B eing singled out for special attention from the police because of your racial or ethnic identity, rather than your actions, seems profoundly inconsistent with the principle of equal treatment under the law. Yet some racial and ethnic groups, most notably African Americans, have protested what they perceive as manifestly unequal treatment from law enforcement agencies. One area of law enforcement that has drawn considerable attention is the enforcement of traffic laws, where there has been a longstanding perception among non-white drivers that they are far more likely than white drivers to be stopped and possibly searched. In other words, they believe they have been subject to racial profiling.

Because drivers from different racial or ethnic groups differ in other ways as well, convincingly demonstrating statistically the presence or absence of racial profiling has been difficult. The purpose of the present study is to apply a new approach to testing for racial profiling in traffic stops to the data for Minneapolis. We find compelling evidence of racial profiling during 2002, with many fewer qualifications than were necessary in previous studies. The new methodology was developed by Jeffrey Grogger and Greg Ridgeway and is described in detail later in this article. Grogger and Ridgeway applied the method to traffic stop data from Oakland, California, and found no evidence of profiling. Subsequently, researchers from the RAND corporation applied the methodology to data from Cincinnati, Ohio, again finding no evidence of racial profiling.

The issue of racial profiling has been prominent in Minnesota. In 2001 the Minnesota Legislature defined racial profiling as “any action initiated by law enforcement that relies upon the race, ethnicity, or national origin of an individual rather than the behavior of that individual” or information that leads police to suspect the individual of a crime (M.S. 626.8471). Two previous analyses of the Minnesota data conducted by the Council on Crime and Justice/Institute on Race

---


and Poverty\textsuperscript{3} and by University of Minnesota graduate student Nathan James,\textsuperscript{4} found evidence of racial profiling using the same data.

The new methodology used in our study circumvents a key statistical problem that undermines the results of these previous studies—namely that the characteristics of drivers of different racial and ethnic groups also differ on dimensions other than race and ethnicity. These other characteristics may in turn be related to drivers’ behavior, both their (legal) driving patterns and their propensity to commit traffic violations. Since few characteristics of the drivers are recorded in the data, other unobserved characteristics can confound the observed relationship between traffic stops and race. This is to say, if a particular racial group is being stopped at a higher rate than their proportion of the underlying population, it is very difficult to ascertain whether the reason for this is racial profiling or is rooted in driving behaviors that vary by race, such as driving older model cars with more mechanical problems that are likely to result in a traffic stop. This confounding problem is described in more detail below.

The research upon which this article is based was supported through a grant from CURA’s New Initiatives program.

\textsuperscript{3} Council on Crime and Justice/Institute on Race and Poverty. “Minnesota Statewide Racial Profiling Report: Minneapolis Police Department.” Report to the Minnesota Legislature, September 22, 2003. This study was mandated by the Minnesota Legislature.


What Would Motivate Racial Profiling?

People make judgments about each other all the time and for many purposes, and we usually do this on the basis of very incomplete information. Sometimes we use information that is not directly related to the matter we are judging. For example, an employer may take into account that a job applicant attended Harvard University even though having attended Harvard is not directly relevant to the job. Why? Perhaps because Harvard graduates are, on average, very capable employees. The employer is not able to observe the applicant’s ability directly, so instead bets on the average, using a signal that is, in itself, irrelevant. The signal need not be as obvious as a Harvard degree; indeed it may even be subconscious.\textsuperscript{5}

Economists refer to this kind of process as statistical discrimination. It is the process of using perceived characteristics of a population to make decisions about individuals. Racial profiling refers to the use of race or ethnicity in this way for law enforcement; race is used as a conscious or subconscious signal of an elevated probability that an individual has violated the law and deserves further scrutiny. Racial profiling in traffic stops could be used to “get a closer look” at certain drivers or their passengers—to glean more information about whether they are intoxicated or indications that a search might be justified and so forth.\textsuperscript{6} A statistical consequence of racial profiling is disproportionate targeting of certain groups.

Although logically similar, racial profiling and hiring Harvard graduates differ in an important way: Society has decided that racial profiling is not a legitimate law enforcement technique. In fact, it violates explicit police department policy in Minneapolis and many other jurisdictions.

It is important to emphasize that racial profiling is not the same thing as racial prejudice, although, in principle, the two can interact. Pure racial profiling may be a logical response to a situation where actions must be taken on the basis of limited information. However, profiling may not be based on correct beliefs about the average


\textsuperscript{6} Our analysis examines only the frequency of the stops themselves. The decision to search is another possible locus of racial profiling, but we do not address it.
The Statistical Problem

The Problem of Unobservable Factors. The simplest way to look at the Minneapolis traffic stop data also seems to many people to be conclusive evidence of racial profiling: 39.7% of drivers stopped were black, but only 15.8% of the driving-age population in 2000 was black; 44.7% of drivers stopped were white, but 69.8% of the driving-age population was white. It is not difficult to come up with a list of reasons why that conclusion is premature. These can be broken into two groups. The first group boils down to reasons why black drivers might commit more traffic offenses than suggested by their share of the driving-age population. For example, African Americans tend to have lower incomes and, therefore, tend to have less driver training and to drive cars that are more likely to have broken taillights or other equipment violations.

The second group of reasons has to do with patterns of driving and policing. For example, in 1998, the Minneapolis Police Department began a new crime-reduction strategy known as CODEFOR, which uses the geographic distribution of 911 calls and police crime reports to guide the allocation of police resources. As a result of this process, a number of census tracts were targeted for high-intensity law enforcement. The residents of most of these census tracts were disproportionately people of color. Because more police officers were patrolling these areas, there was bound to be a disproportionately large number of traffic stops and the drivers were disproportionately likely to be people of color relative to the population of Minneapolis. Because it is a more or less mechanical consequence of CODEFOR, this fact is not useful evidence about either the use of racial profiling or the propensity of residents in these areas to behave in ways that justify traffic stops.

The official report on the Minneapolis traffic stop data conducted by the Council on Crime and Justice and the Institute on Race and Poverty was unable to address many of these statistical challenges because the data are severely limited, particularly with respect to characteristics of the driver. Nathan James, in a masters thesis submitted to the Humphrey Institute of Public Affairs titled “Is Racial Profiling Occurring in the Twin Cities Metropolitan Area? An Analysis of Traffic Stop Data from Nine Minneapolis-St. Paul Law Enforcement Agencies,” applied multivariate techniques that allowed him to control for several factors simultaneously, but he faced the same data limitations.

The position of the Minneapolis Police Department and many others around the country is that racial profiling violates department policy and that officers simply stop drivers when they see evidence of a violation. Furthermore, they argue, it is very difficult to observe the race of a driver before the stop happens. If there are different stop rates, the difference comes from different rates of violations or factors that are not recorded in the data.

The “Veil-of-Darkness” Methodology.

One approach that appears to avoid all of the statistical problems mentioned in the previous section comes from a simple observation: It is impossible to use racial profiling if you cannot observe the race of the driver. This suggests that useful information can be gleaned by comparing stop rates during the day with stop rates during the night. For example, it is true that stop rates for black drivers were disproportionately high in Minneapolis during 2002. If that were equally true at night, when officers are less likely to be able to identify a driver’s race, it would suggest the explanation cannot be racial profiling; if profiling were the cause, stop rates for black drivers should fall at night.

There are two problems with that empirical strategy. First, the conclusions are muddied by the fact that the ability to identify the race of a driver is not as closely tied to the day/night distinction as one might think; for Minnesota as a whole, officers indicated that they were able to identify the race of a driver only 19.4% of the time during daylight and 9.7% of the time during darkness. The accuracy of these reports has been questioned, but a little personal experimentation easily confirms the main message: It is often impossible to identify the race of a driver during the day (think about following a car on a stormy day) and sometimes possible during the night (under streetlights, for example).

The more important drawback, however, is that driving patterns differ between night and day, and might not vary in the same way for different racial and ethnic groups, but it is difficult to document these patterns. Also, different policing patterns during the night and day might mean different responses to these different driving patterns.

---

In their path-breaking paper, Grogger and Ridgeway proposed an approach that is largely immune to the standard criticisms of racial profiling studies. The core idea is a variation on the one discussed above, that darkness interferes with racial profiling (thus the “veil of darkness” name).

There are two components to the veil-of-darkness methodology. The first minimizes the problem of different driving patterns at different times of day. The second is a statistical test for the presence of racial profiling that relies on the idea that darkness would interfere with profiling, but that does not rely on being able to quantify how much it interferes.

The problem of different driving patterns during the day and at night is the result of comparing the treatment of motorists at different times of day. The veil-of-darkness methodology cleverly circumvents that problem by focusing on the evening intertwilight period—that is, the hours when, at the same clock time, it is light on some days of the year and dark on other days. In Minneapolis, the intertwilight period is roughly 5:05 PM to 9:03 PM. Rather than comparing the treatment of drivers during the day and at night, Grogger and Ridgeway suggest comparing what happens to drivers at specific times of day when it will be dark on some days of the year and light on other days of the year.

Figure 1 helps to illustrate this idea. It shows a scatterplot of the date and time of each traffic stop used in our analysis. The blue dots represent stops that took place during times of darkness, whereas the red dots represent stops that took place during daylight. Each stop (dot) meets two criteria. First, the stop did not occur between sunset and the end of civil twilight, a period that is difficult to classify as either “daylight” or “darkness.” In Figure 1, this period is the curving empty space. Second, the stop occurred at a time of day that would be in daylight at some times of the year and in darkness at other times (in other words, there are both red and blue dots on the horizontal line for that time of day in Figure 1).

Our analysis searches for evidence of profiling of black (or other nonwhite) drivers by comparing the probability that a driver stopped during darkness is black with the probability that a driver stopped during daylight is black, holding time of day constant. Because the method relies on this kind of comparison, deliberate misreporting of race would need to be consistently tilted toward daylight or darkness to bias the results.

An important feature of the veil-of-darkness methodology is that it is based on what social scientists call a natural experiment. In this case, the natural experiment is the intertwilight period during which we can observe traffic stops that occur in both daylight and darkness conditions at the same time of day throughout the year. In the language of experiments, darkness is the “treatment,” people driving during darkness are the “treatment group,” and people driving during daylight are the “control group.”

If one is convinced that the only important difference between the treatment and control groups is the treatment (darkness), then there is no need to control for other factors. For example, if nonwhite drivers actually commit more traffic offenses, that would account for a higher stop rate than experienced by white drivers. But the veil-of-darkness approach does not compare the stop rates of white and nonwhite drivers. Instead it looks for a change between light and darkness in the fraction of stops with nonwhite drivers. That comparison is statistically useful if the drivers at, say, 7:30 PM in August (daylight) have very similar characteristics to those at 7:30 PM in November (darkness). This seems very likely, but because the intertwilight period is not a controlled experiment, we address this issue directly later in this article.10

8 Civil twilight is the period of about 30 minutes after sunset when the sun is less than 6 degrees below the horizon.
9 Grogger and Ridgeway define the intertwilight period as the period between the earliest end of civil twilight and the latest. However, for our analysis, we exclude the period between 9:03 PM and the latest end of civil twilight (9:41 PM) because this period is not classified as daylight according to our first criterion.

10 In the language of experiments, what can go wrong with this story is the randomization of the treatment and control groups. Normally, when using either a controlled or natural experiment, one would evaluate the validity of the comparison by comparing the characteristics of the two groups. Here that is impossible because we have no information about the characteristics of drivers in general, only of drivers who were stopped. However, because the treatment (darkness) is seasonal, any problem along these lines must reflect some kind of seasonal variation in either the driving population or policing patterns.
Readers not interested in details of this methodology may wish to skip to the Results section. In technical terms, the veil-of-darkness method estimates a logistic regression that predicts the race of the driver (black vs. white, Latino vs. white, or nonwhite vs. white) using an indicator for darkness and a function of time of day. The presence of racial profiling implies that darkness would have a negative effect on the probability (i.e., a decreased likelihood) that the stopped driver is black (nonwhite). The variability in time of day accounts for the possibility that the mix of drivers on the road changes throughout the intertwilight period.

We consider three functional forms for the time-of-day effects. The simplest is a linear time trend. The second specification, following Grogger and Ridgeway, is a six-knot natural cubic spline. This allows the regression a great deal of flexibility in fitting time-of-day effects, while insisting that these effects be “smooth.” The last specification instead adds dummy variables for each 15-minute interval during the intertwilight period.

The second component of the methodology is a statistical test for the presence of racial profiling that exploits the difference between darkness and daylight as an imperfect proxy for whether the race of a driver is visible. Specifically, the null hypothesis of no racial profiling is equivalent to the null hypothesis that darkness has no effect in the regressions. Because observed effects combine the (unknown) effects of darkness on visibility and the effect, if any, of racial profiling, the test cannot quantify the practical magnitude of profiling. Another way to put it, in more technical terms, is that the test has low power against the null hypothesis of no profiling, and this is because the imperfect (and unknown) correlation between darkness and visibility reduces the ability of the test to detect deviations from the null hypothesis.

### Results

The top row of Table 1 reports our core results when the sample is restricted to white and black drivers. All of the results refer to the evening intertwilight period. (There were few stops during the morning intertwilight period.) The entries measure how the rate at which black drivers were stopped changes between daylight and darkness. The entry in the first column has the most direct interpretation: the fraction of drivers stopped during darkness who are black was 5.5 percentage points lower than the fraction stopped during daylight. That difference is highly statistically significant, and provides compelling evidence of racial profiling.

Since police officers cannot always identify race during the day and sometimes can identify race during hours of darkness, the 5.5 percentage point figure does not quantify the extent of racial profiling (which we would define as the extent to which black or nonwhite drivers are treated differently when their race is known). The estimate in column 1 does, however, offer a new perspective on the disproportionate rate at which black drivers were stopped because this comparison arguably holds approximately constant everything except the presence of sunlight. During daylight in the intertwilight period, 51.4% of drivers who were stopped were black. But during darkness, when profiling is more difficult, the fraction of stopped drivers who were black shrank to 45.9.

The second through fourth columns of the table use more flexible specifications that allow for time of day effects. The table reports the effect of darkness evaluated at 7:00 PM; the predicted effect varies little during the intertwilight period. These more flexible specifications strengthen the conclusion implied by column 1. The effect of darkness, about 7 percentage points, is somewhat larger and the effect of darkness is statistically significant (i.e., the null hypothesis of no racial profiling is rejected) in every case.

The second row of Table 1 summarizes the results when the regressions compare Latino drivers with white drivers. The estimated impact of darkness is somewhat smaller for Latino drivers, but again the hypothesis of no profiling is rejected in all specifications. The third row compares all nonwhite drivers with white drivers. The results are very similar to the first row because most nonwhite drivers were black.

### Sensitivity Analyses

We considered several variations on the empirical strategy described above. First, we allowed the effect of darkness to vary during the intertwilight period by adding interactions between the darkness indicator and the time-of-day functions. Intuitively, the interaction terms allow the effect of darkness to vary among horizontal strips in Figure 1. The darkness-time interactions were not statistically significant in most specifications, and the overall effect of darkness remained statistically significant in every case (i.e., the hypothesis of no profiling was rejected).

Second, we restricted the sample to the period between 6:00 and 8:00 PM on the grounds that there were relatively few stops in darkness before 6:00 PM.

### Table 1. Results of Tests for Racial Profiling

<table>
<thead>
<tr>
<th>Percentage point effect(^b) of darkness on probability driver is...</th>
<th>Functional form for time-of-day effects</th>
<th>Number of stops</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>None</td>
<td>Linear</td>
</tr>
<tr>
<td>Black</td>
<td>−5.5**</td>
<td>−7.0**</td>
</tr>
<tr>
<td>Latino</td>
<td>−4.7**</td>
<td>−5.2**</td>
</tr>
<tr>
<td>Nonwhite</td>
<td>−5.3**</td>
<td>−6.5**</td>
</tr>
</tbody>
</table>

\(^a\) Null hypothesis of no racial profiling is rejected with a p-value less than 0.0001, meaning there is a less than 0.01% likelihood that these results are caused by chance.

\(^b\) Predicted change in probability associated with switching from daylight to darkness, evaluated at 7:00 PM.

\(^c\) Six-knot cubic natural spline.

\(^d\) Indicator variables for each 15-minute interval.
and relatively few stops in daylight after 8:00 PM. The effect of darkness was still highly significant in this specification.

Third, it could be argued that certain types of equipment violations would be less visible after dark (others, such as burned-out headlights, would be more identifiable). If equipment violations are correlated with race (via income, for example), the proportion of stopped drivers who are nonwhite would drop after dark—the same empirical implication as racial profiling. The same argument could apply to registration violations. Therefore, we restricted the sample to stops triggered by driving violations. Again, the results changed very little.

Seasonality in driving patterns could mean that the racial mix of drivers on the road at any given time of day is correlated with darkness, which would reintroduce the confounding problem the veil-of-darkness methodology is intended to circumvent. We addressed the seasonality issue in two ways. First, we added 12 time-of-year dummy variables (approximately corresponding to calendar months) to the specifications that included time-of-day effects. The time-of-year coefficients were not statistically significant, and the effect of darkness remained highly significant. This suggests that seasonal effects do not account for our results.

The second way we looked for seasonal effects was to take advantage of the switches to and from daylight saving time. On days near these switches there is somewhat more than one hour of clock time that switches from daylight to darkness or vice versa over the course of only a few days. Our final sensitivity check involved restricting the sample to stops near these switches. Stops must have occurred within 10 days of the switch to or from daylight saving time and at a time of day that switches from daylight to darkness or vice versa within that 20-day window. The resulting subsample (Figure 2) contains 341 stops.

Using observations near the switches to and from daylight saving time holds time of year nearly constant while still being able to compare stops during daylight and darkness. The logistic regressions for this analysis include only indicators of darkness and whether the stop was near the fall switch. Using this highly restricted sample, the estimated effect of the darkness indicator is negative and statistically significant at the 5% level in the regressions comparing black and white drivers (283 stops) and comparing nonwhite and white drivers (341 stops). In the comparison of Latino and white drivers (203 stops), the coefficient is negative, but not statistically significant. Overall, this suggests that the strong rejection of the hypothesis of no profiling using the full sample is not due to unobserved seasonal variation in driving or policing patterns.

**Limitations.** The validity of estimates from the veil-of-darkness methodology depends on the implicit assumption that driving and policing patterns during the intertwilight period do not vary much with the time of year. There is no easy way to verify the assumption directly, but we found no evidence of time-of-year effects in our analysis.

A second caveat about our results is that it is impossible to distinguish “vehicle profiling” from racial profiling with the available data. Vehicle profiling would take place if police officers used the characteristics of vehicles (rather than drivers) as signals of an elevated probability of a violation of the law. If vehicle characteristics are also associated with race, vehicle profiling could result in the same disproportionate treatment of some racial or ethnic groups as racial profiling. One plausible reason for an association of vehicle characteristics with race is that the relatively low average income of nonwhite minority groups would lead them on average to drive older, less expensive cars than white drivers. If the relevant characteristics of vehicles are less visible during darkness, vehicle profiling would be detected by the veil-of-darkness methodology.

Although vehicle profiling certainly would not be met with the same level of opprobrium as racial profiling, it seems unlikely that most Americans would approve of the practice.

### Conclusions

Our study used a new approach to detecting racial profiling in 2002 traffic stop data for Minneapolis. This method, developed by Grogger and Ridgeway, compares the stop rates for nonwhite drivers during daylight and darkness using only stops that occur at times of day when it is light at some times of the year and dark at other times. The new approach avoids statistical problems that have plagued the study of racial discrimination.
profiling. Our results point strongly to the presence of racial profiling in Minneapolis.

Finally, we note the traffic stop data were collected during 2002 because of contemporary concern about racial profiling in Minnesota. Subsequently, the Minneapolis Police Department took various steps to try to ensure that racial profiling does not take place. Some of these steps were outlined in a federal mediation agreement adopted in December 2003. The Minneapolis Police Department also invited the Council on Crime and Justice to investigate possible institutional reasons contributing to the disparity in stop rates. More recent data would be required to evaluate the success or failure of these efforts.

Joseph A. Ritter is associate professor in the Department of Applied Economics at the University of Minnesota. His recent research focuses on the role of race in U.S. labor markets. David Bael is a doctoral candidate in the Department of Applied Economics at the University of Minnesota, and a graduate student fellow at the University of Minnesota Institute on the Environment. His current research focuses on conservation planning to preserve biodiversity and ecosystem services.

The research upon which this article is based was supported through a grant from CURA’s New Initiatives program. These grants support projects that are initiated by faculty, community organizations, government agencies, or students and that fall outside CURA’s existing program areas.

11 The agreement is available at www.ci.minneapolis.mn.us/police/about/mcu/federal-mediation-agreement.pdf.


New CURA Reporter Publication Schedule

The CURA Reporter has been in continuous publication since 1970 as a quarterly report of research projects undertaken by the Center for Urban and Regional Affairs. Due to budget and staff constraints, the CURA Reporter will be published semi-annually as a double issue for the foreseeable future, beginning with this spring/summer 2009 issue.

Switching to a twice-yearly publication schedule will help CURA save money on printing and postage while still providing regular updates on our latest publications, projects in progress, and new programs and initiatives.

Look for other changes to the CURA Reporter during the next year, including several design changes, a reader survey, and an option to receive the Reporter electronically. As always, readers can access the CURA Reporter online at www.cura.umn.edu/Reporter.php.