Fair Profiling

John Staddon
Duke University

There are several strategies available to police "stopping" suspects. Most efficient is to stop *only* members of the group with the highest *a priori* probability of guilt; least efficient is indiscriminate stopping. An efficient option that satisfies one criterion for fairness is a strategy that matches stop probability to risk probability. But a strategy that chooses stop probabilities so that the absolute number of innocents stopped is equal for all groups is close to maximally efficient and seems fair by almost any criterion.

Profiling is selecting or discriminating, for or against, individuals, based on easily measured characteristics that are not directly linked to the behavior of interest. For example, age, sex or racial appearance are used as partial proxies for criminal behavior. The term "profiling" is usually associated with stop-and-search procedures (see, for example, Callahan & Anderson, 2001), but a similar process occurs in other contexts also. In life, health and motor insurance, for example, people of different ages and sexes are usually treated differently.

The utility and legitimacy of profiling depend on two related characteristics: *accuracy* and *fairness.* How well do the measured characteristic or characteristics predict the variable of interest? And how fair is it to pick on people so identified?

Fairness is not simply related to accuracy. In health insurance, for example, the whole idea of "insurance" arises because sickness cannot be predicted perfectly. But as biological science advances and it becomes possible to predict debilitating genetic conditions with high accuracy, insurance companies may become reluctant to insure high-risk applicants, who may therefore be denied insurance. How fair is this? In general, the greater the ability of an insurer to predict health risk, the more questionable profiling becomes, because the concept of insurance — spreading risk — is vitiated.

On the other hand, few would object to a regression equation that allowed airline passenger screeners to identify potential terrorists with 90% probability. The better police are able to profile, the fewer innocent people will be stopped and the more acceptable the practice will become. But in reality, predictability is never perfect. We are not, like the computer in *Minority Report,* able to predict perfectly who will commit crimes so potential criminals can be arrested in advance.

Almost any kind of selective treatment that is based on a proxy variable is a form of profiling. Intelligence-type tests such as the GRE and SAT, for example, do not measure college performance directly. The test scores are proxies, usually pretty good ones and certainly the best single measures available, for a student's subsequent performance.

Often controversial, profiling nevertheless goes unquestioned in some surprising places. Take *speeding* by motorists, for example. Speeding is an offence and few question laws against it. But a speeder causes no direct harm to anyone. The legitimacy of a law against speeding rests on the accuracy which speeding predicts risk. While it is obvious that an accident at high speed must cause more damage than one at lower speed, the relation between speed and accident *probability* is more contingent. The statistically expected cost of speeding is the product of accident probability times damage caused. If drivers go fast only when it is safe to do so, there may be no, or only a weak or even negative, correlation between speed and the likelihood of an accident. Hence, the incremental expected cost of higher speed may be small or negligible. In which case penalizing — profiling — speeders would be questionable.

The same is true of alcohol and driving. If drunks drive more cautiously (as some do), their proven sensory-motor deficiencies may become irrelevant to accident risk. In these cases also, the fairness of profiling rests on its accuracy. If drinking and speeding really do cause accidents — or at least, are correlated with higher accident probabilities — sanctions against them may be warranted.

Finally there is the issue of personal responsibility. Speed is under the motorist's control, to be sure, just like smoking — which is used in life-insurance profiling. Fewer objections are raised to profiling that is based on proxies that are under the individual's control and for which he can therefore be held responsible. IQ is an interesting intermediate case. Despite some controversy because of racial differences in IQ scores, IQ-type tests are generally regarded as fair, even though IQ is generally assumed to be more or less fixed in adults, hence something over which the individual has no control. Race is of course not something over which the individual has control, which is one reason racial profiling is subject to criticism. On the other hand, neither are age and sex. Yet fewer objections are raised against profiling on these grounds.

It seems unlikely that the desirability or otherwise of profiling rests on a single principle. It is equally clear that its accuracy, and the costs and benefits associated with the practice, must figure in any welfare calculus. With these ideas in mind, the main part of the paper is devoted to a simple quantitative exploration of the accuracy of profiling in "stop-and-search" situations such as driver stops or airport screening. The quantitative analysis in fact allows us to identify a profiling strategy that is both efficient and fair in many cases.

STOP-AND-SEARCH PROFILING

Age and sex profiling are essentially universal: police rarely stop women or old men; young males are favored. The reason is simple. Statistics in all countries show that a young man is much more likely to have engaged in criminal acts, particularly violent acts, than a woman or an older man. The same argument is sometimes advanced for racial profiling, stopping African-American drivers, or airline passengers of Arab appearance, more frequently than whites or Asians, for example, but in this context it is politically controversial. The political and ethical problems raised by profiling and associated practices, and some of the utilitarian aspects of stop-and-search profiling, have been extensively reviewed (see for example, Persico, 2002; Risse & Zeckhauser, 2004). Little can be said that is entirely new. Nevertheless, I will try to show that something novel can be revealed by analyzing in as simple a way as possible the logic behind the practice of profiling.

For simplicity, let's suppose there are just two discriminable groups in the population, A and B. And let us suppose that accurate statistics are available giving the baseline conditional probabilities, p_A and p_B, that a suspect randomly selected from each population is guilty of a crime. Let's stipulate that $p_A \geq p_B$; that is, the probability that a random individual from group A is likely to be a perpetrator is greater than the probability a similar individual from B is a perpetrator. And finally, the proportion of And B individuals in the population is just r and $1\text{-}r$ (e.g., .12 vs. .88, if A is the minority).

I will discuss four possible "stop" strategies:

Indiscriminate

Members of groups A and B can be stopped with equal probability. In this case, the probabilities a given stop will catch a criminal are p_A and p_B, where p_A and p_B are the base probabilities that a given A or B individual has committed a crime. Taking account of the relative frequency of group A and group B individuals in the population, A/(A+B) = r, the net probability of capture per stop is just

$$P_{capt} = rp_A+(1-r)p_B. \quad (1)$$

For example, if $p_A = .2$, $p_B = .1$ and $r = .12$ (i.e., the minority population is twice as likely as the non-minority to be detected in a crime if stopped) the net per-stop probability of a capturing a criminal, P_{capt}, is just **.112**.

In terms of numbers of individuals stopped, the analysis proceeds in this way. Assume a sample population of 10000 individuals (say) and a stop probability s. The number of A's stopped under this strategy will be 10000sr and Bs 10000$s(1-r)$. If s is .1, say, this means that 120 As and 880 B's will be stopped. 24 As (on average) will be guilty and 120-24 = 96 innocent. 88 Bs will be guilty and 880-88 = 792 innocent. Overall, more than 8 times as many innocent B's will be stopped as innocent A's.

Notice that under this strategy a guilty A is more likely to escape detection than a guilty B. The probability a guilty A escapes is just $p_A(1-s)$ and for B, $p_B(1-s)$, where s is the base stop rate. Since $p_A > p_B$, A criminals are proportionately (to the A population) less likely to be caught than B criminals under the *indiscriminate* strategy.

Matched

The police can stop members of each group in proportion to their chance of being guilty (as estimated from prior statistics). Given that $p_A > p_B$, this would mean an innocent person in group B would have a smaller chance of being stopped than one in group A. Using the same symbols as before, the capture probability for the ***Matched-G*** strategy is

$$P_{capt} = \alpha p_A+(1-\alpha)p_B, \quad (2)$$

where $\alpha = p_A/(p_A+p_B)$.

Using the same numbers as the previous example, the net per-stop probability of capture rises to **.167**.

In terms of numbers, given a stop probability of .1, as before, 171 innocent A's will be stopped and 707 innocent B's, for a B:A innocent-stop ratio of 4.1, a substantial improvement over the 8 times ratio of the *indiscriminate* strategy.

Notice that under this strategy, the probabilities that A and B criminals are *not* stopped are equal:

$$p_{esc}(\mathrm{A}) = p_{esc}(\mathrm{B}) = p_A p_B/(p_A+p_B). \quad (3)$$

In other words, there is an equal proportion of undetected criminals in each group under the *matched* strategy.

Matched-I is a variant of the *matched* strategy which matches stops not to the chance of the subject being guilty but to the chance of his being innocent. It is straightforward to show that if the stop probabilities for the two groups are in the ratio

$$\frac{p_{stopA}}{p_{stopB}} = \frac{B(1-p_B)}{A(1-p_A)}, \quad (4)$$

where A and B are the numbers of As and Bs in the population, as before, then the absolute number of innocents stopped will be the same for both groups. Or, in the same form as Equation 2,

$$P_{capt} = \beta p_A+(1-\beta)p_B, \quad (5)$$

where $$\beta = \frac{B(1-p_B)}{A(1-p_A)+B(1-p_B)}. \quad (6)$$

In this case, with the numbers given, the capture probability rises further to **.189.** With an absolute stop probability s = .1, an equal number, 423, A and B innocents are stopped in a sample of 10,000 individuals.

Efficient

The efficient strategy is to stop *only* members of the group with the higher probability of criminality (i.e., $\alpha = 1$ in Equation 2). In this case,

$$P_{capt} = p_{A..} \quad (7)$$

Net capture probability in this case, with the previous numbers is **.2,** which is the highest per-stop capture rate possible.

In terms of numbers, for a total of 1000 stops out of a population of 10,000, as before, this strategy means that 800 innocents in group A (96) will be stopped. None, innocent or guilty, will be stopped in group B. All group-B criminals will escape.

The *efficient* strategy is both inequitable, since group B indviduals are never stopped, and also *in*efficient in a larger sense. It encourages Group B members to feel immune from detection, hence encourages crime in that group. And because it fails to sample the low-crime group cannot update the statistics that allow the estimation of p_B that justifies the strategy itself.

Costs and Benefits

The preceding analysis deals only with capture probabilities. For a complete welfare calculus, the costs and benefits of each strategy must also be included. This is a much more difficult practical and philosophical problem. For example, is the capture of a criminal a *benefit,* because the just retribution he may now suffer for his crime benefits social general welfare? Or it just that *failure* to capture constitutes a *cost* because of the future crimes he will commit? In the first case we are interested in the number of criminals captured, in the second, in the number who evade capture. If the cost per criminal escaped is equal to the benefit per criminal caught, the two measures are equivalent. But if they are not — if, for example, the benefit of a criminal caught is significantly greater than the cost of a criminal escaped — then we need to know both numbers to calculate the net benefit of a given policy. And what of the stopping of innocent people? Each innocent "stop" clearly entails some social cost, but how much? Are the costs of innocent stops just proportional to the rate at which they occur or does the cost accelerate as the rate of stops increases? Are the costs the same for all groups or are stops of people in groups that already feel victimized somehow more costly than others — and should such a difference be recognized by public policy?

Despite these complications, the public policy implications are in fact fairly straightforward. Because all stops entail some social cost, because that cost may be an accelerating one, and because some groups may be highly sensitive to being stopped, our focus should be on minimizing stops of innocent people and making such stops as "fair" as possible.

FAIRNESS

Fairness is not a well-defined concept. It depends to a large degree on perceptions of the members of both groups as well as the specifics of implementation. Do innocent group As recognize the tradeoff between efficiency in catching the guilty versus increased stop risk for them under the *efficient strategy*? Do they appreciate that the *indiscriminate* strategy, which treats innocent members of both groups the same way, nevertheless implies much more efficient crime prevention for low-risk group B than high-risk group A — because a larger fraction of group A criminals than group B criminals will go unstopped? (This is simply a consequence of the facts that A's and B's are equally likely to be stopped, but $p_A > p_B$.) The way in which a stop policy is implemented — the politeness of enforcement officers, the apparent justice of the specific situation in which the stop is made, and a myriad other practical factors, none easily amenable to formal analysis — is perhaps the most important factor in any stop policy.

Nevertheless, no matter how you define the concept, the four strategies make the point that when groups differ in their base rates of criminality, *fairness* is often opposed to *efficiency*. But not always: A little reflection shows that one of the stop strategies is both fair and highly efficient.

The most efficient strategy is also, by one standard, the most unfair, since members of group A are stopped at the maximum rate but group Bs are never stopped; i.e., the *probability* that an A will be stopped is maximal, that a B will be stopped, zero. On the other hand, the total number of stops per capture under the *efficient* strategy is the smallest of the four. And, perhaps most important, the number of *innocent* people stopped is also minimal. But there is a substantial group disparity in both these measures.

At the other extreme, the *indiscriminate* strategy sounds fair and indeed the probability of a stop is the same for both groups. But the total number of stops per capture, and the number of innocent stops, are both large. And the group disparity in innocent stops is large: 8 times more innocent B's are stopped than innocent A's, in the example. Neither the *efficient* nor the *indiscriminate* strategy is very fair.

The two *matched* strategies are much better. In the *matched-G* strategy, for example, the total number of stops is less than in the *indiscriminate* strategy and the ratio of B: A innocent stops is only 4:1 in the example.

But the fairest is the *matched-I* strategy. The numbers of innocent-A and -B stops are explicitly equalized and the total number of stops/capture is less than in either the *indiscriminate* or the *matched-G* strategies. Moreover, the *matched-I* strategy is more efficient (in terms of capture probability) than all the other strategies except the *efficient*: 0.189 vs. 0.2 in the example. And because it allows sampling of both groups, statistics on base (a priori) criminality rates can be kept up to date.

CONCLUSION

When base criminality rates differ between groups, profiling — allocating a limited number of stops so that members of one group are stopped more often than another — captures more criminals than an indiscriminate strategy. The efficiency difference between the two strategies increases substantially as the base-rate difference in criminality increases, which can lead to a perception of unfairness by innocent members of the high-risk group.

Other than the highly inefficient (and unfair by one measure) *indiscriminate* strategy, there is no stopping policy that will equate stop *probabilities* between the two groups. Nevertheless, because no one seeks to minimize the stops of guilty people, it seems more important to focus on the treatment of innocent people rather the population as a whole. And because we live in a democracy, numbers weigh more than probabilities. These two considerations suggest a solution to the fairness problem. The highly efficient *matched-I* strategy stops group A and group B members differentially in a proportion that equates the *numbers* of *innocent* people stopped in both groups. It therefore represents a fair formal solution to the profiling problem.

REFERENCES

G. Callahan & W. Anderson *The roots of racial profiling: Why are police targeting minorities for traffic stops?* Reason, August-September, (2001), http://reason.com/0108/fe.gc.the.shtml

N. Persico, *Racial Profiling, Fairness, and Effectiveness of Policing*, 92(5), The American Economic Review, 1472-1497 (2002)

M. Risse & R. Zeckhauser, (2204) *Racial Profiling*, 32, 2, Philosophy and Public Affairs, Research Library, 131-170. (Spring 2004)