ANALYSIS OF DATA ON THE CONNECTICUT SPEEDING CRACKDOWN AS A TIME-SERIES QUASI-EXPERIMENT

GENE V GLASS

Laboratory of Educational Research,
University of Colorado

IN LATE 1955 IN CONNECTICUT, the number of fatalities per 100,000 population in motor vehicle accidents reached a record high for the 1950s. On December 23, 1955, Governor Abraham Ribicoff took unprecedented action to reduce traffic fatalities. Ribicoff announced that persons convicted of speeding would have their licenses suspended for thirty days at the first offense, for sixty days at the second offense, and for an indefinite period (subject to a hearing after ninety days) at the third offense. Data on traffic fatalities before and after the Connecticut crackdown on speeding can be regarded as a time-series quasi-experiment[1] with some significance for the social sciences. When supplemented with traffic fatality data for the states of Massachusetts, Rhode Island, New York, and New Jersey, the collection of observations can be viewed as

AUTHOR'S NOTE: *The work reported in this paper was supported as Project No. 6-8329 of the U. S. Office of Education. The author is indebted to Donald T. Campbell of Northwestern University and H. Laurence Ross of the University of Denver Law School who supplied the data which are analyzed in this paper.* See *D. T. Campbell & H. L. Ross,* The Connecticut Crackdown on Speeding, 3 LAW & SOC. REV. 33-53 (1968).

1. D. T. Campbell & J. C. Stanley, *Experimental and Quasi-Experimental Designs for Research on Teaching,* in HANDBOOK OF RESEARCH ON TEACHING (N. L. Gage ed. 1963).

D. T. Campbell, *From Description to Experimentation: Interpreting Trends as Quasi-Experiments,* in PROBLEMS IN MEASURING CHANGE, ch. 12 (C. W. Harris ed. 1963).

a multiple-group time-series experiment.[2] The multiple-group time-series design can be diagrammed as follows:

	TIME				
Place	...$n_1 - 1$	n_1		$n_1 + 1$	$n_1 + 2$...
Connecticut	...0	0	T	0	0...
Massachusetts	...0	0		0	0...
Rhode Island	...0	0		0	0...
New York	...0	0		0	0...
New Jersey	...0	0		0	0...

The 0's represent monthly observations of traffic fatalities for the n_1 months prior to T, the treatment, and for the n_2 observations following T. The treatment, T, is the Governor's crackdown on speeding in the state of Connecticut. No comparable alteration of administrative practice took place in the four "control" states.

Evidence of the effectiveness of the Connecticut crackdown on speeding can be gained by comparing the path of the post-T observations of Connecticut with those of the four control states. A precipitous drop in fatalities in Connecticut following T in the absence of similar drops in the control states is compelling evidence of the effectiveness of the crackdown on speeding.

The problem of measuring the abrupt change in level of a time-series and making statistical inferential statements about it is the problem with which the remainder of this report is concerned.

ANALYSIS OF DATA

The Underlying Model

The statistical model upon which analysis of the Connecticut speeding data is based was developed by Box and Tiao.[3] Box and Tiao presented an analytic technique for estimating and making inferences about the change in level of a time-series. The model upon which the analysis is based is a restrictive one; however, many sets of data can be manipulated (by removing cycles, for example) or transformed into

2. *Id.*

3. G. E. P. Box & G. C. Tiao, *A Change in Level of a Non-stationary Time-Series*, 52 BIOMETRIKA 181-92 (1965).

special indices in such a way that the assumptions of the model will be largely met. The statistical model here employed is a special case of the integrated moving average process:[4]

$$z_t = L + \gamma \sum_{j=1}^{t-1} a_{t-j} + a_t, \quad t = 1, \ \ldots, \ n_1 . \qquad (1)$$

L is a "location parameter" descriptive of the overall general level of the series,

γ is a parameter which depends upon the interdepedency of the observations in the time-series, and

a_t is an observation of a random normal variable with mean 0 and variance σ^2.

Formula (1) describes the n_1 observations taken prior to the introduction of a treatment, *e.g.*, the Connecticut crackdown on speeding. The n_2 observations following the introduction of the treatment into the time-series differ from (1) only in that a treatment effect, δ, is present.

$$z_t = L + \gamma \sum_{j=1}^{t-1} a_{t-j} + a_t + \delta, \quad t = n_2, \ \ldots, \ n_1 + n_2 . \qquad (2)$$

The parameter δ is the increment or decrement in the level of the time-series due to the introduction of the treatment. The treatment is assumed to work an immediate and constant effect, δ, upon the time-series.

The fundamental time-series model regards the system as being subjected to periodic random shocks, the a_t (which have zero mean). Furthermore, a proportion, γ, of each shock is assumed to remain in the system to influence the movement of the system through time. Hence, the effect of some extraneous, random influence on the system is not immediately dissipated but continues to work a lessened influence on subsequent observations.

The objective of a statistical analysis is to estimate the value of δ in (2) which is the effect of the introduction of the new law at time $n_1 + 1$ and to determine whether this estimate indicates that the true value of δ—which is unobservable with real, fallible data—is positive,

4. G. E. P. Box & G. M. Jenkins, *Some Statistical Aspects of Adaptive Optimization and Control*, 24 J. ROYAL STATISTICAL SOC'Y B, 297-343 (1962).

negative or zero, *i.e.*, whether the introduction of the law increased, decreased or left unchanged the course of the behavioral index.

Before any analyses of data may proceed, an effort must be made to check the appropriateness of the model in (1) and (2) for the data in hand. This can be done in large part by inspecting the graph of the time-series and of the correlograms[5] of the data. Data which conform to the model in (1) and (2) will have the following properties:

1. There will be an absence of cycles in the graph of the time-series of the data. In addition, the data will appear to fluctuate with only minor or momentary unsystematic drifts away from a general elevation. In other words, sustained "drifts" from a baseline in one direction probably indicate a violation of the model.

2. The correlogram of the original data, the z_t, is free of cycles and shows a random fluctuation around a baseline. (A correlogram is a set of "lag correlations" of the data. The lag 1 correlation for a time-series is the correlation calculated on pairs of observations formed by pairing each observation, z_t, with the observation, z_{t+1}, which follows it by *one* unit of time. The "lag 2" correlation is the correlation calculated on the pairs of observations z_t and z_{t+2} for $t = 1, \ldots, n_1 + n_2 - 2$. The "lag k" correlation is between the pairs z_t and z_{t+k}.)

3. The correlogram of the differences between successive observations in the time-series, *i.e.*, the correlogram for $z_t - z_{t-1}$ ($t = 2, \ldots, n_1 + n_2$) has a lag 1 correlation which is large in absolute value when γ deviates from 1.0 and all higher lag correlations are near zero. In fact, the lag 1 correlation of the differences must be

$$\frac{-(1-\gamma)}{1+(1-\gamma)^2},$$

where γ is an unknown parameter of the model in (1). For example, if $\gamma = 1$, all lag correlations of the differences between successive values are expected to be zero. Fortunately, approximate hypothesis tests are available for testing the significance of the lag correlations.[6]

5. Correlograms are sets of correlation coefficients obtained by pairing the observations in a time-series in different ways and calculating the correlation coefficients that result.

6. M. S. Bartlett, *On the Theoretical Specification and Sampling Properties of Autocorrelated Time-Studies,* 27 J. ROYAL STATISTICAL SOC'Y B (1946).

Investigation of the Fit of the Model to the Data

The basic data were traffic fatalities for the sixty months prior to the Connecticut speeding crackdown in January 1956 and for the subsequent forty months for Connecticut, Massachusetts, Rhode Island, New York, and New Jersey. As the first step in the investigation of the fit of the model in (1) to these data, each monthly fatalities count was divided by the number of miles driven in the state during that month. The transformed raw data thus became "monthly fatalities per 100,000,000 miles driven" for all five states. Such a transformation would effectively eliminate any upward (or downward) trend in the data (no such trend is allowed to appear in the model in equation [1]) due to increases in population, number of drivers, number of cars, etc.

Inspection of the plot of "monthly fatalities per 100,000,000 driver miles" showed marked yearly cycles, as one might expect. The "peaks" of the cycles coincided with the winter months (Dec.-Feb.); the "valleys" occurred during the summer. Such cycles are a clear violation of the assumptions of the model. (The correlogram for "fatalities/100,000,000 miles" for Connecticut showed the "damped sine curve" with a period of twelve months which is characteristic of data possessing yearly cycles. The manner in which the cycles were removed from the data will be discussed later.) It will be instructive for the moment to observe the "monthly fatalities per 100,000,000 driver miles" with the cycles left in. These data for Connecticut appear in Figure 1.

It can be seen in Figure 1 that the fatalities per 100,000,000 driver miles reached the highest point in the period 1951-1955 in December 1955. To the extent that this "emergency" prompted Ribicoff's decision to crack down on speeding in late December 1955, the decline immediately following the crackdown can be partly interpreted as the natural tendency of observations chosen for their extremity to "regress" toward a central value.

There is a marked decrease in fatalities per 100,000,000 driver miles from December 1955 to January 1956. However, there are also decreases in fatalities/100,000,000 miles in six of the eight possible comparisons of a December with the immediately following January for the data in Figure 1. In fact, the drop in fatalities/100,000,000 miles from December 1957, to January 1958, is almost equal to the drop from December 1955, to January 1956. A natural drop from any December to the immediately following January in fatalities/100,000,000 miles is quite apparent in Figure 1. Such cycles are also obvious in the graphs of monthly fatalities per 100,000,000 miles in the four "control" states.

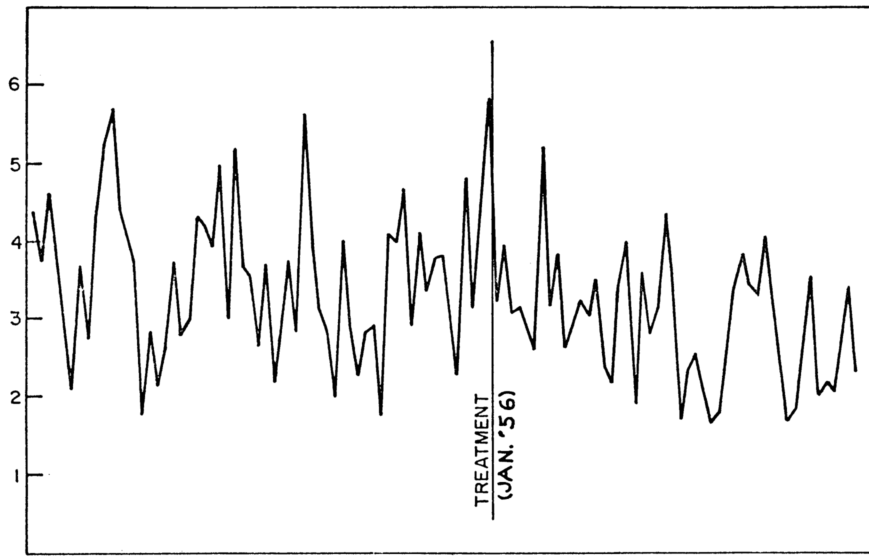

Figure 1. Fatalities/100,000,000 Driver Miles by Months for Connecticut ($n_1 = 60$, $n_2 = 48$)

The following technique was employed to remove the cycles from the data. Since the cycle had a period of twelve months, the average fatalities/100,000,000 miles for each of nine Januaries (1951-1959) was subtracted from each January observation. Similarly, observations on each of the other eleven months were deviated around the average (over nine monthly values) fatalities/100,000,000 miles for that month. This was done for each of the five states. (A constant, 2 or 3, was then added to these transformed scores to make them all positive for convenience in recording.)

These transformed data showed neither apparent cycles nor upward or downward trends. The data for all five states appear in Figures 2-6. In this form, the data appear to satisfy the first condition of equation (1). The next step in the examination of the fit of the model to the data involves the correlograms of the observations in Figures 2-6 and the correlograms of differences between adjacent observations in the series.

Correlograms were calculated on the data in Figures 2-6 for pre-January 1956 ($n_1 = 60$) and post-January 1956 ($n_2 = 48$) data separately. (A marked change in level of a time-series due to a treatment effect would alter the autocorrelations from what they would be in the fundamental process which generates the observations in the time-series; hence, in judging the fit of a model to data from a time-series experiment, correlograms must be calculated separately for pre- and post-treatment observations.) To conserve space, these correlograms are not reproduced here. Each correlogram appeared to be no more than a random array of nonsignificant autocorrelations characteristic of the correlogram to be expected from data conforming to the model in equation (1).

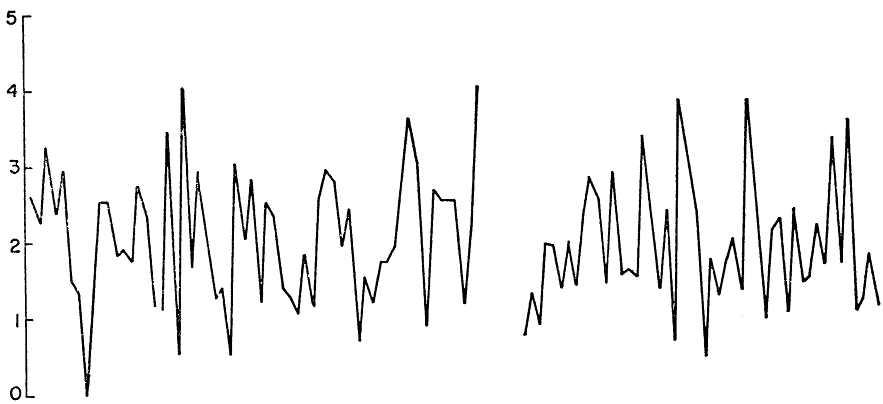

Figure 2. Connecticut Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus 2

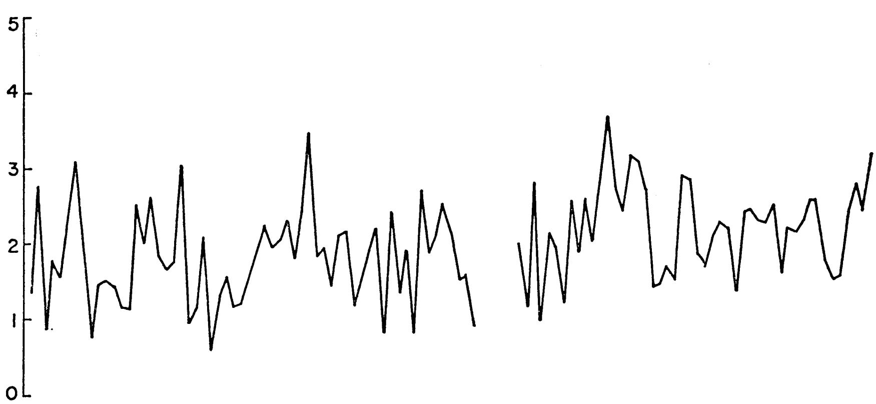

Figure 3. Massachusetts Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus 2

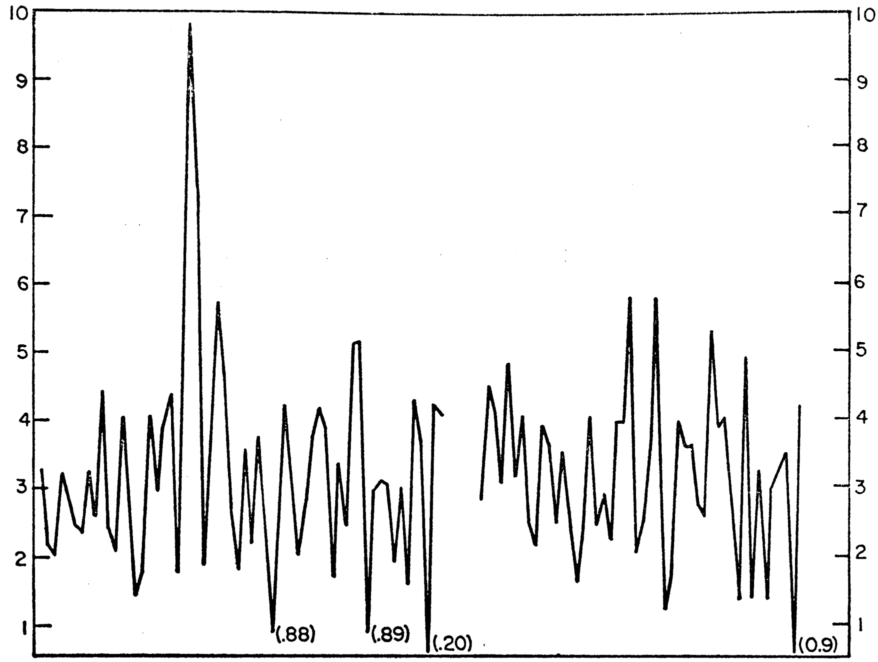

Figure 4. Rhode Island Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus 3

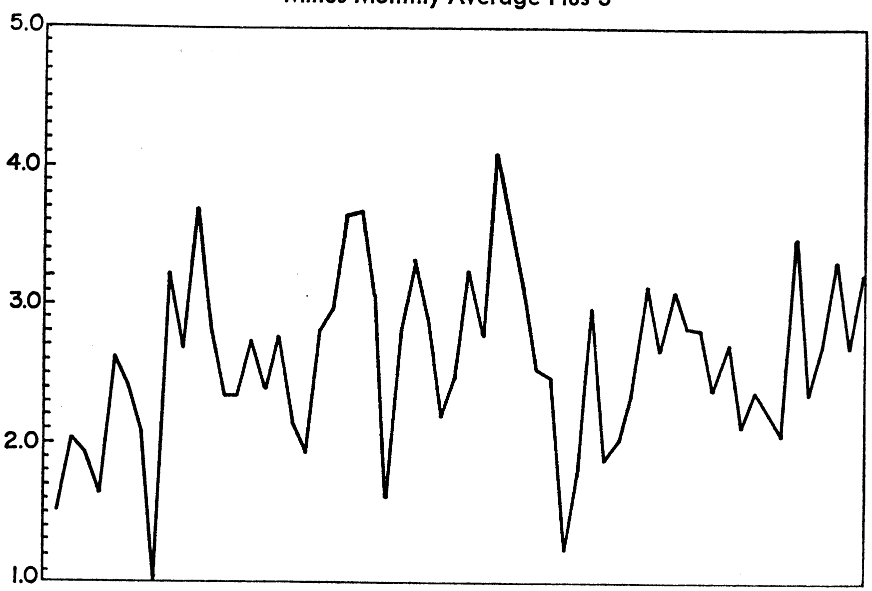

Figure 5. New York Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus 3 ($n_1 = 60$, $n_2 = 48$)

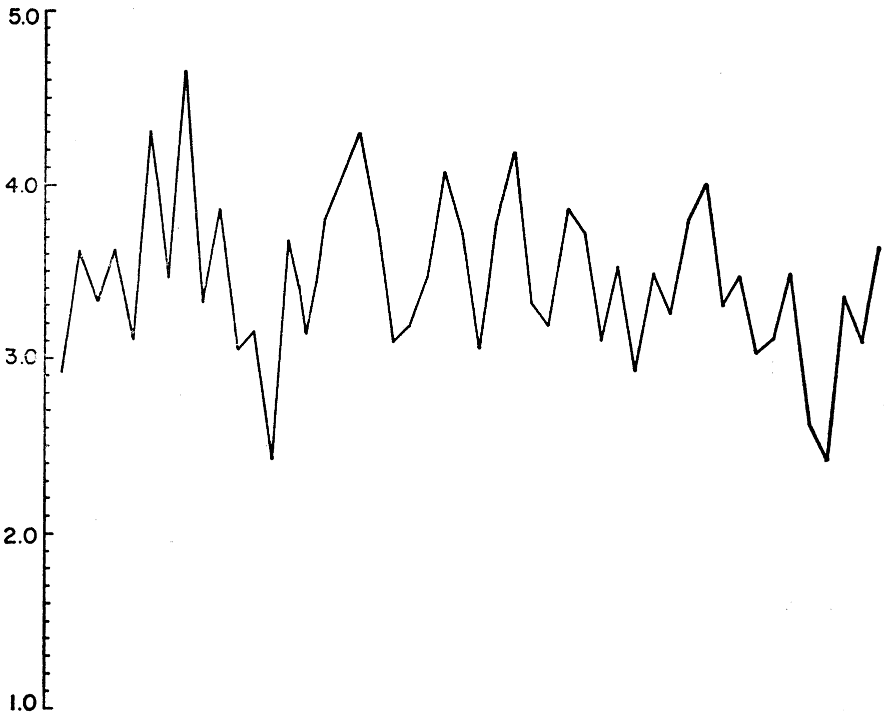

Figure 5 (continued). New York Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus 3 ($n_1 = 60$, $n_2 = 48$)

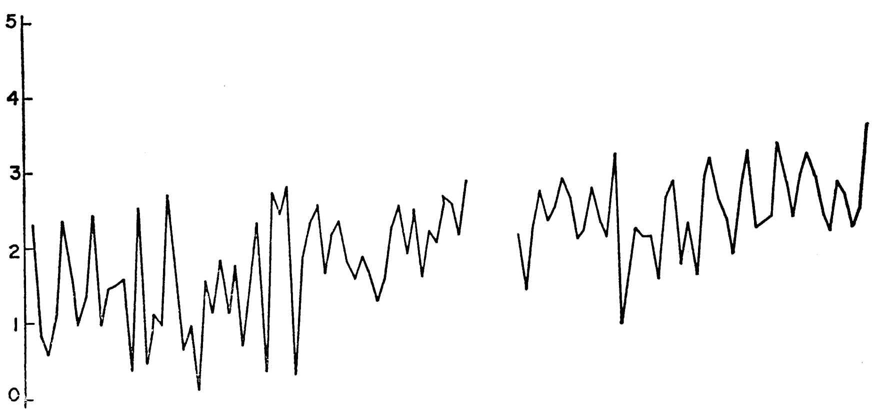

Figure 6. New Jersey Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus 2

The next step in the investigation of the fit of the model in (1) to the data is to calculate the correlogram for the differences between adjacent observations, $z_t - z_{t-1}$. It is necessary to calculate these differences separately for the pretreatment and posttreatment data. Only the correlograms for the sixty pretreatment observations for each state are examined here. As was pointed out earlier, if the model in (1) is satisfied, the lag 1 autocorrelation of the differences $z_t - z_{t-1}$ will equal $-(1-\gamma)/[1+(1-\gamma)^2]$, where γ is an unknown parameter in the model, and the lag 2 and greater autocorrelations of the same data will equal zero. Not knowing γ, it is necessary to obtain an estimate of it. Later it will be seen how the most likely values of γ can be found from the N observations, z_t. The most likely values of γ for the time-series in each of the five states were found to be the following:

State	Most Likely Value of γ	Corresponding Expected Lag 1 Correlation of $z_t - z_{t-1}$ if Model in (1) Holds	Obtained Lag 1 Correlation of $z_t - z_{t-1}$
Connecticut	.01	—.50	—.555
Massachusetts	.01	—.50	—.500
Rhode Island	.01	—.50	—.395
New York	.16	—.49	—.551
New Jersey	.11	—.49	—.612

In light of the above data, the correlograms for the fifty-nine observations of $z_t - z_{t-1}$ for each state should present a lag 1 correlation of approximately —.5 and lag 2 and greater correlations which differ insignificantly from zero. The first such correlogram—for the Connecticut data—appears as Figure 7. The jagged line in Figure 7 is the plot of the lag 1 through lag 30 autocorrelations for the 59 pretreatment observations $z_t - z_{t-1}$ for Connecticut. The lag 1 autocorrelation of —.555 agrees quite closely with the expected value of —.50. Superimposed upon the graph of the correlogram are two curved lines indicating those points which lie two standard deviations from the mean, zero, in the distribution of the lag k autocorrelation coefficient for samples of size 59 from a population in which the coefficient is zero.[7] Only the lag 1 autocorrelation coefficient is significantly different from zero in Figure 7; hence, the conditions of the model—as reflected in the correlogram of $z_t - z_{t-1}$—appear to be met by the Connecticut data.

7. *Id.*

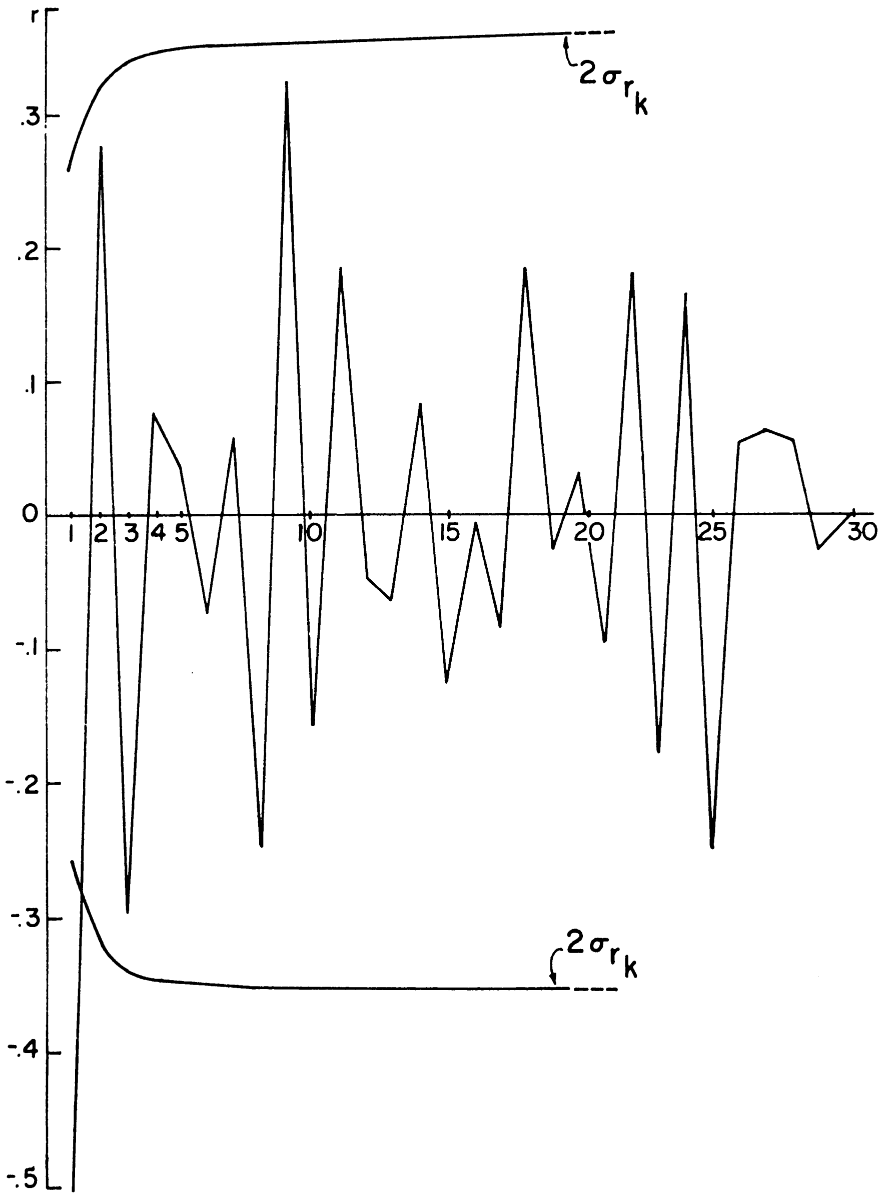

Figure 7. Correlogram of Differences for Pre-Treatment ($n_1 = 60$) Monthly Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus 2 for Connecticut

https://doi.org/10.2307/3052795 Published online by Cambridge University Press

The correlograms (lag 1 through lag 20) for Rhode Island, Massachusetts, New York, and New Jersey for the fifty-nine pretreatment observations $z_t - z_{t-1}$ appear in Figure 8. None of the lag 1 autocorrelation coefficients differs appreciably from the expected values of $-.50$ and $-.49$. The lag 2 and greater autocorrelations are distributed around zero with only three coefficients (viz., lags 18 and 19 for Massachusetts and lag 4 for New Jersey) lying further than two standard errors from zero. (The curved lines marking off two standard errors in the distribution of the autocorrelation coefficients which appear in Figure 7 can be applied to the data in Figure 8 as well.) The import of the data in both Figures 7 and 8 is that the conditions of the model in equation (1) which are reflected in the correlograms as $z_t - z_{t-1}$ are reasonably satisfied by the data for the five states.

After transformation of the data and removal of cycles, the data on fatalities for the five states give no evidence of not satisfying the conditions of the model in equation (1). We shall proceed with the analyses assuming the data are adequately described by such a model.

Analysis for Change in Level of the Five Time-Series

First, we shall consider in turn the individual analyses for changes in level between the 60th and 61st months of the five time-series in Figures 2-6. The analysis of the Connecticut data (Figure 2) will be considered in detail. Summaries of the analyses will be presented for the other four states. After consideration of the individual analyses, the five sources of data will be combined into a single analysis comparing Connecticut with the "control states." The objective of the statistical analysis is to estimate the size of δ, the effect (increment or decrement) of the introduction of the law at time n_1 on the time-series, and to decide whether the true value of δ is positive, negative, or zero. In effect, then, the statistical analysis answers the question whether the observations following the enactment of the law are simply a continuation of the time-series of the preenactment observations or whether they have shifted up or down from the general level of the preenactment time-series. It can be shown that (when one has as many as ten observations in time prior to enactment of the law and ten after) the estimate of the effect (increase or decrease) of the law on the quantitative index being observed is given approximately by the difference between exponentially weighted averages of the observations closely following the enactment of the law and the observations closely preceding the enactment. For example, consider a time-series experiment in which there are $n_1 = 50$

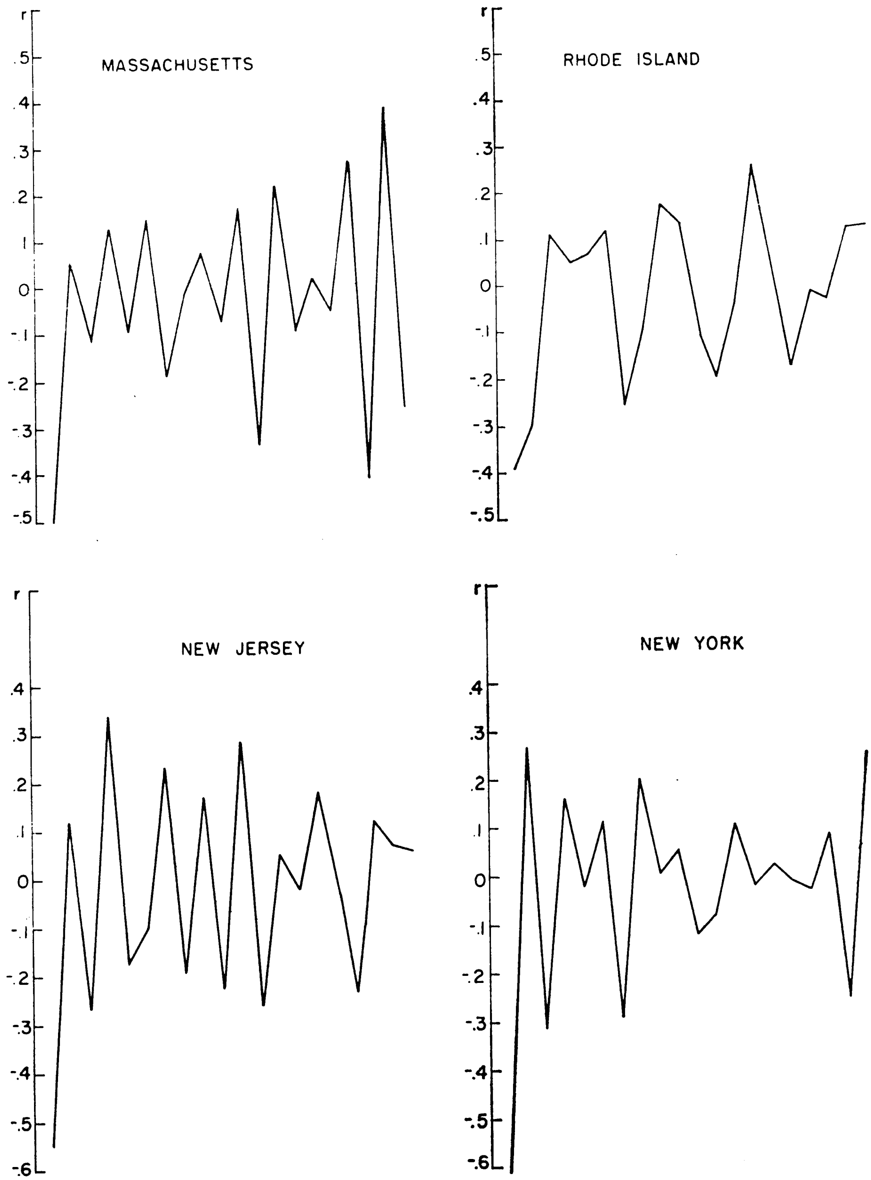

Figure 8. Correlograms (Lag 1 — Lag 20 auto correlations) of Differences, $z_t - z_{t-1}$, for Pre-Treatment Data ($n_1 = 60$) for Mass., R. I., N. J., N. Y.

(See Figures 3-6 for the data which are "differenced" and correlated)

points in time preceding the enactment of the new law and $n_2 = 50$ points following. In some instances (namely when $\gamma = .50$) the estimate of the effect of the law on the index being observed would be given by

$$\hat{\delta} = \ldots + .016z_{55} + .125z_{54} + .25z_{53} + .5z_{52} + z_{51} - z_{50} - .5z_{49} - .25z_{48} - .125z_{47} - .016z_{46}. \ldots$$

Thus the estimate of the effect of the treatment is the sum of weighted posttreatment observations minus the sum of weighted pretreatment observations where the weights diminish as the observations are further removed from the point at which the treatment is introduced. Notice that the weight given to the two observations closest to the treatment is 1 which equals $(.5)^0$. The weights diminish by a factor of .5 for every step taken away from the point at which the treatment was introduced, *e.g.*, four steps away from the treatment the observations are weighted by $(.5)^3 = 0.125$. Observations more than six steps away from the treatment receive virtually no weight in determining δ in this instance. This indicates that the further observations are removed from the treatment, the more their course comes under the influence of random influences irrelevant to the effect of the treatment.

The analysis is complicated by the fact that one of the parameters, γ, of the model in equation (1) is unknown and must be estimated from the data in hand. This parameter is crucial to the determination of which variation of the model adequately describes the data and what postenactment behavior of the time-series should be expected. It is possible to inspect the data in hand (such as that in Figure 2) and determine a distribution of probabilities that the unknown parameter, γ, takes on any particular value between 0 and 2, its possible limits. This "likelihood distribution" is denoted by $h(\gamma|z)$ and can be inspected to determine either the most likely value (the peak of the $h[\gamma|z]$ curve) of γ, called the "maximum likelihod estimate of γ," or a range of likely values of γ. For example, the maximum likelihood estimate of γ for the data in Figure 2 is 0; note that the curve $h(\gamma|z)$ peaks at a value of 0 in Figure 9, which reports the analysis of the data in Figure 2. Thus it appears likely that the model underlying the pretreatment observations in Figure 2 is

$$z_t = L + (0) \sum_{i=1}^{t-1} a_{t-i} + a_t = L + a_t .$$ [8]

8. In this instance, *i.e.*, $\gamma = 0$, the model is equivalent to the model for the "*t*-test" employed in elementary statistics, and the analysis is equivalent to a *t*-test on pretreatment v. Posttreatment observations.

A complete analysis involves analyzing the data for every possible value of γ, these being all numbers between 0 and 2; and, for each value of γ between 0 and 2, observing the value of $h(\gamma|z)$—the likelihood that a certain value of γ is the true one—and of the "t-statistic"—which tells how likely it is that δ is zero. Such an analysis is illustrated in Figure 9, in which is reported the analysis of the data in Figure 2. The curve for the likelihood of γ, $h(\gamma|z)$ and the curve of the t-statistic appear there. The t-statistic is used to measure the probability that δ, the effect of the crackdown on speeding, is zero. (The value of t is equal to the estimated value of δ from the data—which is denoted by δ —minus an hypothesized value of δ of zero divided by the estimated standard deviation of δ.) When the value of the t-statistic deviates from zero by two units, *i.e.*, falls below -2 or above $+2$, the evidence in favor of a nonzero δ is great. In most instances, a t-statistic above $+2$ or below -2 has only about 1 chance in 20 of occurring if there is no effect, δ, of the enactment of the legislation. Even though a t-statistic above $+2$ or below -2 *may* arise when δ is zero, there is only about 1 chance in 20 that it will. The primary reason why "large" (lying more than two units from zero) values of t occur is that δ is not zero, *i.e.*, that the introduction of the law at time n_1 abruptly shifted the level of the time-series either up or down.

Since calculating the estimated treatment effect, δ, involves the unknown parameter γ (which has possible values 0 to 2), δ and the resulting t-statistic are calculated for every value of γ from 0 to 2. Independently, the values which γ is likely to be are found by inspecting $h(\gamma|z)$. Putting the two together, one observes whether the values of t associated with the likely values of γ deviate more than two units from zero (in which case δ is thought to be *not* equal to zero) or whether they lie close to zero (which would support the conclusion that δ is zero). This method of analysis is facilitated by graphing $h(\gamma|z)$ and t against γ on the same graph, as is done in the sections to follow.

Analysis for Change in Level of the Connecticut Data (Figure 2)

The $n_1 = 60$ observations preceding the crackdown on speeding in Connecticut and the $n_2 = 48$ post-crackdown observations were subjected to the analysis outlined in Box and Tiao (1965) for unknown γ. The likelihood distribution of γ given the 108 observations is denoted by $h(\gamma|z)$ in Figure 9. The area under the curve $h(\gamma|z)$ is one unit.

The maximum likelihod estimate of γ is seen to be 0. The curve denoted by t in Figure 9 is the value of $\hat{\delta}/\hat{\sigma}(\hat{\delta})$—read off the right ordinate in the figure—for each value of γ from 0 to 2.

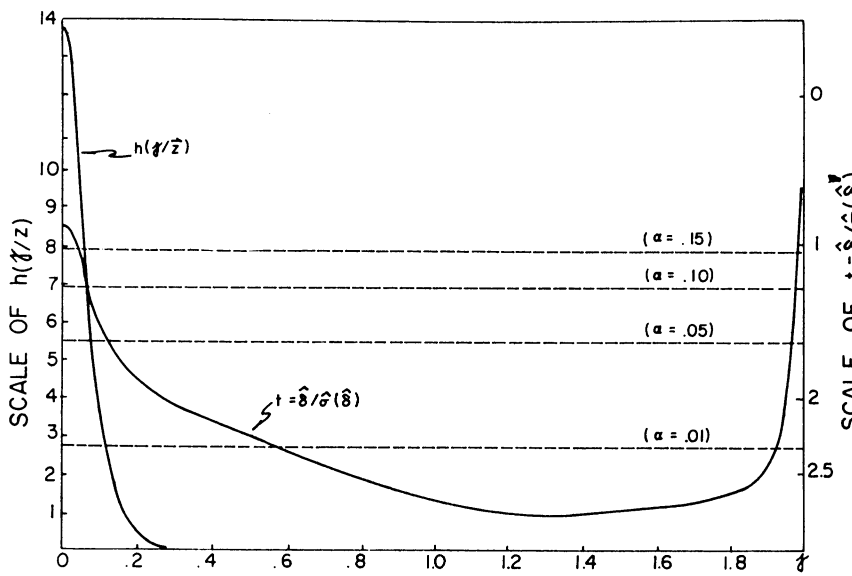

Figure 9. h(γ|z) and *t* for Connecticut Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus 2 ($n_1 = 60$, $n_2 = 48$)

As can be seen by inspection of the two curves in Figure 9 almost all the mass of the likelihood distribution of γ lies between 0 and 0.25, the former being the maximum likelihood estimate and the latter being quite unlikely; over this range (0 to 0.25) the value of the t-statistic for testing the hypothesis that $\delta = 0$ ranges from -0.86 to -2.05 (from about 1 chance in 3 that δ is zero to 1 chance in 20 that δ is zero). If γ is set equal to its maximum likelihood estimate, namely 0, the corresponding t of -0.86 indicates that δ is zero.[9]

The analysis reported in Figure 9 will support neither a confident acceptance nor rejection of the hypothesis that δ is zero, *i.e.*, that the crackdown on speeding had no effect on the fatality rate. The analysis proved sensitive to the unknown value of γ.

9. Inspection of the graphs is facilitated by the dotted lines which mark off the values of t ($df = 106$) required for significance at the .01, .05, .10 and .15 levels for a one-tailed test of the hypothesis that $\delta = 0$. For the four control states the alternative hypothesis is that $\delta > 0$.

Analysis for Change in Level of the Massachusetts, Rhode Island, New York, and New Jersey Data

In Figure 10, the likelihood distributions and t-statistics for testing whether δ can be considered zero are presented for the four "control" states. In all analyses, the likely values of the unkown parameter γ fall below .30. The maximum likelihood estimates of γ are .01 for both Massachusetts and Rhode Island. For New York and New Jersey, the maximum likelihood estimates of γ are .16 and .11, respectively.

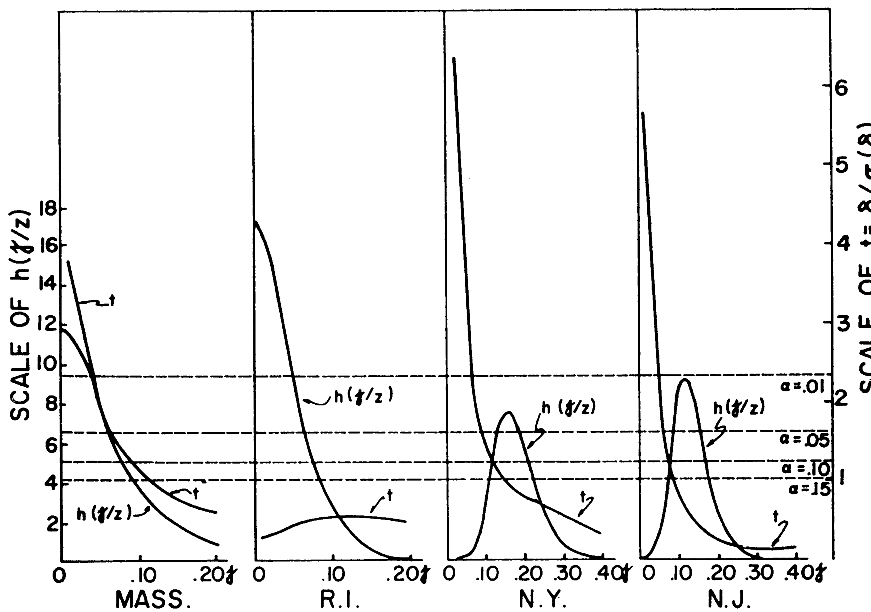

Figure 10. $h(\gamma|z)$ and t for Fatalities/100,000,000 Driver Miles Minus Monthly Average Plus a Constant for Mass., R. I., N. Y., and N. J.

Considering only the value of t for the maximum likelihood estimates of γ, the Massachusetts data yield the only value of $\hat{\delta}$ which differs significantly from zero. In fact, if $\hat{\delta}$ were zero for Massachusetts a t-statistic as large or larger than that obtained for the Massachusetts data would occur less than one time in 100. The t-statistics for Rhode Island, New York, and New Jersey do not attain statistical significance.

Considering the value of t over the ranges of likely values of γ, neither the Rhode Island, New York, nor New Jersey data present any evidence for a value of $\hat{\delta}$ significantly different from zero. The results for the Massachusetts data are equivocal. At the maximum likelihood estimate of γ, t is between 2.3 and 1.6. The value of t drops below $+1$ above the point on the γ-scale above which lies approximately 25 per cent of the area under $h(\gamma|z)$.

None of the analyses of the four control states yields compelling evidence of any abrupt change in fatality rate associated with the events in that state immediately prior to January 1956. The evidence ranges from definitely not supporting the presence of abrupt change in the case of Rhode Island to slightly equivocal as evidence for an abrupt change in the case of Massachusetts.

The Analysis of a Planned Comparison of Connecticut With the Control States

An inspection of the analyses of the control states' data showed that the estimated treatment effects, δ, are all positive,[10] though statistically insignificant. The estimated effect of the tightening of the law in Connecticut was negative, indicating a slight decrement in the fatalities rate. This state of affairs suggests that the slight decrement in the fatalities time-series for Connecticut after the 60th month should be viewed in light of the slight tendency of the fatalities rate to rise in the four control states. Essentially, we wish now to compare δ for Connecticut with the average δ for the four control states.

If the time-series for each state can be regarded as independent of the others, well-known inferential statistical techniques can be employed in making comparisons between Connecticut and the four control states. Accordingly evidence was sought concerning the degree of dependence among the time-series for the different states.

Given the normality assumption of the model in (1) and (2), the independence of the various time-series can be demonstrated if the series show no intercorrelation. To reduce the burden of data analysis without a serious reduction in the sensitivity of the test of the hypothesis of no intercorrelation, data for the first fifty months for Connecticut, Massachusetts, and New York were used. Using "months" as the unit across which correlations were computed, the three intercorrelations of these states were computed for the variable "fatalities/100,000,000 miles minus monthly average." The intercorrelation matrix was as follows:

10. Bear in mind that no "treatment" was actually applied in these states.

	Connecticut	New York	Massachusetts
Connecticut	1	—.105	—.061
New York	—.105	1	—.207
Massachusetts	—.061	—.207	1

A test was made of the hypothesis that the fifty triplets of observations were a random sample from a tri-variate normal distribution in which all intercorrelations are zero;[11] the test supported the plausibility of the hypothesis of zero correlations.

A single planned comparison will serve to evaluate the significance of the change in level of the time-series for Connecticut as compared to the changes or lack thereof in the four control states. This comparison has the following form:

$$\psi = \delta_C - (\delta_M + \delta_{RI} + \delta_{NY} + \delta_{NJ})/4$$

The value of ψ is estimated by replacing the parameters with their least-squares estimates; the variance of the comparison is estimated from the common residual variance for the five states multiplied by $[1^2 + 4(\frac{1}{4})^2]$. The estimated value of ψ divided by an estimate of its standard deviation is a *t*-statistic.[12] However, since the estimated change of level effects and residual variances differ for different values of γ, we shall estimate and test the significance of the comparison for the maximum likelihood estimates of γ for each state and for reasonable upper and lower limits to the value of γ for each state.

The hypothesis to be tested is that $\psi = 0$, *i.e.*, that the "shift" in level of the Connecticut time-series is equal to the average "shift" in the four control states, against the alternative hypothesis that $\psi < 0$, *i.e.*, that the shift in the fatality rate in Connecticut at January 1956 is less than the average shift of the four control states.

Because the γ's are unknown, we shall specify a range between which each γ probably lies as well as the maximum likelihood estimate of each γ. The lower limit to the range for each state will be that value of γ below which approximately 25 per cent of the area under the likelihood distribution of γ lies; the upper limit will mark off approximately the upper 25 per cent of the likelihood distribution. The data for estimating and testing the comparisons appear in Table 1.

Note in Table 1 that for the maximum likelihood estimate of γ, it is estimated that the fatalities/100,000,000 miles rate showed a *downward*

11. M. S. Bartlett, *Tests of Significance in Factor Analysis*, 3 BRIT. J. PSYCHOLOGY, statistical section, 77-85 (1950).

12. W. L. Hays, STATISTICS FOR PSYCHOLOGISTS, ch. 14 (1963).

shift of .152 fatalities/100,000,000 miles coincident with the crackdown on speeding.

TABLE 1

VALUES OF $\hat{\delta}$ AND $\hat{\sigma}(\hat{\delta})$ FOR THE MAXIMUM LIKELIHOOD ESTIMATE AND REASONABLE UPPER AND LOWER LIMITS OF γ FOR ALL FIVE STATES

State		Maximum Likelihood Estimate of γ	Reasonable Upper Limit for γ	Reasonable Lower Limit for γ
1. Connecticut	γ	.01	.10	.01
	$\hat{\delta}$	—.152	—.594	—.152
	$\hat{\sigma}(\hat{\delta})$	.176	.391	.176
2. Massachusetts	γ	.01	.15	.01
	$\hat{\delta}$	.472	.259	.472
	$\hat{\sigma}(\hat{\delta})$	.126	.341	.126
3. Rhode Island	γ	.01	.10	.01
	$\hat{\delta}$	.079	.326	.079
	$\hat{\sigma}(\hat{\delta})$	.276	.617	.276
4. New York	γ	.16	.25	.10
	$\hat{\delta}$	.275	.247	.337
	$\hat{\sigma}(\hat{\delta})$	.289	.375	.233
5. New Jersey	γ	.11	.20	.07
	$\hat{\delta}$	.198	.093	.331
	$\hat{\sigma}(\hat{\delta})$	.292	.391	.236

TABLE 2

RESULTS OF PLANNED COMPARISONS OF $\hat{\delta}$ FOR CONNECTICUT WITH THE AVERAGE $\hat{\delta}$ FOR MASSACHUSETTS, RHODE ISLAND, NEW YORK, AND NEW JERSEY

	Maximum Likelihood Estimate of γ	Reasonable Upper Limit for γ	Reasonable Lower Limit for γ
$\hat{\psi}$	—.408	—.825	—.457
$\hat{\sigma}_{\hat{\psi}}$	.269	.484	.241
$t = \hat{\psi}/\hat{\sigma}_{\hat{\psi}}$	—1.517	—1.705	—1.896
Prob $[t_{530} < t]$	.065	.045	.030

For a given set of five values of γ (one for each state, ψ is estimated by subtracting the average $\hat{\delta}$ for Massachusetts, Rhode Island, New York, and New Jersey from the value of $\hat{\delta}$ for Connecticut. The residual variance, assumed to be equal for all five states, is estimated from the average of the residual variances for all states. The values of $\hat{\psi}$, $\hat{\sigma}_{\hat{\psi}}$ and t which correspond to the maximum likelihood estimates of and reasonable upper and lower limits to γ are reported in Table 2. The bottom row of Table 2 is the probability of a Student t-variable falling below the value of $\hat{\psi}/\hat{\sigma}_{\hat{\psi}}$. The probabilities in the last row of Table 2 can be interpreted as the probabilities of obtaining a value of $\hat{\psi}$ as small as or smaller than the one obtained when ψ, the true difference between δ_C and $(\delta_M + \delta_{RI} + \delta_{NY} + \delta_{NJ})/4$, is zero. When the maximum likelihood estimates are used for each state, it is estimated that the shift of the Connecticut time-series is .408 units (fatalities/100,000,000 driver miles) greater in a *negative* direction than the estimated average shift for the four control states. A differential this large or larger would occur only 65 times in 1,000 ($p = .065$) if ψ were truly zero. Hence, there appears to be statistically reliable evidence of an abrupt diminution of the traffic fatality rate in Connecticut in January 1956.

CONCLUSION

It can be seen in Table 2 that one may conclude that there is a statistically significant reduction as of January 1956 in fatalities/100,000,000 driver miles for Connecticut as compared with the four control states.

The above conclusion must not be accepted without due consideration of a source of potential invalidity in the experiment. As Campbell and Ross have quite correctly pointed out, the fact that Governor Ribicoff was prompted to take action in late 1955 by the alarmingly high fatality rate for that period *introduces the possibility of a regression effect from the observations immediately preceding his actions to the observations immediately following.* If one observes a time-series for a period of time and selects that observation which appears quite extreme, subsequent observations are likely to be relatively less extreme. None of the analyses performed here "corrected" in any way for the possibility of this regression effect. It is not clear to the author how one might intentionally do so. The inferential techniques applied in this paper are designed to assess the statistical stability of an alteration in the course of a time-series; they are indeed quite blind to the underlying causes of such alterations. Subjecting the data to inferential statistical analyses is a properly cautious procedure when dealing with the unstable, fallible (showing random variation) data which the "real world" yields. But even after the stability or reliability of a statistical result is established, the cautions which Campbell and Ross have raised about the interpretation of such a result (*e.g.*, is it merely an instance of the regression effect) are themselves fully proper.