based discrimination among NBA referees is more consistent with implicit discrimination than explicit racial animus partly because referees' make split-second decisions. In a context more closely related to the present research, Payne, Lambert, and Jacoby (2002) found that evidence of implicit racial stereotyping was time sensitive in an experimental task of determining whether an individual was holding a weapon or a tool.

[30] It is possible that a Hawthorne effect contributed to the propensity for conscious deliberation about the search since officers were aware that they were involved in a study of racial profiling. But a Hawthorne effect would not give police officers an incentive to equate daylight and darkness search rates for drivers of one race.

no direct evidence about bias in searches, only an explanation of why the bias is not connected to information available before the stop. Police officers might still assess post-stop suspicion cues in a racially biased way due to language or other behavioral differences (Lang, 1986; Engel and Johnson, 2006; Smith and Alpert, 2007).

8 Conclusions

This paper develops a model of traffic stops and searches, which emphasizes how changing information about a driver's race affects stop rates and search rates when police officers use race in their decisions. Empirical results support the model's prediction about stops, but not the predictions about searches. Except under very restrictive assumptions, this implies that police officers' choices were inconsistent and, therefore, were not the result of statistical discrimination or optimal exercise of a taste for discrimination. The pattern of empirical results could be the consequence of cognitive biases and implicit discrimination active during the time-sensitive stopping decision, but the evidence about this interpretation of the failure of optimization is indirect.

References

Antonovics, Kate and Brian C. Knight, "A New Look at Racial Profiling: Evidence from the Boston Police Department," *The Review of Economics and Statistics,* v. 91, no. 1, (2009).

Anwar, Shamena and Hanming Fang, "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence," *American Economic Revew,* v. 96, no. 1 (2006).

Bertrand, Marianne, Dolly Chugh, and Sendhil Mullainathan, "Implicit Discrimination," *American Economic Review,* v. 95, no. 2 (2005).

Close, Billy R. and Patrick L. Mason, "After the Traffic Stops: Officer Characteristics and Enforcement Actions," *The B.E. Journal of Economic Analysis & Policy,* v. 6, no. 1 (2006).

Council on Crime and Justice and Institute on Race and Poverty, "Minnesota Statewide Racial Profiling Report: Minneapolis Police Department," Report to the Minnesota Legislature, September 22, 2003.

Davis, Angela. "Race, Cops, and Trafic Stops." *University of Miami Law Review,* 51(425) (1997).

Dominitz, Jeff. "How Do the Laws of Probability Constrain Legislative and Judicial Efforts to Stop Racial Profiling?" *American Law and Economics Review,* v. 5, no. 2 (2003)

Engel, Robin S. and Richard Johnson. "Toward a Better Understanding of Racial and Ethnic Disparities in Search and Seizure Rates." *Journal of Criminal Justice,* v. 34, no. 6 (2006).

Grogger, Jeffrey, and Greg Ridgeway, "Testing for Racial Profiling in Traffic Stops from behind a Veil of Darkness," *Journal of the American Statistical Association,* v. 101, no. 475, (2006).

Harris, David A. " 'Driving while Black' and All Other Traffic Offences: The Supreme Court and Pretextual Traffic Stops." *The Journal of Criminal Law and Criminology,* v. 87, no. 2 (1997).

Hernández-Murillo, Rubén and John Knowles, "Racial Profiling or Racist Policing? Bounds Tests in Aggregate Data," *International Economic Review,* v. 45, no. 3 (2004).

Horrace, William C. and Shawn M. Rohlin, "How Dark is Dark? Bright Lights, Big City, Racial Profiling." *Review of Economics and Statistics,* v. 98, no. 2 (2016).

Johnson, Richard. "The Development of Criminal Suspicion by State Troopers during Motor Vehicle Stops." Doctoral dissertation, University of Cincinnati (2007).

Kahneman, Daniel, *Thinking, Fast and Slow,* Farrar, Straus, and Giroux, 2011.

Knowles, John, Nicola Persico, and Petra Todd, "Racial Bias in Motor Vehicle Searches: Theory and Evidence," *Journal of Political Economy,* v. 109, no. 2 (2001).

Lang, Kevin. "A Language Theory of Discrimination," *Quarterly Journal of Economics,* v. 101, no. 2 (1986).

Myers, Samuel L., Jr., "Analysis of Racial Profiling as Policy Analysis," *Journal of Policy Analysis and Management,* v. 21, no. 2, 287-300 (2002).

NAACP, "Born Suspect: Stop-and-Frisk Abuses & the Continued Fight to End Racial Profiling in America," National Association for the Advancement of Colored People, Baltimore (2014).

Ogden, Cynthia L., Cheryl D. Fryar, Margaret D. Carrol, and Katherine M. Flegal. "Mean Body Weight, Height, and Mody Mass Index, Unived States, 1960-2002." National Center for Health Statistics, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services. (2004)

Payne, B. Keith, Alan J. Lambert, and Larry L. Jacoby. "Best laid plans: Effects of goals on accessibility bias and cognitive control in race-based misperceptions of weapons." *Journal of Experimental Social Psychology,* v. 38, no. 4 (2002).

Pope, Devin G., Joseph Price, and Justin Wolfers. "Awareness Reduces Racial Bias." NBER Working Paper 19765 (2013).

Price, Joseph and Justin Wolfers. "Racial Discrimination among NBA Referees." *Quarterly Journal of Economics,* v. 125 no. 4 (2010).

Ritter, Joseph A. and David Bael, "Detecting Racial Profiling in Minneapolis Traffic Stops: A New Approach," *CURA Reporter,* Spring/Summer 2009.

Schell, Terry, Greg Ridgeway, Travis L. Dixon, Susan Turner, and K. Jack Riley, "Police-Community Relations in Cincinnati: Year Three Evaluation Report," Rand Corporation Technical Report 535, 2007.

Smith, Michael R. and Geofrey P. Alpert. "Explaining Police Bias: A Theory of Social Conditioning and Illusory Correlation." *Criminal Justice and Behavior,* v. 34, no. 10 (2007).

9 Appendix: Derivation of Prediction 2

To simplify notation I temporarily suppress both the time-of-day notation and the traffic-enforcement benefit, A. Let $G(J^*|J)$ be the distribution of J^*, conditional on J. For any value of A and any t in the intertwilight interval,

$$\begin{aligned}
&P(\text{search}|\text{stop}, b, \ell = 1) - P(\text{search}|\text{stop}, b, \ell = 0) \\
&\quad = \delta P(\text{search}|\text{stop}, b, b) + (1-\delta)P(\text{search}|\text{stop}, b, u) \\
&\qquad - \eta P(\text{search}|\text{stop}, b, b) - (1-\eta)P(\text{search}|\text{stop}, b, u) \\
&\quad = (\delta - \eta)\Bigg[\frac{1}{1-F(J_b)}\int_{J_b}^{\infty}[1 - G(J_b^*|J)]\, dF(J) \\
&\qquad\qquad - \frac{1}{1-F(J_u)}\int_{J_u}^{\infty}[1 - G(J_b^*|J)]\, dF(J)\Bigg] \\
&\quad = (\delta - \eta)\Bigg[\frac{1}{1-F(J_b)}\int_{J_b}^{J_u}[1 - G(J_b^*|J)]\, dF(J) \\
&\qquad\qquad + \left(\frac{1}{1-F(J_b)} - \frac{1}{1-F(J_u)}\right)\int_{J_u}^{\infty}[1 - G(J_b^*|J)]\, dF(J)\Bigg] \\
&\quad = (\delta - \eta)\Bigg[\frac{1}{1-F(J_b)}\int_{J_b}^{J_u}[1 - G(J_b^*|J)]\, dF(J) \\
&\qquad\qquad + \left(\frac{F(J_b) - F(J_u)}{(1-F(J_b))(1-F(J_u))}\right)\int_{J_u}^{\infty}[1 - G(J_b^*|J)]\, dF(J)\Bigg].
\end{aligned}$$

Assumption 8 implies that the first integrand is less than $[1 - G(J_b^*|J_u)]$ and that the second integrand is greater than $[1 - G(J_b^*|J_u)]$. Therefore, since $F(J_b) < F(J_u)$,

$$\begin{aligned}
&P(\text{search}|\text{stop}, b, \ell = 1) - P(\text{search}|\text{stop}, b, \ell = 0) \\
&\qquad < (\delta - \eta)[1 - G(J_b^*|J_u)]\left[\frac{F(J_u) - F(J_b)}{1-F(J_b)} + \frac{F(J_b) - F(J_u)}{1-F(J_b)}\right] = 0
\end{aligned}$$

The previous inequality holds for every $t \in T$. Averaging over the intertwilight interval, T yields

$$\int_{t\in T} [P(\text{search}|\text{stop}, b, \ell = 1, t) - P(\text{search}|\text{stop}, b, \ell = 0, t)]\, \tilde{B}(t)\, dt < 0, \quad \text{(A-1)}$$

where $\tilde{B}(t) = B(t)/\int_{s\in T} B(s)\, ds$ is the density of observations of b drivers by police officers at t. Since inequality (A-1) holds for any A, integrating over the distribution of A produces the inequality in part (i) of the prediction.

For w drivers, similar manipulations yield

$$\begin{aligned} &P(\text{search}|\text{stop}, w, \ell = 1) - P(\text{search}|\text{stop}, w, \ell = 0) \\ &\quad = (\delta - \eta)\Bigg[-\frac{1}{1 - F(J_u)} \int_{J_u}^{J_w} [1 - G(J_w^*|J)]\, dF(J) \\ &\qquad + \left(\frac{F(J_w) - F(J_u)}{(1 - F(J_w))(1 - F(J_u))}\right) \int_{J_w}^{\infty} [1 - G(J_w^*|J)]\, dF(J)\Bigg]. \end{aligned}$$

The first integrand is pointwise less than $[1 - G(J_w^*|J_w)]$, while the second integrand is pointwise greater than $[1 - G(J_w^*|J_w)]$. In this case the quantity in parentheses is positive. Therefore

$$\begin{aligned} &P(\text{search}|\text{stop}, w, \ell = 1) - P(\text{search}|\text{stop}, w, \ell = 0) \\ &\quad > (\delta - \eta)[1 - G(J_w^*|J_w)] \left[-\frac{\big(F(J_w) - F(J_u)\big)}{1 - F(J_u)} + \frac{\big(F(J_w) - F(J_u)\big)}{1 - F(J_u)}\right] = 0, \end{aligned}$$

Again, the inequality holds for every $t \in T$, and averaging over T and A proves part (ii).