Determining the Social Effects of a Legal Reform

The British "Breathalyser" Crackdown of 1967

H. LAURENCE ROSS
University of Denver
DONALD T. CAMPBELL
Northwestern University
GENE V GLASS
University of Colorado

The social effects of a legal reform are examined in this paper utilizing the Interrupted Time-Series research design, a method of analysis that has broad potential use in studies of legal change more generally. A previous demonstration of the applicability of this design to the sociology of law concerned the Connecticut crackdown on speeders (see Campbell and Ross, 1968; Glass, 1968). In that study, the substantive findings were that the crackdown had little effect on the highway death rate, and that it introduced certain unexpected and undesirable changes into the legal process in Connecticut. The present study concerns a similar attempt to lower the highway death rate through changes in the law, specifically the British Road Safety Act of 1967. Critical scrutiny of the data indicates that in this instance the legal change quite impressively achieved its goal.

The British crackdown attempted to get drunken drivers off the road, and thus took aim at a scientifically demonstrated correlate of automobile accidents. The Connecticut crackdown, in contrast, was based on commonsense considerations unsupported even by correlational studies. Its sponsors claimed success prematurely, before such possibilities as random variation and statistical regression could be ruled out as explanations of an apparently striking decline in accident rate. In the present study, similar claims turned out to be justified.

Authors' Note: *This study was supported in part by National Science Foundation Grant G51309X.*

In presenting this report, we hope for two consequences. Substantively, we hope that officials concerned with traffic safety will consider adopting a legal reform which has proved in one notable instance to be effective in reducing traffic deaths; methodologically, we hope to increase awareness of the need for hard-headed evaluation of legal and administrative reforms, and of the value of experimental and quasi-experimental designs for this purpose.

INTERRUPTED TIME-SERIES ANALYSIS

The Interrupted Time-Series is a quasi-experimental design (Campbell and Stanley, 1966; Campbell, 1969) for studying the effect of a given "treatment" on a variable that is repeatedly measured over a period of time before and after the application of the treatment. Like all quasi-experimental techniques, the time-series design is a substitute for an unfeasible true experiment. The true experiment requires randomized assignments of subjects to experimental and control groups, but the time-series design can be used, albeit with greater equivocality, in situations lacking randomization.

The essence of an Interrupted Time-Series design is the extension of a typical before-and-after study to a series of observations at various times removed from the experimental treatment, both before and after. To illustrate, the typical before-and-after study concerns only points immediately prior and subsequent to the treatment, as in Figure 1, which compares accidental deaths in Connecticut before and after the crackdown on speeding. It is very difficult to interpret any change from before to after the treatment for various reasons, discussed in more detail in our full presentations. Briefly, these reasons are:

(1) History. The change observed may be due to simultaneous events other than the experimental treatment.

(2) Maturation. The change may be part of some long-term trend.

(3) Instrumentation. The measured change may be based on a change in the means of measuring, rather than in the thing being measured.

(4) Testing. The change may be caused by the initial measurement rather than by the treatment.

(5) Instability. The apparent change may be no more than chance or random variation.

(6) Regression. If the group was selected because it was extreme on some measure, statistical reasoning indicates that it will appear less extreme on subsequent tests, even though the intervening treatment may be completely ineffectual.

A study of Figures 1 and 2 of Campbell and Ross (1968: 38, 42) will illustrate the relevance of time-series data to four of these six threats to validity. In this Connecticut case, maturation and testing are pretty well ruled out by the extended data series inasmuch as *both* posit processes that would have existed in prior years and inasmuch as the 1955-1956 drop is not interpretable as a continuation of trends manifest in 1951-1955. History and instrumentation are not controlled by this design, but an examination of plausible alternative causes such as winter weather and possible changes in record-keeping make this implausible as rival explanations of the 1955-1956 drop. It is on the threats of instability and regression that the time-series presentation exposes weaknesses invalidating the public pronouncements of the Connecticut experiment.

Instability was a possibility totally neglected. *All* of the 1955-1956 change was attributed to the crackdown; the Governor of Connecticut stating, "With a saving of forty lives in 1956, a reduction of 12.3% from the 1955 motor vehicle death toll, we can say the program is definitely worthwhile." When the prior years are examined it become obvious that the 1955-1956 shift is typical of the usual annual shifts, rather than being exceptionally large. The problem of regression was likewise overlooked. When a treatment is applied because of extremity on some score (e.g., remedial reading courses applied to persons because of their low reading comprehension scores) it is likely that subsequent scores will on the average be less extreme due to statistical regression alone, even if the treatment has had no effect. The problem of regression is not easy to communicate briefly. It will be helpful to think of a time-series that fluctuates completely at random. If one moves along the series, selecting points that are extraordinarily high, on the average subsequent points will be lower, less extreme and closer to the general trend. In the Connecticut case it appears certain that the great 1954-1955 increase instigated the crackdown. Thus the point where the treatment was instigated was selected for its height. Therefore, a good part of the 1955-1956 decrease must be attributed to statistical regression.

Legal change is a subject for which the Interrupted Time-Series design seems eminently suited. True experiments can seldom be performed in the law because all persons receive the treatment at the same time or because, even if only some receive it, legal or practical considerations prevent the necessary randomization. If a policy strikes a legislature or an administrative body as being a good idea, it is adopted wholesale; if it seems unpromising at first glance, it may not be tried at all. Moreover, even when a change is adopted "experimentally," it is seldom applied at random to one group of people and not to another similarly situated group. The experimental change is typically put into full-scale effect for either an arbitrarily limited time or for a single jurisdiction chosen nonrandomly from among many others. The time-series design is appropriate in these circumstances.

The opportunity to work with time-series design in studies of legal change is enhanced by the fact that there are numerous series of data that are routinely gathered by governmental bureaus and agencies. Examples are general and specific crime rates, institutional commitments, case loads, economic indexes, and accident rates. Because these data are routinely gathered, their measurement is not taken by participants as a cue that a study is being done (Webb et al., 1966). Generalization to other groups involves fewer theoretical problems than for laboratory experiments because of the much greater similarity between field of experimentation and field of application.

The special relevance of time-series data for questions of legal impact has no doubt frequently been recognized, even though simple before-and-after figures, or percentage change from the previous year remain the commonest means of reporting. Time-series data have been employed by Stieber (1949) and Rose (1952) in studies of the effects of compulsory arbitration; by Wolf, Luke and Hax, (1959), Rheinstein (1959) and Glass (forthcoming) in studies of divorce law; and by Walker (1965) and Schuessler (1969) in studies of the effects of capital punishment. But the formal development of the method, the analysis of its strengths and pitfalls, the development of appropriate tests of significance, are all too recent for method to have received the widespread application it deserves. We have previously reported a negative application, primarily rejecting the Connecticut claims. In the present paper we report an optimistic one, in which effects claimed in press releases stand up under scientific scrutiny.

ALCOHOL AND TRAFFIC ACCIDENTS

THE LEGISLATION AND BACKGROUND

The sponsors of the British Road Safety Act of 1967 based their action on a voluminous scientific literature which showed association between accidents, particularly serious ones, and blood alcohol, particularly in high concentrations. In a recent review of the literature, three studies of fatal accidents were cited in which the proportion of drivers with alcohol in their bloodstreams ranged from 55 to 64%. In single-vehicle accidents, three other studies revealed alcohol in from 71 to 83% of the victims (Automobile Manufacturers Association, 1966). One of the latter studies matched the deceased drivers with a sample obtained later of drivers in the same location at the same hour. Only 23% of the controls had a concentration of .02% or more of alcohol in their blood, compared with 71% of the deceased drivers.

There were similar findings in reports of several correlational studies of nonfatal accidents. The U.S. Department of Transportation recently issued a report containing the following summary:

> Scientific investigation of actual crashes and the circumstances in which they occur and laboratory and field experiments show very clearly that the higher a driver's blood alcohol concentration:
>
> –the disproportionately greater is the likelihood that he will crash;
>
> –the greater is the likelihood that he himself will have initiated any crash in which he is involved; and
>
> –the greater is the likelihood that the crash will have been severe.
> [House Committee on Public Works, 1968: 15]

The British government, then, had a good theoretical basis on which to form their program of control. The attempt was further justified by claims of success in similar programs in the Scandinavian countries (Andanaes, 1966). The state of knowledge about alcohol and accidents is quite different from the existing knowledge about the effect of speed, which indicates no simple relationship with accidents.

Since 1925, it had been an offense in Britain to drive while under the influence of alcohol. However, as one British lawyer explained:

> I knew only too well how easy it was to secure acquittal from a charge of drunken driving in the United Kingdom. The form one adopted for the defense was always to insist on a jury trial; the evidence as to drunkenness was always given by the Police Surgeon who had made the drunken man carry out some rather extraordinary tests, many of which perfectly sober people could not carry out. You would inevitably find that your jury consisted of people like myself, honest, law-abiding citizens who both drove motor cars and also drank alcohol. The inevitable reaction of juries faced with a case of this nature was "there but [for] the grace of God go I . . . Not Guilty" [Insurance Institute for Highway Safety, 1968: 40].

Legislation in 1962 permitted blood and urine tests, with certain presumptions to be raised in the event of the driver's refusal to cooperate. The stimulus for additional legislation was a continued rise in automobile-related deaths and serious injuries. Deaths had peaked in 1966 at 7,985, a culmination of a steady rise throughout the 1950s and 1960s. Injuries peaked a year earlier, but remained quite high (384,000) in 1966.

The new legislation, put into effect on October 9, 1967, was not particularly radical as compared, for instance, with Scandinavian procedures, or even with the laws in several American states. However, the Act was well publicized in Britain and included the following features:

(1) The criterion of impairment was set at a blood alcohol level of .08%. This is a more stringent standard than that prevailing in most American states, but less so than that prevailing in Norway and Sweden (.05%) or in Czechoslovakia, Bulgaria, and East Germany (.03%). A blood alcohol level of .08% might be barely reached if a 160-pound man drank three drinks in quick succession on an empty stomach (Campbell, 1964).

(2) Police were authorized to give an on-the-scene breath test. This test, called the Breathalyser, gave its name to the crackdown in the British press. The test may be administered to a driver if "the constable has reasonable cause– (a) to suspect him of having alcohol in his body; or (b) to suspect him of having committed a traffic offense while the vehicle was in motion." The test may also be given to any driver involved in an accident. A driver who fails the breath test is brought to the police station for a (more accurate) blood or urine test, on the basis of which a charge is made.

(3) A mandatory punishment was instituted, consisting of "disqualification" (license suspension) for one year and a fine of £100 or imprisonment for up to four months, or both. Severe penalties were also instituted for failure to submit to the breath test or to either the blood or urine tests.

(4) The specific starting date for the new regulations was given advance publicity. This provides an essential aspect making the study interpretable. A very gradual change of enforcement would have produced results indistinguishable from a gradual change in long-term trends.

Although official publicity campaigns greatly increased public awareness of the new procedures and penalities, particularly of the on-the-scene breathalyzer test, enforcement was probably not much increased. During the first six months after the act was initiated, only 20,000 drivers had to take the test, and fewer than half of them failed it. A report commissioned by the Insurance Institute for Highway Safety states "that in reality [the British driver's] chances of being apprehended for driving after drinking are no greater than they were before" (Bennett and Westwick, 1968: 10).

CLAIMED RESULTS

As in the case of the Connecticut crackdown, the fact of fewer casualties in the period immediately following the institution of the reform was interpreted as evidence of an effect. The Ministry of Transport in its official press releases was considerably more restrained than the governor of Connecticut had been, but its claims were based on much the same kind of reasoning. For instance, a press release of March 21, 1968, was headlined: "Road Casualties in 1967 Lowest Figure for Nine Years." This release documented the fact that in the last three months of 1967 casualties had declined by sixteen percent and deaths had declined by twenty-three percent; readers were reminded that the Road Safety Act came into force on the 9th of October. On December 11, 1968, the Ministry of Transport issued a press release headline: "First Twelve Months of 'Breath Test.' 1,152 Fewer Dead on Roads." Although the term "cause" was never used, the report contains statements about "casualty savings" and "gaining safety from the new legislation." The magnitude of the shift, particularly in the night hours when the casualty rate declined by a third, makes the British interpretation less offensive than the official line in Connecticut. Our statistical analyses, in fact, support the press releases. But the claims failed to indicate that thought had been given to such obvious alternative causes of the decline as instability of the casualty rate, regression from peak statistics, and other safety-related events taking place at the same time.

INTERRUPTED TIME-SERIES ANALYSIS OF THE BREATHALYSER CRACKDOWN

THE STATISTICS

A graphic presentation of some of our time-series analysis of the Breathalyser crackdown is shown in Figure 1. Our analyses are based upon statistics made available to us by the British Ministry of Transport, including breakdowns going beyond the data reported in their press releases, some of which were made especially at our request.[1] We report

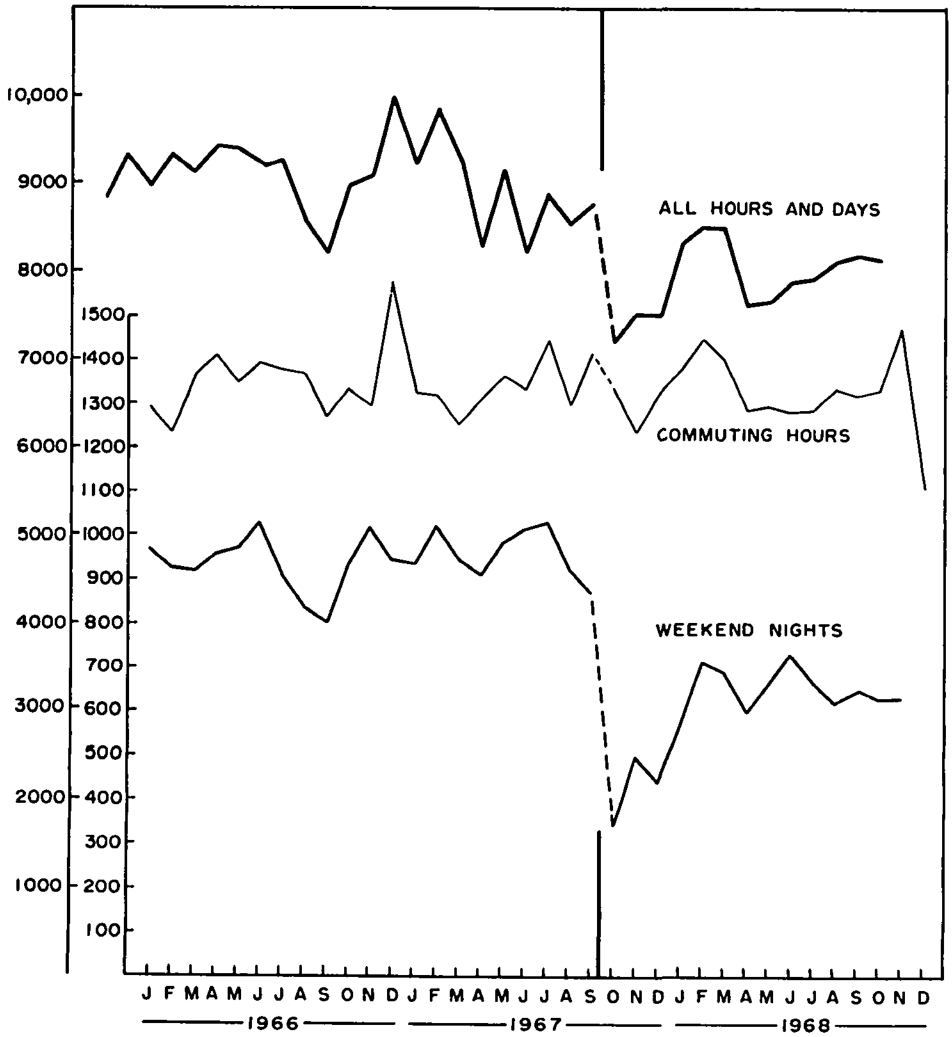

Figure 1.

[**Authors' Note (added in press)**: We have just been informed by the Ministry of Transport that, through error, the data we were supplied for weekend nights are actually for the periods Thursday midnight to 4:00 a.m. Friday; Friday 10:00 p.m. to 12:00 p.m.; Friday midnight to 4:00 a.m. Saturday morning; and Saturday 10:00 p.m. to 12:00 p.m. By the inclusion of the late Thursday night (Friday morning) data and by the omission of the late Saturday night (Sunday morning) data, we have underestimated the weekend night effects. We have also been informed that the data labeled "commuting hours" should have been labeled "closed hours," as it includes closed hours for Saturday and Sunday as well as for the weekdays.]

here only a portion of our analyses, selected so as to display the major results and to illustrate the method. For the full presentation, including alternative analyses, see Glass et al. (forthcoming).

All hours and days. Data on total monthly casualty rates by seriousness of casualty are available back to 1961. For the present analysis, we have focused upon the combination "fatalities plus serious casualties," since specific hour and day analyses, to be discussed later, were only available in that form. These data have been smoothed in two stages: first, the rates for months of 28, 29, and 30 days in length have been extrapolated to 31-day equivalents (by multiplying the obtained rates by 31/28, and so forth). Second, a yearly cycle of seasonal variation has been removed. (In this cycle, January is lowest, August through December high.) The average monthly rate prior to the crackdown has been used to compute a monthly correction such that the mean annual correction is zero. These corrections have been added to, or subtracted from, the 31-day rates. The crackdown began on October 9, 1967. In order to plot October as a purely posttreatment value, an additional prorating has been employed.[2] Without the prorating, October 1967 would have had a plotted value of 7681, instead of 7226 shown.

Visual inspection of the all-hours-and-days graph supports the hypothesis that the crackdown had an effect. A simple nonparametric consideration offers further confirmation: the September-October drop of 1967 is the largest one-month shift not only for the three years plotted in Figure 1, but also for the total series going back to January 1961. The odds against this are 93 to 1. This holds true even if the uncorrected value for October 1967 is used. The most sophisticated test of significance in this situation is that developed by Box and Tiao (1965).[3]

Taking out the annual seasonal cycle is a problematic matter (and even more so in the shorter series that follows). Our procedure is only one of many, all of which are unsatisfactory in one way or another. Can the effect be noted in these all-hours-and-days data without the annual cycle being removed? Only if one compensates for it by eye, or by seasonally controlled comparisons. The largest shifts in the uncorrected data tend to be the annual December-January drops, four out of seven larger than our focal September-October 1967. On the other hand, the 1967 September-October drop of 1,654 (or of 1,199 if October be left with the eight precrackdown days uncorrected for) greatly exceeds the seven other September-October shifts, being three times as large as the largest of them.

Week-end nights. Going back continuously until January 1966, the Ministry of Transport has monthly statistics by hours of the week. From their analyses and press releases, it was apparent that Friday and Saturday nights, from 10:00 P.M. to 4:00 A.M. the following morning, were the hours in which the effect of the Breathalyser was strongest. The bottom time-series in Figure 1 depicts these data. Here, the figures have been prorated to four Fridays and four Saturdays each month. Rather than attempt to estimate seasonal cycles on the basis of these few data, it seemed appropriate to use the monthly corrections based on all hours and days, proportionately reduced in magnitude.

These data provide striking evidence of efficacy. The casualty rate seems to initially drop some forty to forty-five percent, and to level off with a net reduction of perhaps thirty percent.

The September-October 1967 drop is four times as large as any other month-to-month change. The Box and Tiao statistic produces t in excess of 6.50 for likely magnitudes of effect, for which the chance probability is less than .0000001.

For these weekend nights, the data are convincing visually even when seasonal corrections are not made. Even then, the 1967 September-October drop far exceeds any other month-to-month change, including the December-January ones.

We have chided the typical administrator for being too quick to announce success without taking into account instability, and without adequate sampling time periods before and after the legal change. But when does the administrator have enough evidence? This is in part a function of the prior instability of the series, and in part a function of the magnitude of the change. In the present instance, using the Box and Tiao test and the common acceptance level of a chance probability of less than .01, the administrator could have announced a significant drop after only one posttreatment month. As a matter of fact, after only one month the t value was 8.63 where $p < .01 = 2.86$, $p < .001 = 3.88$. Even for the all-hours-and-days data, he would have had to wait only one month, at which time the t value was 3.27.

Commuting hours as a control. While both series of data considered so far indicate that the crackdown had an immediate effect, it becomes important to know to what extent that effect has been sustained. Such considerations involve inferences as to what the long-term trends would have been without the crackdown. On the basis of increased traffic volume, one would expect a steady rise. On the basis of increased availability of divided and limited access highways, one would expect a decline. The trend was actually downward 1961-1963, markedly upward 1963-1965, and slightly downward from January, 1966 until the crackdown. Thus there are no grounds here for extrapolation.

What one needs in such cases is a "control group" or some other control comparison. In the Connecticut case, we were able to use data from adjacent and similar states for this purpose, assuming similar weather, vehicles, and safety changes in the absence of a crackdown. Such comparisons never achieve the effectiveness of the randomly assigned control groups of true experiments, but are nonetheless useful. Because of differences in drinking and closed hours, as well as rate of automobilization and highway construction, Irish or Belgian data would be of less use as a control than were other states for Connecticut, but they would still be of value.

But control data series need not come solely from different persons, groups or populations. In the present situation, a valuable comparison would come from those high accident hours least likely to be affected by drinking. Commuting hours during which British pubs and bars are closed seemed ideal. Casualties on the five working days between the hours of 7:00 to 10:00 A.M. and 4:00 to 5:00 P.M. were chosen (pubs close after lunch at 2:30 P.M.). These monthly rates were prorated to 23 working days per month. These data showed a distinctly different annual cycle than did the all-hours-and-days data; rather than January being the lowest month, August was, whereas August was the highest in the all-hours-and-days cycle. November and December were high in both cycles. These differences made it inappropriate to use the 1961-1967 annual cycle used for the other two curves. Since the commuting-hour data showed much the same cycle each of the three years (except that the high was November rather than December in 1968) and since there was only trivial indication of effect, the three years of these data were averaged to get the annual cycle, which was then removed from the series.

In Figure 1, the middle line represents the resulting commuting-hours series. There is visibly no effect of the crackdown, nor does the Box and Tiao test show one, when applied to the series as graphed, or to an alternate way of removing the annual cycle. (The graphed approach would have some bias in the direction of minimizing the September-October 1967 shift.)

Ideally, this commuting-hour series would provide a control comparison against which we could decide whether or not the Breathalyser enforcement was being maintained or had abated. Insofar as it is relevant for this purpose, the crackdown had a maximum impact for the first three or four months and has leveled off since. But at the end of 1968 there was still a definite saving of some thirty percent in the weekend-night rates.

The appropriateness of the comparison is weakened by the dissimilarity shown in its annual cycle. Yet it is the nearest thing we have. If it is to be of value we need to do better than has been done here with the annual trend. Data subsequent to the crackdown continue to be collected, and four or five years from now we will have available a better estimate of the commuting-hour annual cycle.

THE THREATS TO VALIDITY

In the presentation of the Interrupted Time-Series design at the beginning of this paper, we listed six threats to validity. In the presentation so far of the British crackdown, we have paid attention primarily to the threat of instability—the only one, it should be remembered, to which tests of significance are relevant.

Reviewing the other threats, *maturation* seems out: the October 1967 drop is not plausibly interpretable as part of a general trend manifested prior to the crackdown. *Testing* and *instrumentation* seem unlikely: the procedures for recording and publicizing traffic casualties were well established prior to the crackdown and did not change on account of the crackdown. But this is not a trivial matter. The official categories of "seriously injured" and "slightly injured" obviously call for a judgment the threshold for which could change if the record-keepers were strongly motivated to make a good show. Crime rates, for example, have shown such fluctuations (Etzioni, 1968; Campbell, 1969: 415). In this regard it is comforting to note that for the all-days-and-hours figures, for which fatalities are separately available, they show as marked effects as do serious injuries. (In the crime studies cited, homicides and murders were markedly less susceptible to recording bias than were lesser crimes. See Campbell, 1969.) *Regression* seems implausible here, for, in marked contrast to the Connecticut case, the crackdown was not a reaction to a peak crisis, but rather to a chronic condition, as inspection of the series indicates.

There remains the catchall category labeled *history*—discrete events other than the experimental treatment that occurs simultaneously with them. In quasi-experimental thinking, when a set of hypotheses cannot be ruled out mechanically through design, the researcher bears the burden of seeking out the reasonable hypotheses included therein and ruling them out or allowing for them individually. The following explanations have been suggested as possible alternative or additional explanations of the change in the British casualty rate in October of 1967 (Bennett and Westwick, 1968).

(1) *The publicizing of crackdown.* The government conducted a two-phase publicity campaign concerning the crackdown, from September 25 through December 21, 1967. This large-scale effort involved several hundred thousand pounds spent for paid advertising, in addition to donations of large amounts of free time by public radio and television. The campaign publicized and explained the crackdown.

Although the publicity campaign may have helped the crackdown produce its effect—indeed, it may be considered as a part of the crackdown—the continued lower casualty rate is inconsistent with the idea that the publicity campaign acted independently. It seems reasonable to posit that the publicity campaign made the crackdown more effective, and to expect that the effect of the crackdown might be increased with additional publicity campaigns. An additional reason for doubting the independent effect of the publicity campaign is the known ineffectiveness of most safety publicity; a similar safety campaign conducted in Britain in 1964, on the same scale and with the same media as the 1967 campaign, had no notable effect on the casualty rate.

(2) *Improvements in traffic controls.* Within the past two or three years, there have been some important improvements in traffic control in Britain. For instance, the priority of vehicles at traffic circles has been resolved, and signs posted accordingly; "halt" and "yield" signs had been posted to control entry to major arteries; and intersections known to be dangerous had been reworked.

Perhaps part of the observed change in the casualty rate is due to these efforts, but the introduction of reforms in traffic control can best be conceived as a gradual program rather than as a sudden one, whereas the change in the data is abrupt. In addition, traffic signs would not be expected to have a greater effect at night than during the day.

(3) *Tire inspection.* New tires must now meet the standards of the British Standards Association. However, since the proportion of vehicles with new tires increases very gradually, the comments concerning the traffic control program apply here and rule out explaining much of the observed change in these terms.

(4) *Reduction in two-wheeled vehicles.* Motorcycles and motor scooters have a high accident rate; the number of these vehicles in use is alleged to have decreased very sharply in 1967, one estimate being as much as thirty percent (Bennett and Westwick, 1968). The reduction is said to be due to a temporary increase in the purchase tax on these vehicles, which was rescinded in 1968.

The factual basis of this explanation is challenged by statistics maintained by the Ministry of Transport showing that the use of motorcycles and motor scooters declined only about fourteen percent in 1967. This decline was part of a general, long-term decline in the use of these vehicles, and was about average in amount. A decline of this form is unlikely to produce an abrupt effect in a causally related variable. Just to be sure, we have examined all-hours-and-days figures separately for cars and for two-wheeled motor vehicles. The sharp October 1967 drop exists in both series, but is much more marked for four-wheeled cars.

(5) *Improvement in traffic law enforcement in London* has been suggested as a cause of the decline. Since there is no demonstrated sharp and direct relationship between law enforcement and accident rates, this explanation can be discounted.

(6) *Highway traffic* has grown less rapidly in Britain since 1965 than before that date. However, since growth has continued, albeit at a slower pace, an absolute decrease in the number of accidents does not seem reasonably explained by this fact. The actual volume of traffic in Britain increased by six percent in 1967.

(7) *British insurance companies* offer an enormous discount for claim-free driving. However, this is no innovation, and any effect that it might have on casualties would not be expected to follow the form of our data.

CONCLUSION

The Interrupted Time-Series design used in this study of the British crackdown on drinking and driving has ruled out a wide variety of potential alternative explanations of the observed decline in casualties. The only serious contenders to the hypothesis that the crackdown saved lives and injuries are a group of hypotheses each of which refers to a simultaneous event that might be expected to have a similar effect. However, close attention to each of these rules them out as plausible explanations of much of the change observed at the time of the crackdown. Our conclusion is that the crackdown, with its attendant publicity, did save lives and prevent injuries, and that it continues to have an important beneficial effect on British highways.

Substantively, we have shown that a relatively simple and inexpensive legal reform has produced the results for which it was intended. We believe that the British Act, with appropriate modifications, would meet the requirements of constitutionality in the United States; and although direct generalization is not possible, we can see no reasons why such action would not have a similarly beneficial effect in this country. Officials charged with responsibility for highway safety might well be urged to consider this adoption.

Methodologically, we have demonstrated a technique for evaluating the effect of social changes generally and legal changes in particular. This technique ought to be used more frequently than it is at present by both pure and applied social research. The student of society does not need experimental control to assess the effect of a change, providing he knows the limits of the techniques he uses and proceeds sensibly rather than mechanically. If the resulting knowledge is imperfect, the same problem applies in a slightly lesser degree to the best controlled laboratory experiments when one tries to generalize beyond the laboratory. In contrast, the ability to generalize to a large population outside the laboratory is inherent in this and other quasi-experimental techniques where the basic experiment itself is conducted in a similar field situation. Uniquenesses in such settings make it of course desirable to have replications and cross-validations.

The administrator who wants to adopt an innovation such as this should introduce it in such a way that its effectiveness can be reconfirmed in his own setting. For this purpose, where the Interrupted Times-Series is all that is feasible, rules should be kept in mind. First, an abrupt, strong, dateable point of impact should be sought, since gradual innovation cannot be distinguished from secular trends. Second, the available time-series records should be continued so as to preserve comparability. Third, the innovation should be introduced when the problem is at a chronic level, rather than in response to crisis. Fourth, the administrator should seek out control series, from adjacent political units or from subset data within his own polity.

NOTES

1. For these data, we are indebted to N. F. Digance and J. M. Munden, Directorate of Statistics, Ministry of Transport, London. Their help is gratefully

acknowledged.

2. This prorating procedure assumed that the rate for the first nine days of October 1967 was characteristic of October 1961-1966 and of the year 1967. The average October value was 9042. January-September of 1967 ran 1.058 times the average January-September 1.058 x 9042 = anticipated October 1967 of 9566, 8/31 of which is 2468. The actual total for October 1967 was 8269, of which we assume that 8269 - 2468 = 5801 occurred during October 9 to 31. Expanding 5801 for 23 days to a 31-day month produces 7814. (This is then corrected for seasonal trend by subtracting 588, to achieve the plotted point of 7226.) It is obvious here, in the monthly prorating to 31 days, and in the prorating of weekends below, that for a scientific or legislatively authoritative analysis, we should have been given access to records by days rather than by months. However, this was not feasible.

3. Their model assumes that the time-series is subjected to an influence at each time which tends to move the series up or down, and that in the long run these influences—if they could be examined individually—would follow a normal distribution. Though a new influence enters maximally at each point, the effect of the influence is felt on the series at points beyond its initial appearance. Thus the statistical model specificially takes into account the nonindependence of adjacent observations in a time-series. It is with respect to this typical nonindependence of real data that attempts to solve the problems of time-series analysis with simple regression models fail. Data which conform to the statistical model will not show regular periodic cycles. Since most systems which are partially affected by weather and other annual phenomena show yearly cycles, it is necessary to remove such cycles in the data before analysis. Subsidiary autocorrelation analyses verify the absence of cycles. Glass, Tiao, and Maguire (1970) have modified the model to allow for the data to show constant rates of "drift," increase or decrease, over time. It is this modified model which has been used here, applied to the total series, beginning January 1961. For all the likely values of the effect, the t values are 4.0 or larger, which indicates that the shift is of a magnitude that would occur by chance less than once in 10,000 similar series.

REFERENCES

ANDANAES, J. (1966) "The general preventive effects of punishment." Univ. of Pennsylvania Law Rev. 114 (March): 949-983.

Automobile Manufacturers Association, Inc. (1966) The State of the Art of Traffic Safety: A Critical Review and Analysis of the Technical Information on Factors Affecting Traffic Safety. Cambridge: Arthur D. Little.

BENNETT, R. O. and E. H. WESTWICK (1968) "A report on Britain's road safety act of 1967." Prepared for the Insurance Institute of Highway Safety.

BOX, G. E. P. and G. C. TIAO (1965) "A change in level of non-stationary time series." Biometrika 52: 181-192.

CAMPBELL, D. T. (1969) "Reforms as experiments." Amer. Psychologist 24 (April): 409-429.

--- and L. ROSS (1968) "The Connecticut crackdown on speeding: time-series data in quasi-experimental analysis." Law & Society Rev. 3 (August): 33-53.

CAMPBELL, D. T. and J. C. STANLEY (1966) Experimental and Quasi-Experimental Designs for Research. Chicago: Rand-McNally.

CAMPBELL, H. E. (1964) "The role of alcohol in fatal traffic 'accidents' and measures needed to solve the problem." Michigan Medicine 63 (October): 699-703.

ETZIONI, A. (1968) "Shortcuts to social change?" The Public Interest 12: 40-51.

GLASS, G. V (1968) "Analysis of the Connecticut speeding crackdown as a time-series quasi-experiment." Law & Society Rev. 3 (August): 55-76.

--- H. L. ROSS, and D. T. CAMPBELL (forthcoming) "Statistical analyses of the impact of the British road safety act of 1967."

GLASS, G. V, G. C. TIAO, and T. O. MAGUIRE (forthcoming) "Analysis of data on the 1900 revision of the German divorce laws as a quasi-experiment." Law & Society Rev.

Insurance Institute for Highway Safety (1968) Highway Safety, Driver Behavior: Cause and Effect. Washington.

RHEINSTEIN, M. (1959) "Divorce and the law in Germany: a review." Amer. J. of Sociology 65: 489-498.

ROSE, A. M. (1952) "Needed research on the mediation of labor disputes." Personnel Psychology 5: 187-200.

SCHUESSLER, K. F. (1969) "The deterrent influence of the death penalty," pp. 378-388 in W.J. Chambliss (ed.) Crime and the Legal Process. New York: McGraw-Hill.

STIEBER, J. W. (1949) Ten Years of the Minnesota Labor Relations Act. Minneapolis: Industrial Relations Center, University of Minnesota.

U.S. House Committee on Public Works (1968) 1968 Alcohol and Highway Safety Report. Washington: U.S. Government Printing Office.

WALKER, N. (1965) Crime and Punishment in Britain. Edinburgh: Edinburgh Univ. Press.

WEBB, E. J. et al. (1966) Unobtrusive Measures: Nonreactive Research in the Social Sciences. Chicago: Rand-McNally.

WOLFF, E., G. LÜKE, and H. HAX (1959) Scheidung und Scheidungsrecht: Grudfrägen der Ehescheidung in Deutschland. Tübingen: J.C.B. Mohr.