OPERATIONS RESEARCH CENTER

working paper

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

THE EFFECT OF THE 18-YEAR OLD DRINKING AGE ON AUTO ACCIDENTS*

by

Stephen Cucchiaro
Joseph Ferreira, Jr.
Alan Sicherman

OR 034-74 May, 1974

Massachusetts Institute of Technology
Operations Research Center
and
Laboratory of Architecture and Planning
Cambridge, Mass.

* This research was supported in part by NSF Grant GI 38004, by the MIT Undergraduate Research Opportunities Program, and by the U. S. Army Research Office (Durham) under Contract No. DA HC04-73-C-0032.

The Effect of the 18-Year Old
Drinking Age on Auto Accidents

by

S. Cucchiaro, J. Ferreira, and A. Sicherman
M.I.T., Cambridge, Mass.

Abstract

The effect of Massachusetts' reduced drinking age on auto accidents is examined by employing an interrupted time series analysis of monthly accident data covering the period January, 1969, through September 1973. The data were stratified by driver age, accident type and (to a limited extent) operating-after-drinking. These raw data were adjusted using monthly mileage and seasonal indices and, where possible, a control group not affected by the drinking law. Correlograms of the adjusted series were computed to check for remaining systematic bias. Finally, the average accident rates for the adjusted, well-behaved series before and after the March 1973 change were compared using standard t-tests.

Accident rates among 18-20 year olds did increase significantly--about 40% for involvement in fatalities. Nevertheless, the results are consistent with the hypothesis that, as a result of the reduced drinking age, 18-20 year old driving-after-drinking behavior has become comparable to that of older drivers.

The Effect of the 18-Year Old Drinking Age on Auto Accidents

Table of Contents

List of Tables

List of Figures

ACKNOWLEDGEMENT

Support for this research has come from several sources. The work began during the summer of 1973 as an undergraduate research opportunities project (UROP) for Stephen Cucchiaro (then a senior in mathematics) under the supervision of Prof. Joseph Ferreira of the M.I.T. Department of Urban Studies and Planning. The initial work was supported by the UROP office and NSF Grant G138004 for "Innovative Resource Planning in Urban Public Safety Systems" at the M.I.T. Operations Research Center and the Laboratory of Architecture and Planning. Alan Sicherman, a graduate student in Operations Research, joined the group in the fall of 1973 and was supported by Army Research Office under Contract No. DA HC04-73-C-0032.

We would also like to thank a number of individuals at the Massachusetts Registry of Motor Vehicles for interesting us in the question and providing us with the raw data and related statistical information. In particular, we appreciate the cooperation of James Velis, Assistant to the Registrar, Fred Cody, Administrative Assistant to the Registrar, and Evelyn Trefrey of the Registry's Statistical Bureau.

THE EFFECT OF THE 18-YEAR OLD DRINKING AGE ON AUTO ACCIDENTS

1. BACKGROUND

On March 1, 1973, Massachusetts put into effect a law which lowers the drinking age limit from 21 to 18 years of age. There has been considerable concern and discussion in the state legislature, in the Registry of Motor Vehicles, and in the press over the extent to which lowering the drinking age has caused an increase in motor vehicle accidents. Preliminary comparison of raw data for the few months before and after the March 1 change have produced much-quoted statistics such as a "131% rise in road deaths linked to teenagers and liquor."*

Though the increase is dramatic, these data are not adjusted for road use condition or for other factors that might have changed during the months involved. Also, the actual number of drinking related fatalities involving young operators are subject to large statistical fluctuations since they are small compared with the total number of highway fatalities and the total number of teenager-involved accidents during those months.** A more thorough analysis of the data is needed to determine the extent to which the continuation past trends, fluctuations due to chance factors and other such considerations account for portions of the changes observed in the raw data.

This paper studies in more detail several kinds of accident data and attempts to develop a broader and more reliable picture of how and to what extent reducing the drinking age has affected teenage driving and highway accidents. The analysis uses monthly property, injury and fatal

---

* See, for example, Boston Herald American, 12/26/73, p. 3.

** A Boston Herald American article on February 1, 1974, indicated that persons fatally injured as a result of accidents involving drinking drivers between 18 and 20 years of age totaled 79 between March 1, 1973 and January 24, 1974, and 35 during the corresponding period of 1972-73.

accident data for Massachusetts for the period January, 1969, through September, 1973. Several standard statistical techniques for examining times series data are used to test various hypotheses.

Lack of access to finer breakdowns of the data and to data beyond September 1973 limited the scope of the study. However, the advent of the energy crisis in the fall of 1973 would reduce the usefulness of the data beyond October, 1973. The data were sufficient to provide estimates of the law's effect on several accident rates. Fatal accident rates showed the largest increase -- up approximately 40% for the 18-20-year-old group. The findings support the general conclusion that, as a result of the law, the 18-20-year-old driving-after-drinking frequency is now comparable to that of older drivers.

2. HYPOTHESES AND AVAILABLE DATA

A variey of hypotheses concerning the effects of the new Massachusetts drinking law are possible. One is that there has indeed been a sharp increase in the accident rate for 18 to 20 year olds as a result of the lower drinking age limit.. Another is that motorists under 18 have experienced an increase in accidents since they can now obtain liquor more easily by themselves through their 18 to 20 year old friends. Motorists between 21 and 23 might also experience an increase since younger 18 to 20 year old friends with whom they associate will want to spend more time drinking.

A more complicated type of hypothesis might be that there was a transient increase in the accident rate immediately after the March 1 change but, after a few months, the rate dropped to a lower, long-term level. Yet another theory is that an increased accident rate after March 1, 1973 reflects a continuation of a trend that had been in progress prior to the inactment of the new drinking law. Still another suggests that the number of citations issued to 18 to 20 year olds for alcohol-related violations has increased due to a more stringent police attitude in response to the enactment of the relaxed Massachusetts drinking law.

Further complications arise if differential "types" of 18 to 20 year olds (e.g., females, just received licenses, etc.) have experienced different changes in accident rates. Also, different geographic areas (e.g., urban, near state border, etc.) may have experienced different effects. The rates for different types of accidents (e.g., fatal, property damage only, single car accidents, etc.) might change in different ways.

In order to test these hypotheses, much more data is needed besides before and after figures for accidents involving drinking operators aged 18 to 20. Certain data are needed to control for mileage driven, seasonality, past trends and the like. Then quarterly or monthly accident data broken down by age of operator, related citations, and accident severity are needed to test the basic hypotheses.

The Massachusetts Registry of Motor Vehicles provided us with certain monthly data for the January 1969 through September 1973 period.* The following types of information were available on a monthly basis:

(1) The number of citations** issued for "operating under the influence," and "operating after drinking" violations.

(2) The number of reported accidents classified as fatal, non-fatal, or property damage types.

(3) Driver age breakdowns for the above data.

(4) Breakdowns of (2) indicating whether "operating under the influence" or "operating after drinking" citations were issued.

(5) Estimates of the number of miles driven in Massachusetts.

Data beyond September were not available and, unfortunately, we could not obtain the accident data broken down both by age <u>and</u> drinking citations.

The fatal, non-fatal, and property damage only breakdowns enable differential changes in accident rates related to accident severity to be identified. The age breakdowns permit focusing on specific age groups. They are also helpful in controlling for other factors that might influence the number of alcohol related accidents since changes in the monthly number of accidents involving operators over 25 are not likely to be related to changes in the drinking law.

---

* These data were provided by the Accident Records Section of the Mass. Registry of Motor Vehicles.

** Operating after drinking and operating under the influence "violations" appearing in these Registry statistics refer to information on reports for investigated accidents. The investigator cites an involved operator for "driving under the influence" if he determines that the operator's blood alcohol content was at least 0.1%. Driving after drinking is indicated if the operator admits drinking or has a blood alcohol content less than 0.1%.

The drinking citation data provide the only available Registry indicators of the number of accidents related to liquor consumption. However, such citation data may be misleading to the extent that the March 1 change altered the circumstances under which police issued drinking citations to youths.

Eighteen time series could be developed using the four age groups or the two citation categories together with the three accident types. Table 1 identifies these 18 time series plus an additional series (A4-T1) for all monthly fatalities involving operators of all ages. The numbers and abbreviations for these time series given in Table 1 will be used throughout the report. Figures 1, 2 and 3 graph the unadjusted monthly data for several of these series.

Table 1. The Time Series Used in the Analysis

Driver and Accident Categories

OPERATOR INVOLVED**

*D1: Received an "Operating after Drinking" Citation

*D2: Received an "Operating under the Influence" Citation

A1: Under 18 years of age

A2: 18 to 20 years of age

A3: 21 to 23 years of age

A4: over 23 years of age

AA: All ages combined

ACCIDENT TYPE

T1: Involved a Fatality

T2: Involved only non-fatal personal injury

T3: Involved property damage only

Time Series Labels

1) D1-T1	7) A1-T1	13) A3-T2
2) D2-T1	8) A2-T1	14) A4-T2
3) D1-T2	9) A3-T1	15) A1-T3
4) D2-T2	10) A4-T1	16) A2-T3
5) D1-T3	11) A1-T2	17) A3-T3
6) D2-T3	12) A2-T2	18) A4-T3

19) AA-T1: Total number of deaths resulting from auto accidents.

---

* Includes drivers of all ages

** Multiple car accidents will appear in more than one category

Figure 1: RAW DATA FOR SELECTED TIME•SERIES (Part 1)

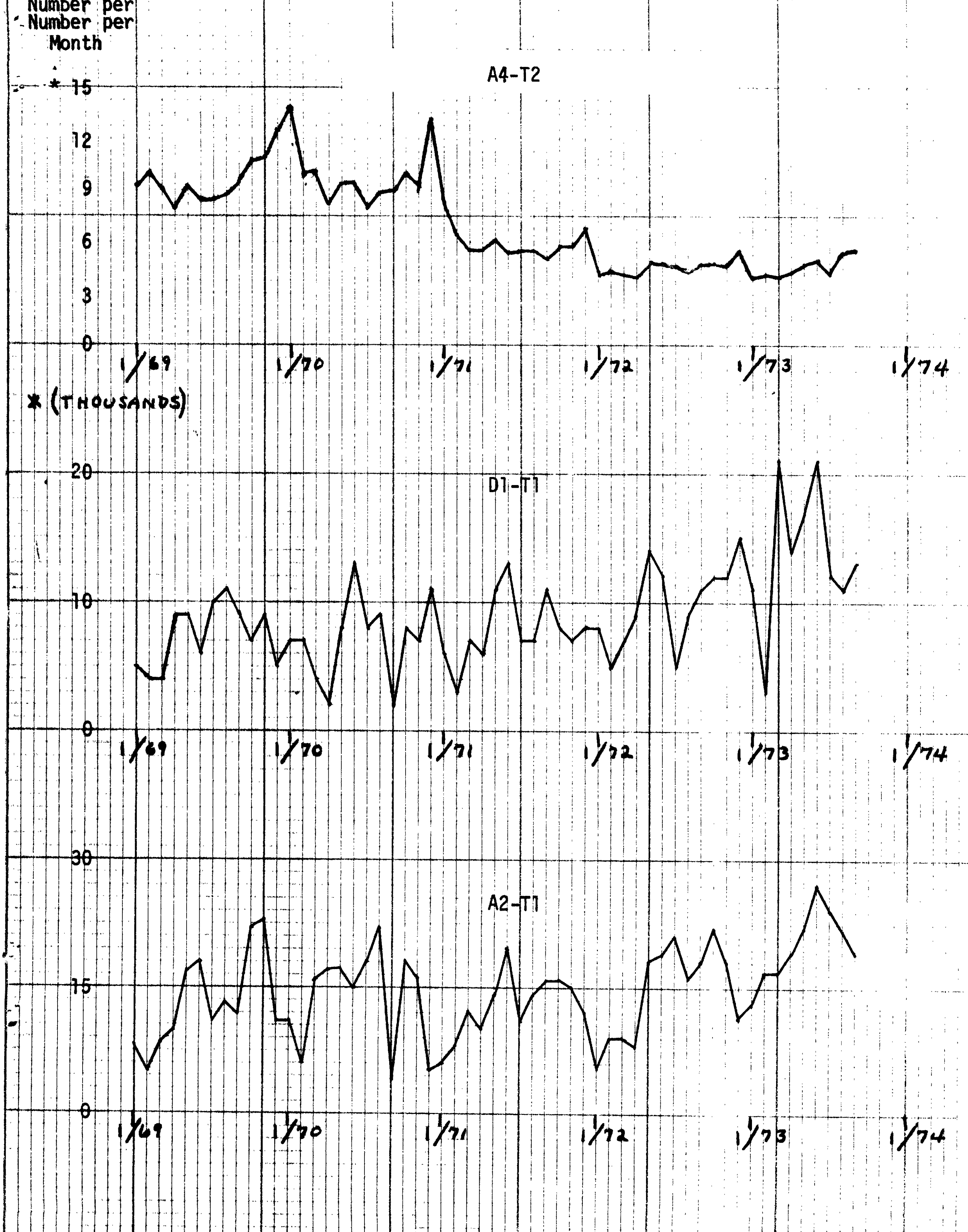
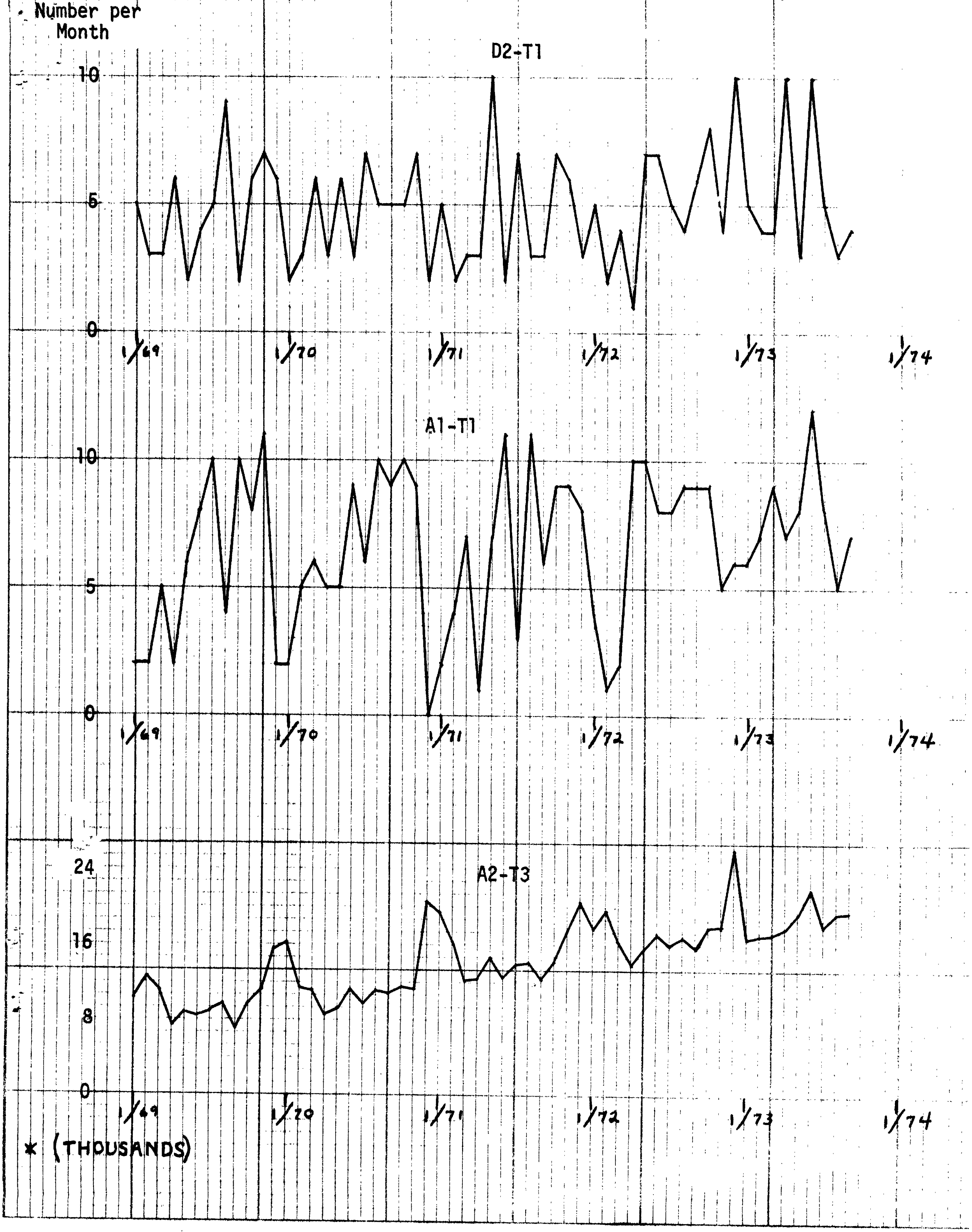


Figure 2: RAW DATA FOR SELECTED TIME SERIES (Part 2)



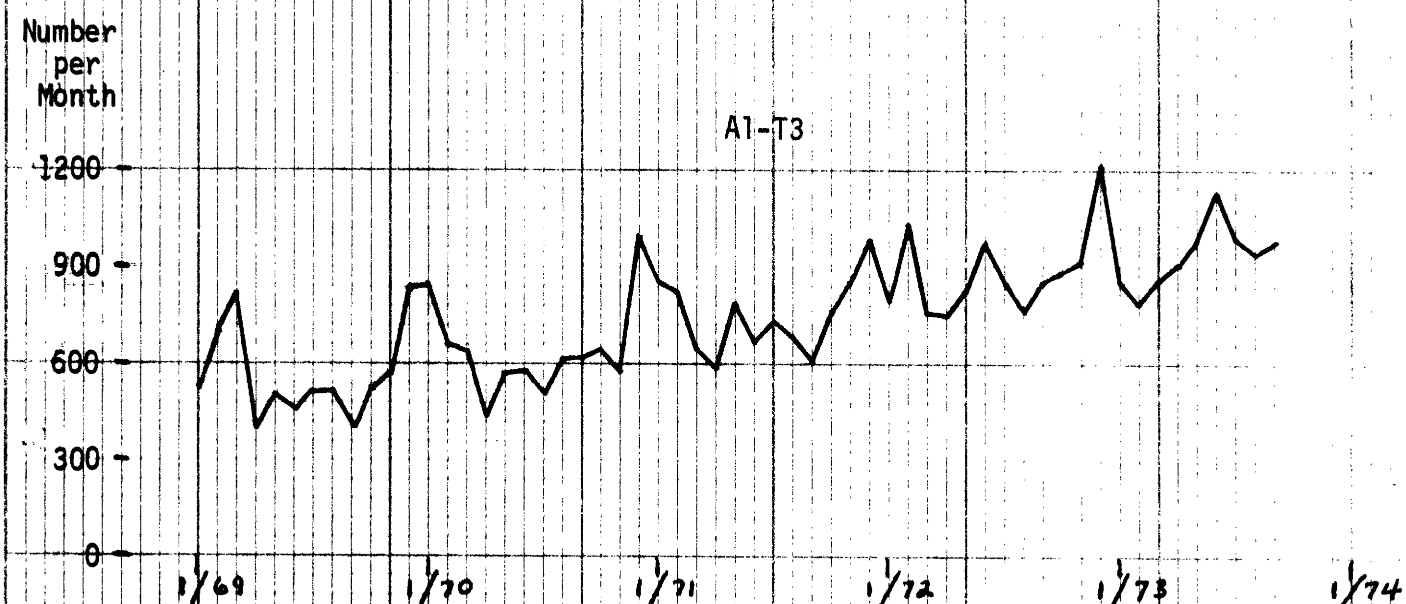


Figure 3: Raw Data for Selected Time Series (Part 3)

3. THE ANALYSIS

The behavior of the eighteen sets of monthly accident and citation data will be interpreted as interrupted times series.* After adjusting the series to remove suspected, systematic biases, we will identify those series which appear to have significantly different average accident rates before and after March, 1973. Finally, various hypotheses about the law's effects will be checked for consistency with the data and a theory of what has happened will be developed.

Several standard statistical techniques** are used to adjust the data for miscellaneous factors to correct for trends, and to identify those adjusted time series for which the policy change appears to be associated with significantly different behavior. Two methods were used to attempt to factor out the influence on the accident rate of changes in miscellaneous factors unrelated to the new drinking law such as the amount of driving, and seasonal trends. Both methods are described in detail in Appendix II.

One method scaled the monthly data using estimates of monthly mileage, seasonal trends and other such factors. The second method regarded the over 23 year old drivers as a control group whose accident data could be used to standardize the 16 to 23 year old experience. Assuming that miscellaneous factors affected all age groups similarly, any significant difference in the ratio of the monthly accident experience for the over 23 group and each of the other groups could be attributed to the changed drinking law.

---

* Appendix I briefly reviews the use of an interrupted time series design.

** The techniques are similar to those used elsewhere in studying highway accident data. See for example the work of Glass and Campbell on the effect of Connecticut's speed limit enforcement [1,4].

All eighteen series were adjusted using method one. Method two was used only for series #7, #8 and #9. (See discussion in Appendix II.) Figures 4, 5 and 6 compare three series before and after adjustments. The adjusted figures are scaled to represent the equivalent accident rate for each standardized month per 2.3 billion motor vehicle miles (the 1969-1973 average mileage per month). Thus, the 9 reported fatal accidents involving "after drinking" citations during April of 1969 correspond to a seasonally adjusted rate of 10.7 per 2.3 billion motor vehicle miles. Similarly, the 9 such accidents during August of 1970 correspond to 8 after the method I adjustments.

The adjustments account for some but certainly not all of the monthly fluctuations in the accident rates.* Nevertheless, the remaining "noise" is easily handled if it represents random fluctuations around average before and after accident rates. Ideally, the mileage and seasonality adjustments or the use of the over 23 control group account for all systematic biases in the series except for the March 1973 change. For series such as D1-T1, the only sharp or systematic change in the adjusted series appears to have occured in March 1973. For others, such as A3-T1, no trends are obvious and it is not clear whether there is a March 1973 interruption. For the A4-T2 series, interruptions at other times appear more important.

To help identify those "well behaved," adjusted series which appear to exhibit trends or interruptions in March 1973 (or not all all), we examined the correlograms of the adjusted time series. As explained in Appendix II, the correlogram of a particular series measures the extent to which accident rates K months apart are correlated. Correlograms of a few series (for the period before the March 1973 change) are shown in the Appendix for K values

---

* For example, the mean number of monthly citations for the D1-T1 series was 8.02. The standard deviations before and after mileage and seasonality corrections were 3.08 and 2.53.

Figure 4: SERIES D1-T1 BEFORE AND AFTER ADJUSTMENTS

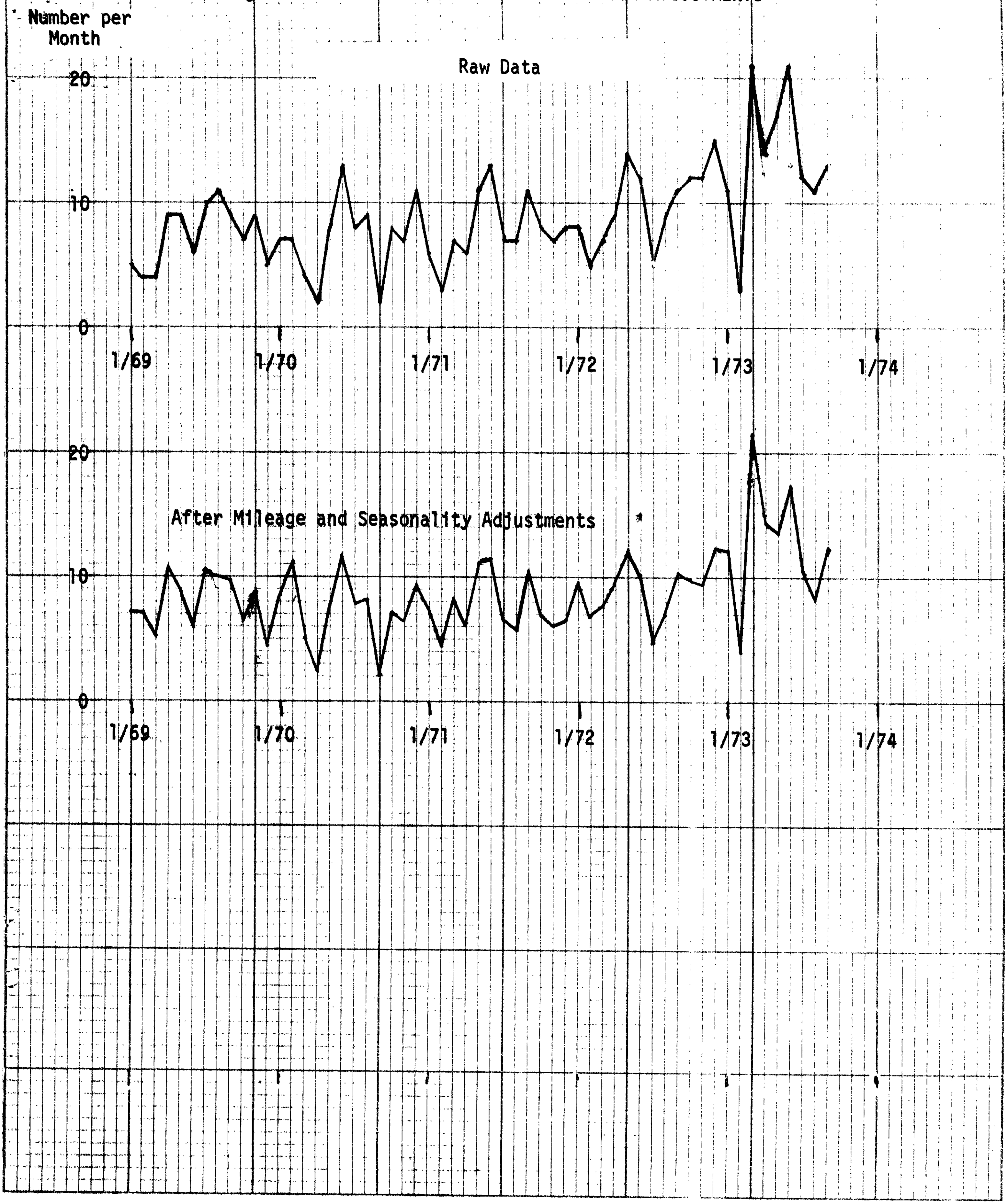


Figure 5: SERIES A2-T1 BEFORE AND AFTER ADJUSTMENTS

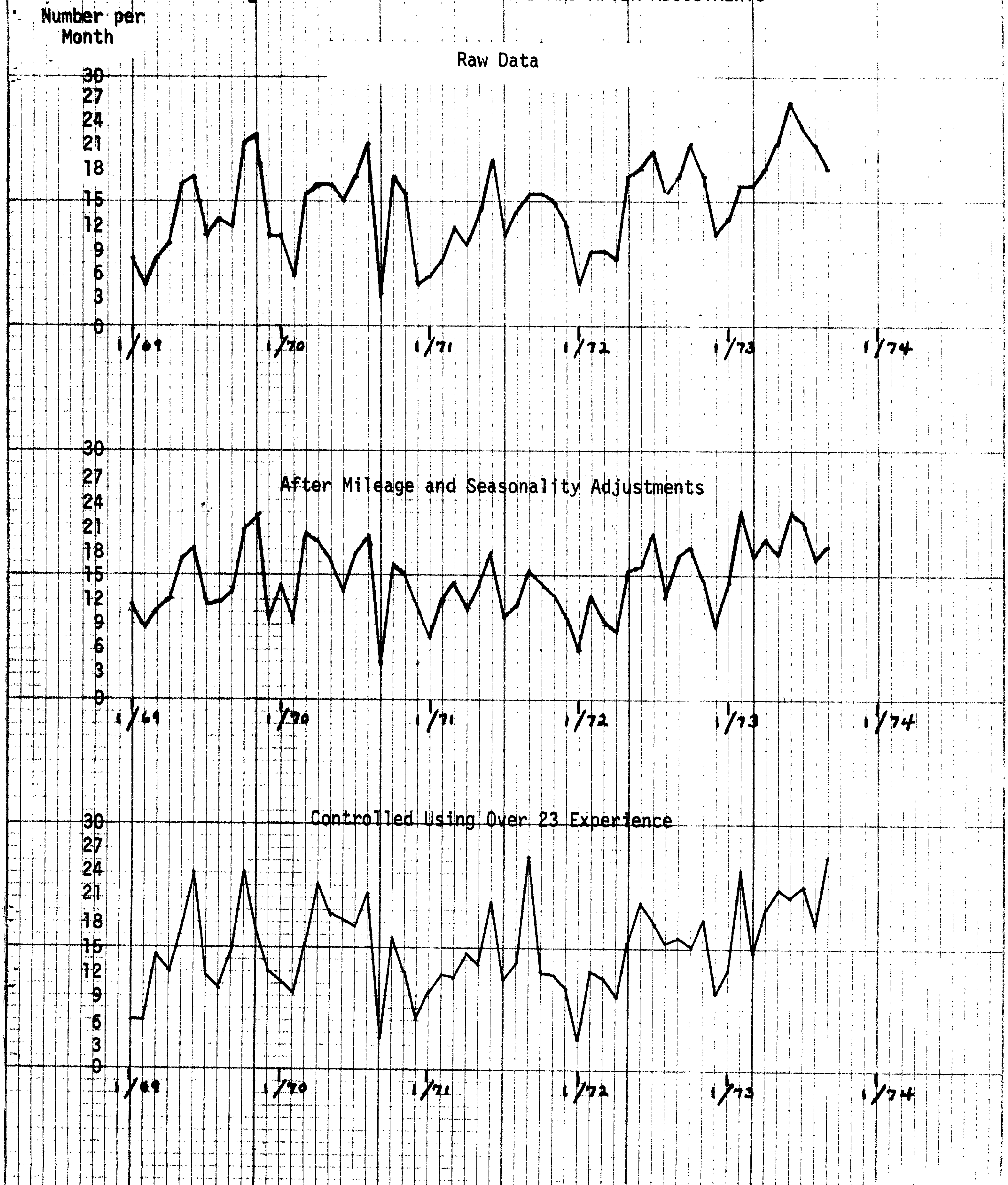
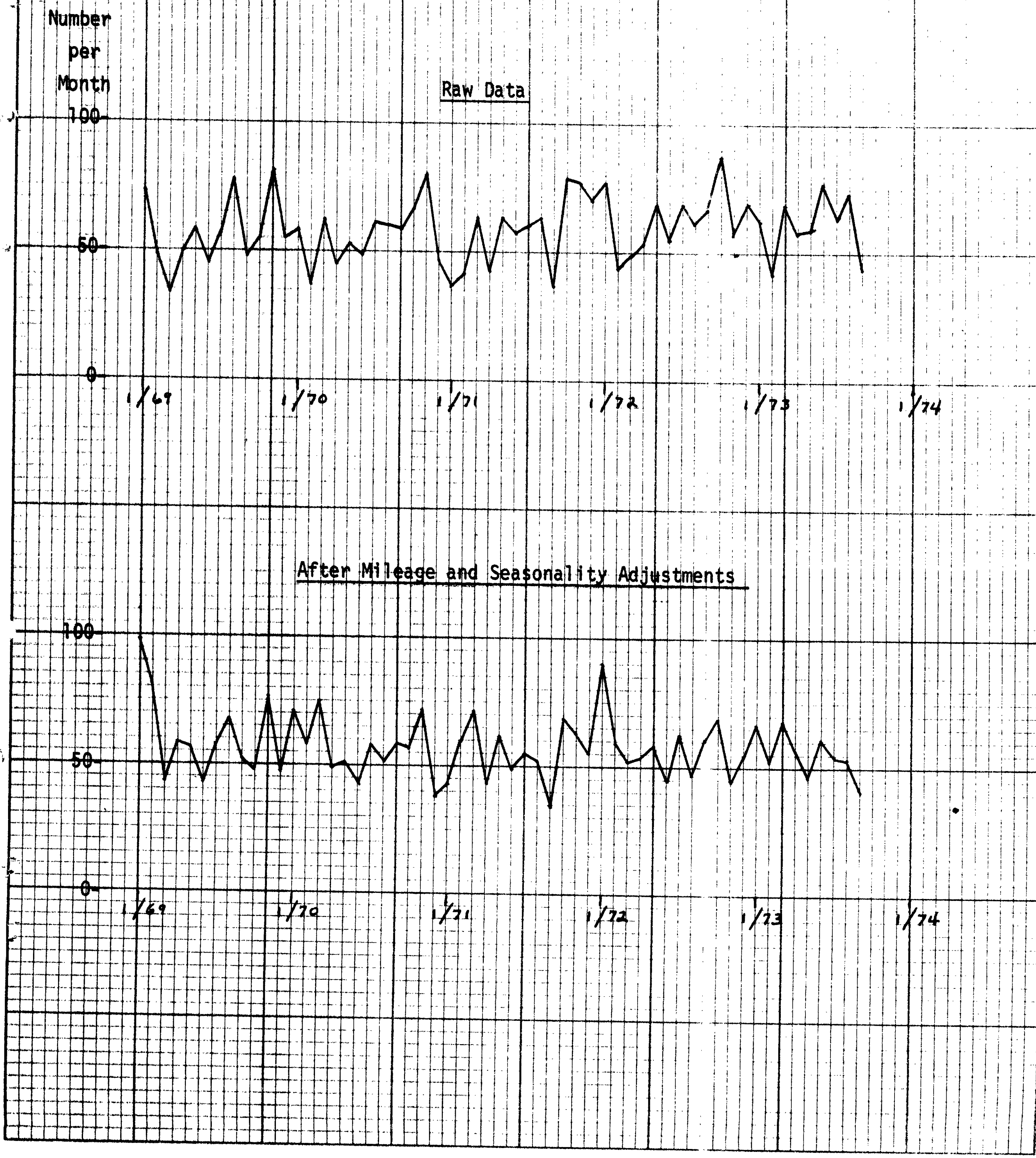


Figure 6: SERIES A4-T1 BEFORE AND AFTER ADJUSTMENTS



between 1 and 50. A series exhibiting linear growth with no noise would produce a linearly increasing correlogram. A stable series with uncorrelated, random noise would produce a stable correlogram that fluctuated randomly around 0.

As expected, series A4-T2 in Figure 6 above produced a poorly behaved correlogram. Both series D1-T1 (#1) and A2-T1 (#8) were judged well-behaved. In all, nine series were acceptable (see Appendix II for a discussion of the criteria used to make these judgements). Of those adjusted using method I, #1, #3, #5, #8, #9, #10 and #19 were accepted. Of those adjusted using the over 23 experience, #8 and #9 were accepted.

In summary, these nine series--after adjustments and controls to remove certain systematic biases--appear to fluctuate randomly about a stable level for the months prior to March 1973. To the extent that this is the case, the hypotheses suggested in section 2 can be tested by comparing the average accident rates before March 1973 with the experience after that date.*

Table 2 and figures 7, 8 and 9 present the results. For each of the eight series, averages for the adjusted numbers of accidents per month were computed for four time periods:

Period A:	January	1969	through	February	1973
Period B:	August	1972	through	February	1973
Period C:	March	1973	through	September	1973
Period D:	April	1973	through	September	1973

---

* Conceivably more complicated adjustments and statistical techniques could be used to interpret some of the remaining 10 series. However, we felt that the short time period following the change, the uncertainty about alternative adjustments and the level of unexplained noise in the data limited the practicality of more detailed analysis.

Considering Periods B and D separately permits possible time varying effects to be identified. If the average accident rate for Period B were significantly different from the average during Period A, then a trend which began before passage of the law might account for any Period A to C change. Periods C and D are distinguished to enable identification of any transient effect of the law during March 1973. (With only seven months of data following the change it was not possible to investigate longer transient effects.)

Since all nine adjusted series vary substantially from month to month throughout the 5 years, small changes in the before and after average accident rates do not necessarily suggest a March 1973 interruption in the series. To establish a criterion for determining the smallest change that should be considered significant, we employed standard t-tests for the difference between two means. The usual assumptions about the normality, independence and stationarity of the "noise" in the series were made and 95% confidence intervals were used for each of the eight series. Appendix II explains the assumptions and the technique.

4. THE RESULTS

The results for all nine time series are summarized in Table 2 and graphed in Figures 7, 8, and 9. The letter beside each bar indicates the time period for which the average accident rate was calculated.* The dotted vertical lines are related to the tests for significant differences before and after March 1973. For each series, the dotted line indicates the lowest possible period C averages that would still be judged significantly higher than the period A average accident rate. Similar lines might be drawn to compare means for other combinations (A-D, B-C, B-D, and A-B). However, B-C and B-D comparisions are based on too little data (6 or 7 months each) to have narrow confidence intervals and are not used. Corresponding lines for A-D and A-B comparisions are not sufficiently different to warrant separate lines on the graphs (Appendix II discusses the differences in more detail).

The bar graphs in Figures 7, 8 and 9 indicate the average number of operators (or deaths in one instance) involved in particular types of accidents during each month. The magnitude of the bars are scaled so that they correspond to the average number of operators (or deaths) for a deseasonalized month per 2.3 billion motor vehicle miles (the January 1969 through February 1973 average). For example, the Figure 7 result for fatal accidents involving operators over 23 years old has a magnitude of 59.8 during period A. Thus an "average" month between January 1969 and February 1973 had 59.8 such operators involved in fatal accidents. Similarly, an average of 13.7 operators aged 18-20 were involved in fatal accidents during each month of period A.

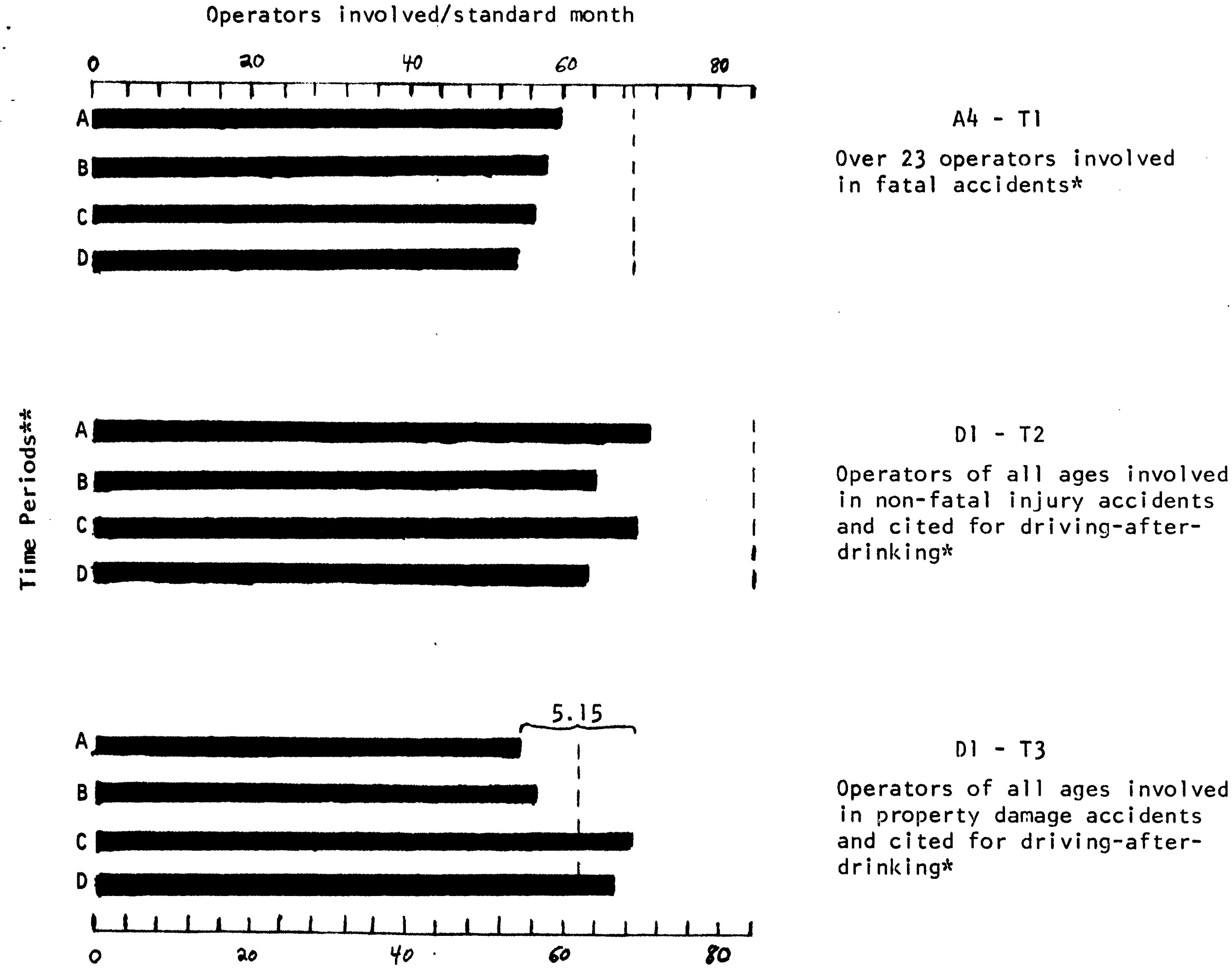
---

* Period A includes January 1969 through February 1973
Period B includes August 1969 through February 1973
Period C includes March 1973 through September 1973
Period D includes April 1973 through September 1973

Table 2: The Results -- Before and After Averages

TIME SERIES NUMBER	TIME SERIES DESCRIPTION	Monthly Averages of Adjusted Data during TIME PERIODS: A	B	C	D	Smallest possible significant increase (% of A)	Significant Increases $(100 \frac{C - A}{A})$
#1 (D1-T1)	FATAL - OPERATING AFTER DRINKING	8.03	9.34	14.03	12.79	1.77 (22%)	75%
#3 (D1-T2)	NON-FATAL - OPERATING AFTER DRINKING	70.73	64.19	68.66	63.18	13.23 (18.7%)	
#5 (D1-T3)	PROPERTY DAMAGE - OPERATING AFTER DRINKING	54.23	56.10	68.49	66.08	7.82 (14.4%)	24.5%
#8 (A2-T1)	FATAL 18-20	13.69	15.46	18.98	19.25	3.07 (22.0%)	38.5%
#8 (A2-T1)	FATAL (using control group) 18-20	13.69	15.47	19.78	20.76	3.51 (25.6%)	44.7%
#9 (A3-T1)	FATAL 21-23	12.47	13.02	12.58	12.28	2.85 (22.9%)	
#9 (A3-T1)	FATAL (using control group) 21-23	12.48	13.11	13.11	13.40	3.51 (28.1%)	
#10 (A4-T1)	FATAL OVER 23	59.85	58.28	56.30	53.89	9.15 (15.3%)	
#19 (AA-T1)	ALL DEATHS	78.0	79.5	80.0	77.5	8.08 (10.4%)	

Figure 7. Average Monthly Rates Before and After Change (Part I)

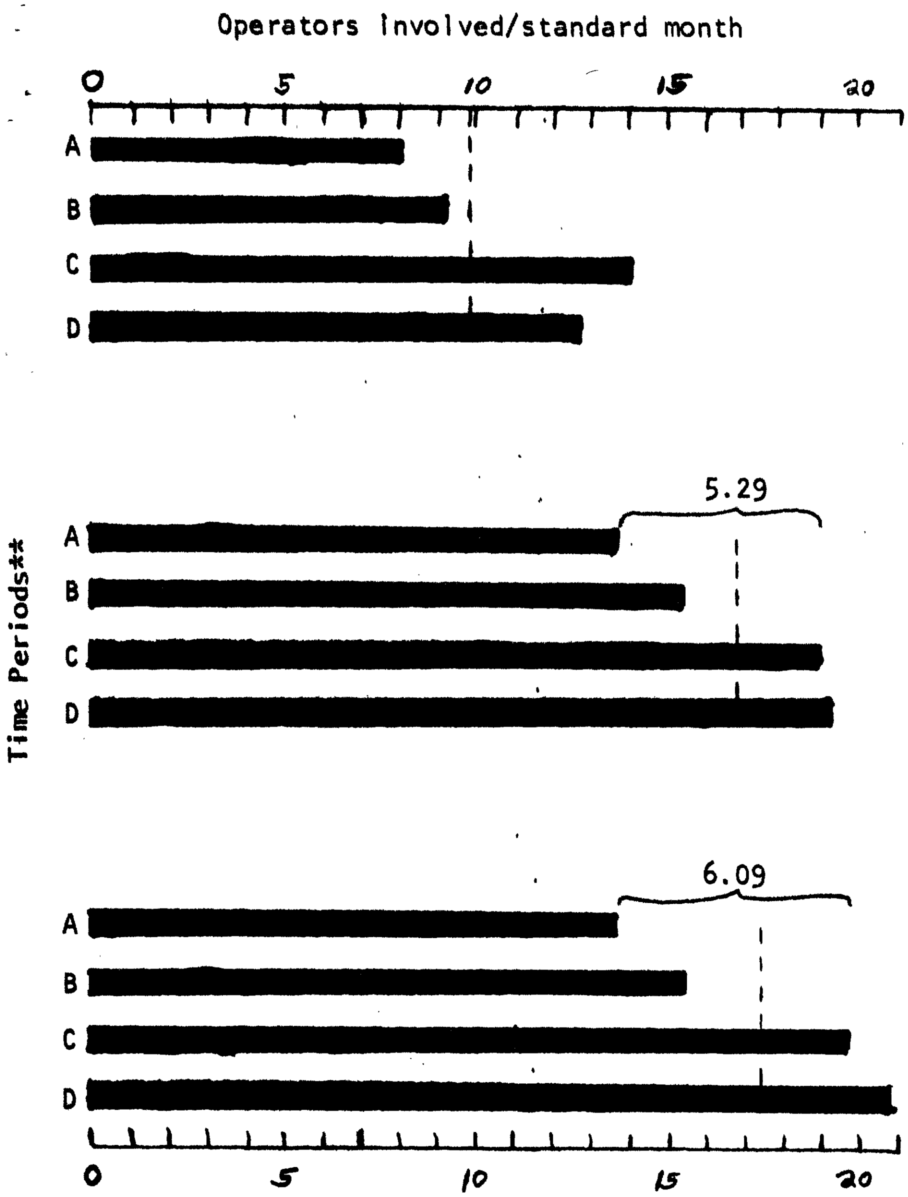


*The standard monthly averages are computed from the monthly figures adjusted for mileage and seasonality and are scaled to indicate the rate per 2.3 billion miles (the monthly average).

**Period A: January 1969 through February 1973
Period B: August 1969 through February 1973
Period C: March 1973 through September 1973
Period D: April 1973 through September 1973

---Period C averages extending to the right of the dotted line are considered significantly different from the period A average. (Within the accuracy of the graphs, the same is approximately true for A-B and A-D comparisons.)

Figure 8. Average Monthly Rates Before and After Change (Part 2)



D1 - T1

Operators of all ages involved in fatal accidents and cited for drinking-after-driving*

A2 - T1

18-20 year olds involved in fatal accidents*

A2 - T1

18-20 year olds involved in fatal accidents*** (using control group)

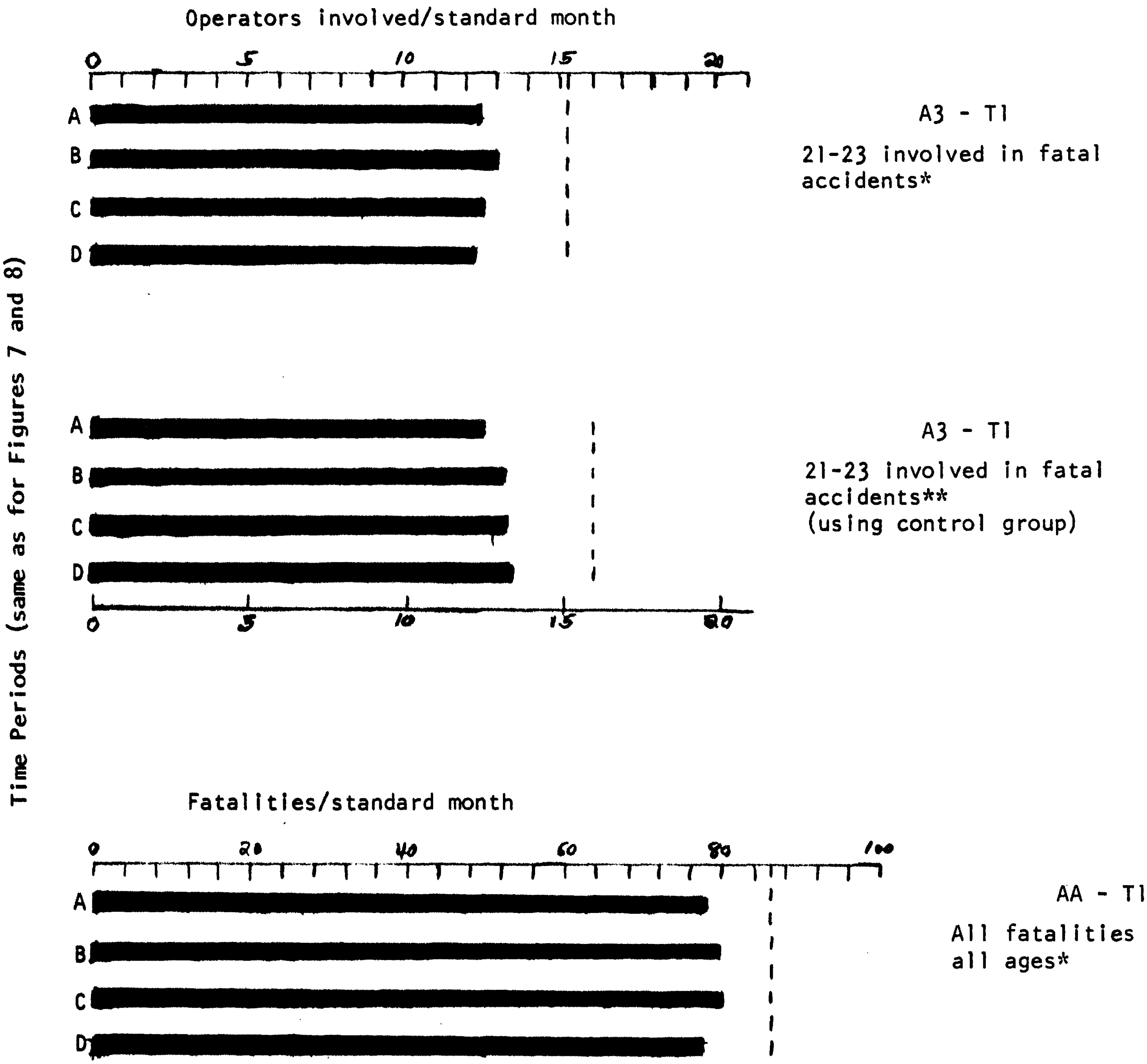
*The standard monthly averages are computed from the monthly figures adjusted for mileage and seasonality and are scaled to indicate the rate per 2.3 billion miles (the monthly average).

**Period A: January 1969 through February 1973
Period B: August 1969 through February 1973
Period C: March 1973 through September 1973
Period D: April 1973 through September 1973

***The standard monthly average is computed from the monthly ratios of the 18-20 year old experience to the over 23 year old experience multiplied by (1/5.75) so that the period A average matches that obtained in series A2-T1.

---Period C averages extending to the right of the dotted line are considered significantly different from the period A average. (Within the accuracy of the graphs, the same is approximately true for A-B and A-D comparisons.)

Figure 9. Average Monthly Rates Before and After Change (Part 3)



*The standard monthly averages are computed from the monthly figures adjusted for mileage and seasonality and are scaled to indicate the rate per 2.3 billion miles (the monthly average).

**The standard monthly average is computed from the monthly ratios of the 18-20 year old experience to the over 23 year old experience multiplied by (1/5.75) so that the period A average matches that obtained in series A3-T1.

---Period C averages extending to the right of the dotted line are considered significantly different from the period A average. (Within the accuracy of the graphs, the same is approximately true for A-B and A-D comparisons.)

For the bar graphs corresponding to series adjusted using the control group, the interpretation is slightly different. The number of over 23 year old operators involved in fatal accidents in each month is used as the standard of comparision. The bars represent the average over each period of the ratio of the 18-20 year-old experience to the over 23 year-old experience. The length of the bars are scaled so that the value during period A matches the period A mean obtained using method I adjustments (series #8 or #9).

The sizable monthly fluctuations that remain in the adjusted time series limit our ability to determine whether changes in the A to C or A to D averages are due to the March 1973 change or to random fluctuations in monthly figures. Thus, for the 18-20 fatalities of series #8, changes of less than 3 involvements in fatalities per month between periods A and C would not be judged significant (using the t-test described in Appendix II at the 95% confidence level) even though a change of 3 would amount to 22% of the period A monthly average. Similarly, for the over-23 fatality figures, changes of less than 9 operator involvements per month (in the period A to period C averages) would not be judged significant. Despite these large noise levels a number of significant differences were found.

Depending upon which averages are compared, the estimated magnitude of the changes can differ. For series #1 the estimates vary between 37% (comparing period B to period D) and 75% (comparing period A to period C). We shall use changes in the period A to period C averages as our (point) estimate of the changes in those series judged to be "interrupted." These changes are thought to be the best (point) estimates of the effects of law. The period B averages were based on the 7 months just preceeding the enactment of the law and are subject to large standard errors. However, the period A averages are based on several years data after corrections for seasonality and trends and are not substantially different from the period B averages. Period C figures are used for post March 1973 estimates since so few months data were available after the change and period C and D estimates were not significantly different. Series #1 estimates are most sensitive to the choice--a fact that may be due to changed enforcement levels for issuing driving after drinking violations. Such possibilities will be discussed later.

Results for each series are discussed separately in the remainder of this section. General conclusions and interpretations are developed in section 5.

Series #1: Operating after drinking fatalities

Period C and D averages were significantly different from those during period A. The estimated 75% increase from 8.03 to 14.03 cases per month corresponds to 6 more operators per (standardized) month involved in fatal accidents when cited for driving after drinking. In more graphic terms, the relationship between the number of operators involved in fatal accidents and the number of deaths during the years 1967-1972 has been close to 6/5.[15] For the series #1 estimates this amounts to 5 more deaths per month associated with "driving-after-drinking" accidents.

Series #3: Operating after drinking injuries

No significant increase was noted. Some studies have shown that alcohol is less of a factor in property damage and non-fatal injury accidents[11] and this may account for the different results for series #1 and #2. Other possibilities are changes in the enforcement of driving after drinking restrictions and the fact that period A to period C changes would have to exceed 13 (about 20%) before being judged significant.

Series #5: Operating after drinking involvements--property damage only

Here a significant increase of 24.5% is found. The change is much less than for fatal accidents involving drinking--consistent with the theory mentioned above. It corresponds to an increase of about 8 in the number of operators involved in property-damage-only accidents and cited for driving after drinking.

Series #8: Fatal accidents involving 18-20 year olds

A significant increase of about 38.8% is noted. This amounts to 5.3 more operator involvements per month and is very close to the difference observed in series #1. Such a match suggests that 18-20 year old operators account for the increase in drinking-related fatalities.

Series #9: Fatal accidents involving 21-23 year olds

No significant increase is observed suggesting that this age group was not affected by the law. The 1% increase was far from the 25% increase needed in order for the change to be judged significantly above the noise level.

Series #10: Fatal accidents involving operators over 23

As expected, no significant change is observed. The 6% drop is well within the normal range of fluctuations.

Series #20: 18-20 year old fatalities--using the control group

A significant increase is noted. The 44.8% change amounts to 6 additional operator involvements per month. These figures closely match the series #8 figures for 18-20 year old operators involved in fatalities when adjustments were based on seasonality and miles driven.

Series #21: 21-23 year old fatalities--using the control group

No significant increase is observed reconfirming the results for the 21-23 year olds using mileage and seasonality considerations. It should be noted that the figures using the control group are numerically close to their counterparts using mileage and seasonality. The concurrence strengthens the argument that changes are due to the law and not to extraneous factors not considered in the analysis.

Series #19: All auto accident-related fatalities

No significant change is observed. However, increases less than 10% per month could not be distinguishable from noise in the data. The impact on the total number of deaths of the increase in the number of 18-20 year olds involved in fatalities is not great enough to be distinguished from the impact that random or unknown factors have upon the total number of deaths. This is to be expected since the 18-20 year old group account for only about 15% of the total number of operators involved in fatalities. Even a fairly large change in this group may not substantially affect the total mean.

5. Summary and Conclusions

Of the eighteen sets of 57-month time series, only the seven series involving fatalities or driving-after-drinking violations were sufficiently well behaved to be used in the analysis. The non-fatal and property-damage data were distorted -- apparently due to changes in accident reporting under the no-fault insurance law. The driving-under-the-influence and under-18-year-old-fatality data were subject to large monthly fluctuations because of the small monthly totals involved and their correlograms were unacceptable. The correlograms for the driving under the influence citations were rejected because they exhibited marked 12 month periodicity even after the standard seasonality connections.*

The results for the seven usable series support the hypothesis that reducing the drinking age resulted in an increase in drinking related accidents among 18-20 year old drivers. Among these usable series, the two which did not include the experience of 18-20 year old operators did not appear to experience interruptions in March of 1973. Fatal accidents involving 21-23 year olds (series A3-T1) and operators over 23 (series A4-T1) were apparently unaffected by the law, but those for 18-20 year olds (series A2-T1) jumped 40%.

Two of the three series concerning accidents involving operating after drinking citations increased significantly -- 75% and 24% respectively for driving-after-drinking fatalities and property-damage-only accidents. These results indicate a substantial March 1973 change in the frequency of drinking related accidents (and, possibly, in the circumstances whereby

*Evidently seasonality adjustments for accident occurrence do not correspond to those for accidents involving driving-under-the influence citations. We did not have enough data to permit meaningful adjustments to correct for this periodicity.

drinking citations are issued). Since the increase in reported operating-after-drinking fatalities (75% corresponds to 6 more per month) matched the numerical increase in 18-20 year old involvement in fatal accidents (40% corresponds to 5.3 more per month), and since no older age groups experienced an increased rate of involvement in fatalities, one concludes that there was a significant increase in the frequency of drinking-related accidents among 18-20 year olds.

Judging from the results for the operating-after-drinking series, the increase in accident rates among 18-20 year olds more most pronounced for fatal accidents. Precise figures for non-fatal and property damage accidents could not be developed, but the effect for non-fatal accidents appears smaller by a factor of about three. It is quite possible that the changed law produced a transient effect as well as a (smaller) long term increase in accident rates. However, the data did not include enough months after the change to enable detailed investigation of these transient effects.

Despite the dramatic increase in fatalities among 18-20 year olds and in the number of drinking-related citations, some care must be exercised in using the results to draw conclusions about the drinking behavior of youths. From a public policy point of view, one would like to know how the accident and drinking behavior of 18-20 year olds before and after March 1973 compares with that of older motorists. Arguments for repealing the law would be strengthened if the post March 1973 accident rate for the youths reflected a much higher drinking, driving and accident involvement rate than existed among older age groups.

It is quite tempting to draw such an inference from the above results. The logic is as follows: (1) The increase of 5 per month in the number of fatalities involving 18-20 year olds was caused by the relaxed drinking law and reflects an increase in driving-after-drinking accidents. (2) Since this increase matches the increase in operating-after-drinking citations, the latter are a good indication of drinking-related accidents. (3) Even if we assume that no operating-after-drinking related fatalities prior to March 1973 involved 18-20 year olds,* the 5 or 6 new monthly 18-20 year old drinking-related fatalities amount to more than 35% of all post-March 1973 operating-after-drinking citations. (4) Since 18-20 year olds comprise only 8% of all drivers, the group accounts for more than 4 times its share of drinking-related fatalities.

The problem with this reasoning is that all operators in all age groups are assumed to receive drinking-related citations whenever involved in fatal accidents after drinking. Before March 1973, only 15% of all operators involved in fatalities received drinking citations. But numerous studies have indicated that upwards of 50% of all auto fatalities are alcohol related.[11] Raw data from the Massachusetts State Police Laboratory suggest similar numbers (55%-60% for both over and under 21 year old age groups after March 1973).** Another indication of changed reporting is that, even though such a small percentage of fatalities produced drinking violations, the number of monthly operating-after-drinking violations after March 1973 increased by <u>more</u> than the increase in 18-20 year old fatalities -- and that doesn't count any additional operating-under-the-influence violations.

In light of this possibility that drinking related citations may be an incomplete and biased indication of alcohol-related fatalities among all

*Preliminary data from the Department of Public Safety (reported, for example, in the February 1, 1974, issue of the Boston Herald American) associate 17% (30 out of 174) of the drinking-related citations issued between March 1, 1972 and January 24, 1973 with 18-20 year old operators.

**More extensive discussion of the State Police Laboratory data may be found in Appendix III.

age groups, let us reexamine our results. The most reliable estimate developed in this paper is the 40% increase in 18-20 year fatalities. Fatality data are generally considered to be the most complete accident data and, unlike the drinking violation data, are not affected by citation practices. Since an increase in fatalities was observed after March 1973 only for the 18-20 year-old group, the additional 5.3 involvements per month is our best indication of the increase in drinking-related fatalities due to the law.

These additional 5.3/month amount to 28% (5.3/19) of the post March 1973 18-20 year old fatalities. But 18-20 year-olds had been regularly involved in alcohol related fatalities prior to the change in law. Suppose the 18-20 year olds now drove after drinking as often as other age groups and in accordance with the studies mentioned above were involved in drinking-related fatalities that accounted for about half the fatal accidents associated with their age group. Then about 10 (53%) of the roughly 19 monthly (post March 1973) fatal accidents involving 18-20 year olds would be alcohol related. Of these 10, about 5 would result from the change in the law implying 5/month prior to March 1973. Thus we expect that 5/13.7 or 36% of the 18-20 year old fatalities before March 1973 involved alcohol. Between 2 and 3 drinking-related citations were issued monthly to 18-20 year olds involved in fatalities before March 1973.* Allowing for a certain amount of under-reporting, the 5/month estimate for pre-1973 months is reasonable.

This picture of 18-20 year old behavior can be consistent with the drinking-related violations data as well. Suppose alcohol-related fatalities and citations for other age groups remained unchanged before and after

*From the Department of Public Safety figures cited earlier.

March 1973 and that 18-20 year olds involved in drinking-related fatalities were issued operating-after-drinking citations half the time before March 1973 and three-fourths of the time afterwards. Then they would account for $13.7 \cdot 0.36 \cdot 0.50 = 2.5$ citations per month before March 1973 and $19 \times 0.53 \times 0.75 = 7.6$ citations after.* The difference of 5.1 citations per month is based on several individual estimates, each of which is subject to substantial fluctuations. Nevertheless, it is close to the observed increase of 6 operating-after-drinking citations per month.

The above argument shows how the observed data could result if alcohol were related to the same fraction of fatal accidents for 18-20 year olds as for older drivers. However, 18-20 year olds typically have more than their share of fatal accidents. Though they include 8% of all drivers, they account for about 15% of all involvements in fatal accidents between 1969 and 1972.[15] Whether this is due to more driving or driving exposure, to less driving experience and skill, or to some other reason is not clear, but relatively frequent driving-after drinking is not likely to be the cause.

In conclusion, our interpretation of the consequences of the reduced drinking age supports a re-evaluation of the arguments for repealing the law. While the 18-20 year olds do have a higher fatal accident involvement rate than older drivers, the same fraction appears to be drinking-related -- implying that 18-20 year olds now drink and drive about as often as older drivers. Hence, arguing for a prohibition on 18-20 year old drinking solely in order to avoid the 5/month increase in fatal accidents involving 18-20 year olds appears unduly discriminatory against this age group -- why not prevent 30-40 year olds

*The second term in the calculations is the previously estimated fraction of fatal accidents that involve alcohol.

from drinking and protect their families as well as themselves?*

6. Future Research

Finally, some comments about unanswered questions and future research. Appendix III discusses some of these issues in more detail. No-fault insurance appeared to limit the usefulness of non-fatal injury and property damage data, thereby eliminating the possibility of estimating the law's effect on less serious accidents. The best one could do was get a crude indication from drinking violation data that the effect is much smaller than for fatalities. Transient effects may be involved (see in particular the DI-TI data in Figure 4 for operating-after-drinking fatalities) but cannot be adequately studied without more data following March 1973. Unfortunately, data after October 1973 must be adjusted to include "energy crisis" effects -- a particularly difficult correction if various age groups are affected differently.

The most logical next step would be to break down the drinking-related citations by age and repeat the above analysis. Accurate estimates of the fraction of 18-20 year old fatalities that involve liquor are central to policy recommendations regarding the drinking age law. However, the State Police Lab and drinking-related citation data are not unbiased, consistent measures of drinking-related driving. More detailed information is needed about the circumstances whereby drinking-related citations are issued and about the relationship between operator blood specimens tested and the number and type of fatal accidents involved.

*Since 18-20 year olds have more than their share of fatalities, increasing the driving age might be justifiable (unless the difference is due to experience), but such an option is not at issue here.

Bibliography

1. Campbell, Donald T., "Measuring the Effects of Social Innovations by Means of Time-Series," Statistics: A Guide to the Unknown, J. Tanur, et al. editors, Holden-Day, Inc., San Francisco, 1972, pp. 120-129.

2. Department of Scientific Research and Road Research Laboratory, Research on Road Safety, London: Her Majesty's Stationary Office, 1963.

3. Gibra, Issac N., Probability and Statistical Inference for Scientists and Engineers, Englewood Cliffs, N.J., Prentice Hall, Inc., 1973.

4. Glass, Gene V., Analysis of Data on the Connecticut Speeding Crackdown as a Time Series Quasi-Experiment, Law and Society Review 311, (1968-1969), pp. 55-76.

5. Glass, Gene V., A Program for the Analysis of Certain Time-Series Quasi-Experiments, Educational and Psychological Measurement 27 (1967), pp. 743-750.

6. Haight, Frank A., "Do Speed Limits Reduce Traffic Accidents?," Statistics: A Guide to the Unknown, Holden-Day, Inc., San Francisco, 1972, pp. 130-136.

7. Haddon, W., Accident Research--Methods and Approaches, New York: Harper and Row, publishers, 1964.

8. Holtzman, Wayne H., "Statistical Models for the Study of Change in a Single Case," in Problems in Measuring Change, C. W. Harris, editor. The University of Wisconsin Press, Milwaukee, 1963.

9. Kendall, M. G., Time Series, New York: Hafner Press, 1973.

10. Liquor Handbook, New York: Gavin-Jobson Associates, Inc., 1968.

11. Little, A. D., The State of the Art of Traffic Safety, New York: Praeger Publishers, Inc., 1970.

12. National Safety Council, Accident Facts, Chicago, Illinois (1967 and 1969 editions).

13. National Safety Council, Traffic Safety, Chicago, Illinois, March and September issues of 1969-1973.

14. Rebello, Robert E., editor, Econometric Software Package Users Manual (ESP), Cambridge, Mass., M.I.T., November, 1972.

15. Registry of Motor Vehicles, the Commonwealth of Massachusetts, Publication of Statistical Facts, Boston: Registry of Motor Vehicles Statisticians Office, 1969-1972.

16. Spiegel, Murray P., Theory and Problems of Statistics Outline Series, New York: McGraw-Hill Book Company, 1961.

Appendix I -- THE TIME SERIES

Part A: The Interrupted Time Series Design

The interrupted time series design is one useful method which allows us to draw inferences from statistical data. Graphically, the statistic being tested is plotted for a significant amount of time before and after a policy change is introduced. (See Figure A-1 below.) Assuming everything except the policy under study remains unchanged, we can then deduce whether a change in the statistic occurring after the policy change is associated with significantly different behavior (LINE A), is a continuation of a trend that had been in progress prior to the policy change (LINE B), or is just part of an unstable "zig-zag" (LINE C).

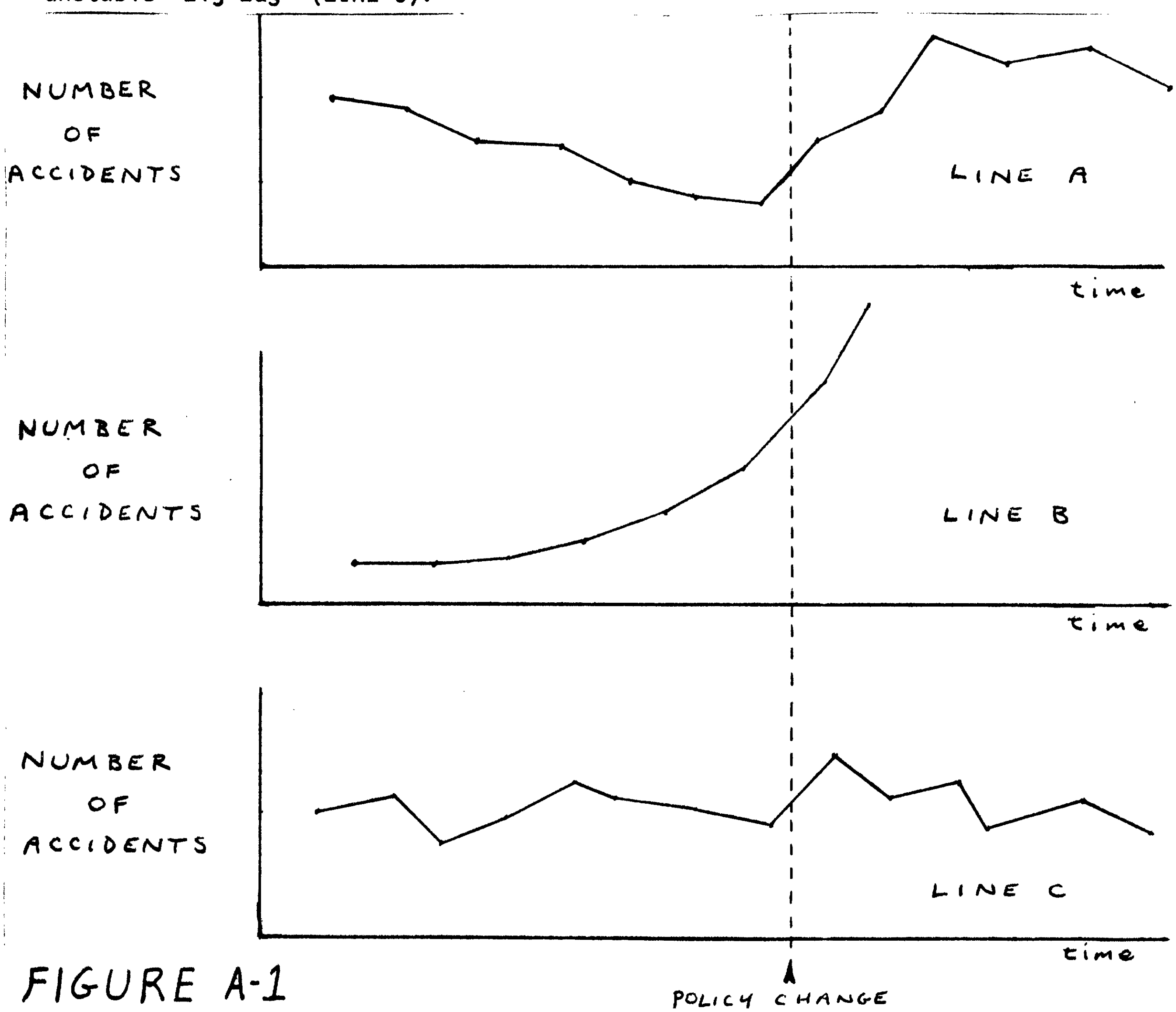


FIGURE A-1

Notice that if we had observed the time period just before and just after the policy change there would be no way of telling which of the three cases mentioned above would be a valid representation of the change being studied.

Attempting to identify the causal relationship between a change in policy and a specific measure of effectiveness is complicated by the fact that many other factors besides the change in policy under study may influence the specific measure of effectiveness being tested. Since it may not be feasible to isolate the change in policy under study and its specific measure of effectiveness from other factors which may influence the observed results, several hypotheses may have to be examined using a variety of statistics and controlling for various third-variable effects.

Part B: Miscellaneous Factors That May Influence the Accident Data

Several factors might influence the accident statistics in addition to the policy change concerning the eighteen year old drinking law. Before associating a change in the accident statistics with the change in policy, the effect of these other factors must be considered.

Any change in the driving environment or in the pattern of road use between 1969 and 1973 might affect the number of Massachusetts reported accidents and complicate the analysis of the drinking law's effects. If changed driving patterns were only caused by new 18-20 year old drinking habits, adjustments might not be required. However, other circumstances might introduce systematic bias in the monthly accidents figures. While most aspects of Massachusetts' driving environment are unlikely to have changed substantially and systematically during 1969-73, factors such as increased vehicle safety and seat belt use or a steady increase in miles driven are possible exceptions.

As is commonly done, we shall adjust for the number of miles driven per month. National Safety Council figures indicate a 5% per year increase in mileage but a steady fatality rate per (hundred million) vehicle miles through the sixties [12,13]. During four years, the effect of increased mileage on the number of accidents is substantial.

We shall also adjust for seasonality. Included under this heading are such monthly changes as weather, holidays and liquor consumption habits. Some of the seasonality effects manifest themselves in changes in the number of miles driven in certain months relative to others. Examples are months having holidays. However, National Safety Council figures show that the number of fatalities per hundred million miles driven is not the same for each month. Apparently other factors produce monthly changes in accident risk. Since we have only seven months data following the change in the drinking law, corrections for seasonally different months are important.

Another way accident statistics can change is when the way accidents are reported change. Different standards for what is considered an accident may be in effect during different years. Also, more or less accidents may be reported depending on the reporting mechanism. The advent of No Fault Insurance in Massachusetts beginning in January 1971 apparently had a significant impact on the number of reported non-fatal accidents. This fact might complicate the interpretation of certain accident statistics before and after 1971.

We assume that other factors influencing the monthly accident figures did not introduce any systematic bias in the data but contributed to random statistical fluctuations in the data.

Part C: Corrections for Mileage, Seasonality and No-Fault

As explained in the text, two methods of correcting for mileage, seasonality and no-fault were employed. Method I involved scaling monthly figures by estimated mileage and seasonality trends. Method II used the experience of drivers over 23 years of age as a control group.

Method I

Visual examination of several series suggested that the No-fault law produced a marked reduction in reported accidents beginning in January 1971. For non-fatal injury accidents the effect was most pronounced. (See the graphs of raw data in Section 2 of the text.) The impact of no-fault appeared to produce additional changes during 1972 and we felt that 1971-1972 was too short a time span to gauge what the 1973 impact ought to be. Accordingly, the non-fatal accident data for the four age groups appeared most affected and were not further analyzed.*

The monthly figures for the remaining sets of data were normalized to monthly figures per billion vehicle miles by dividing each monthly figure by the estimated number of miles driven each month. The estimates [15], based on gasoline consumption, are given in Table A-1 and reflected an approximately linear increase in mileage of 0.4% per month.

Next, a seasonal index for each of the twelve months of the year was developed for particular classes of accident data. These reflected the ratio of the particular monthly accident rate to one-twelfth of the yearly average. For example, the seasonality factor for February (for operators involved in fatal accidents) is 0.75 indicating the number of reported operators involved in fatal accidents during February (with 28 days) is three-quarters of the rate for a "normal" month with no seasonal effects. By dividing each mileage-

---

* Interestingly enough, no-fault appeared to have a different affect on reported accidents involving young and old drivers. A larger drop in reported non-fatal accidents apparently took place for older age groups.

Table A-1 Estimated Miles Driven

Period	Estimated Monthly Vehicle Miles* (Billons)												Estimated Average Miles Driven per Gallon of Gas*
	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	
1/69 - 12/69	1.94	1.75	1.98	2.01	2.23	2.23	2.21	2.31	2.12	2.25	2.06	2.30	12.39
1/70 - 12/70	2.21	1.91	2.11	2.14	2.25	2.43	2.32	2.36	2.25	2.29	2.12	2.41	12.38
1/71 - 12/71	2.19	2.02	2.24	2.24	2.20	2.51	2.41	2.53	2.36	2.36	2.34	2.47	12.36
1/72 - 12/72	2.25	2.19	2.43	2.30	2.55	2.64	2.45	2.71	2.42	2.50	2.50	2.52	12.37
1/73 - 9/73	2.48	2.30	2.58	2.37	2.76	2.69	2.61	2.79	2.40				12.35

*Vehicle mile estimates are based on Massachusetts Registry figures for monthly gasoline consumption and their estimated average miles driven per gallon of gas.

adjusted monthly figure by its seasonality factor, all the monthly figures were deseasonalized.

Seasonality factors were developed for particular sets of data using several sources and criteria. Because the series only spanned four years (1969-1972, plus January through September of 1973), and because large fluctuations in each year's monthly figures were present, it was not possible to calculate reliable seasonal factors from the data themselves. Information from the following sources were also used:

a) Seasonality factors developed by using national data for fatalities per mile driven between 1966 and 1968 [12].

b) Monthly fatality data for Massachusetts between 1966 and 1970 [15].

c) Monthly hard liquor sales during 1967.

d) Monthly non-fatal injury accident data for Massachusetts between 1966 and 1970 (the five years immediately before no-fault insurance).

To calculate seasonality factors using (b) and (d) above, monthly figures for each year were divided by one-twelfth the yearly total. For each month, the <u>median</u> ratio of the five years was chosen as the seasonal factor. (Using the median avoided sensitivity to extreme values.) Although these monthly figures were not based on accidents per miles driven but rather on the absolute number of accidents, it was felt that the monthly seasonal variation in miles driven was negligible compared to the seasonal effect. All age groups were assumed to have the same seasonal effect.

Examining the appropriate raw time series suggested that series involving fatalities and/or alcohol showed the same gross seasonality properties. For these data, sources (a), (b), and (c) above were used to calculate seasonality factors. All of the factors in sources (a), (b), and (c) resembled each other in relation to which months were high and low as well as in the value of the factors themselves. The median of the factors of (a), (b), and (c) was chosen for the overall factor for the alcohol related and/or fatality related statistics being tested. The final figures arrived at in order of the calendar months were .84, .75, .87, .95, 1.04, 1.03, 1.00, 1.09, 1.00, 1.11, 1.15, 1.12.

Factors derived from (d) were used for the non-alcohol related, non-fatal accident figures for all ages. The averages of the month to year ratios for 1971 and 1972 multiplied by twelve for total property damage figures were used as seasonal factors for all non-alcohol-related property damage accident figures. The values are not listed here because these sets of data were not used in the final analysis. (See the text for a list of those sets used in final analysis.)

Summarizing for Method I, series judged not to be grossly affected by No-Fault were adjusted

(1) By dividing each monthly figure by the estimated number of miles driven in that month;

(2) By dividing the result of step 1 by the appropriate seasonality factor.

<u>Method II</u>

The alternative method for adjusting the raw data involved using accident data for operators aged 24 and over as a control. Monthly figures for each type of accident involving young drivers were divided by the corresponding over-23 year old figures. The idea is that the over-23 year old group exhibits the same monthly mileage and seasonal fluctuations (and, perhaps the same no-fault insurance characteristics) as the younger age group--but is unaffected by the drinking law. Hence, resulting ratios would exhibit only random monthly fluctuations or systematic changes due to the new drinking law.

Examination of the gross properties of the time series for the different age groups lends credence to this assumption. Also, the yearly ratios remained fairly constant. The table below indicates the percentage of all fatal accidents involving operators in each age group.

Table A-2: Operators and Accident Involvements by Age Groups

Age Group	% of All Operators in Each Group			% of Fatal Accidents involving an Operator in Each Group			
	1969	1970	1971	1969	1970	1971	1972
under 18	2.6	2.8	2.6	6.6	7.0	7.2	6.8
18-21	7.6	7.7	8.0	14.8	15.1	14.4	14.7
21-23	8.2	8.3	8.1	14.5	14.2	13.1	14.0
over 23	81.6	81.2	81.3	64.1	63.7	65.3	64.5

Source: Massachusetts Registry of Motor Vehicle Annual Statement Reports

Figures 4, 5 and 6 in the text graph the D1-T1, A2-T1 and A4-T2 time series (numbers 1, 8 and 14) before and after adjustments for mileage and seasonal factors. Figure 5 compares the raw data for A2-T1 with the addustments using both Methods I and II. These figures are discussed in the body of the text.

Appendix II -- INTERPRETING THE TIME SERIES

Part A: Correlation and Interruptions

In the time series approach, we have a set of temporally ordered (adjusted) observations divided into two groups; those pertaining to the time period preceding the policy change and those pertaining to the time period following the policy change.

The statistical methods appropriate for analyzing the data depend heavily on the characteristics of the data. If the figures are uncorrelated with one another, e.g., independent and randomly fluctuating, then a variety of standard statistical procedures could be applied. However, if some correlation among the data points is present, e.g., some residual seasonality or some kind of significant upward trend, then much more complicated methods are required and minor interruptions become harder to identify. A common measure of the degree of correlation among points in a time series is the autocorrelation coefficient of order k, where k is the number of time units separating the points. If enough observations of the time series are available, then the autocorrelation coefficient ρ_K can be estimated by using the formula [8]

$$r_K = \frac{\sum_{i=1}^{n-K} (X_i - \bar{X})(X_{i+K} - \bar{X})/(n - k)}{\sum_{i=1}^{n} (X_i - \bar{X})^2/n} ,$$

where

X_i = i^{th} observation, r_K = estimate for autocorrelation coefficient ρ_K

n = number of observations and $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X$.

Theoretically, $-1 \le \rho_K \le 1$. When $|\rho_K|$ is near 1, a high degree of correlation is present. When $|\rho_K|$ is near 0, little or no correlation is present. A plot of r_K versus k for a time series is called a correlogram. The general appearance of a correlogram along with the values of r_K tell a lot about the characteristics of a time series.

An independent sequence of normally distributed observations would have a correlogram with values r_K fluctuating close to 0. A periodic series would exhibit a periodic correlogram with high values of r_K repeating at various intervals. A series with a trend might exhibit large values for some of the r_K and an appearance of a sloping line as the general shape of the correlogram. (See the examples in Figures A-2 through A-5.)

To test for significant correlation in a correlogram, we use r_K as an estimate of ρ_K and test the hypothesis that $\rho_K = 0$. Confidence limits for r_K are available if we assume independent mormally distributed fluctuations and an r_K that behaves in the same way as serial correlation.[3, 8] The independent, normal assumptions are often used for accident data[4] since it seems reasonable that random fluctuations in the number of accidents during one month would be independent of fluctuations in other months. For a 50 term series, we would accept the hypothesis that $\rho_K = 0$ if $-0.28 \le r_K \le 0.28$ for any particular value of K.

In fact, some additional complications arise. As k increases, the confidence limit gets wider since fewer and fewer terms are available for calculating r_K . Thus, in the case of a series with fifty terms, about the first thirty values of r_K are worth calculating. Since the confidence limits apply to a single r_K value, all thirty values are not expected to fall within the 95% linits. If the values of r_K were independent and ρ_K did in fact equal

0, then the probability that more than three fell outside would be about 0.06. Accordingly, we shall accept the $\rho_K = 0$ hypothesis if no more than 3 are outside, if these are not too large (e.g., if $|r_K|$ is still < 0.4) and if the shape of the correlogram is not pronouncedly periodic or suggestive of a trend.

A correlogram can be calculated for each of the time series for the fifty months preceding the policy change. If the hypothesis $\rho_K = 0$ is accepted, we assume that the basic characteristics (i.e., independent normally distributed fluctuations) in the data apply to the points following the policy change. These points were not included in the original correlogram since we expect something might shift due to the policy change and the points by themselves, did not contain enough observations to compute r_K.

The pre-policy change and post-policy change are then compared to determine whether they are significantly different. The null hypothesis is that they have the same mean and standard deviation (i.e., the policy change had no effect). Once again a 5% confidence level is used. The criterion was defined by a one-tailed t-test for the difference between two means since we suspect that the most likely alternative to equal means should be for the mean of the post-group to be larger than the mean of the pre-group. In mathematical terms, the hypothesis that the period A mean is the same as the period C mean* for any well-behaved series is rejected if

$$\bar{X} - \bar{Y} > t_{95,55} \, V_p \sqrt{\frac{1}{50} + \frac{1}{7}} \; ,$$

where $\bar{X} = (1/50) \sum_{i=1}^{50} X_i$ and $\bar{Y} = (1/7) \sum_{i=51}^{57} X_i$ and X_i is the observed

---

* Period A covers January 1969 through February 1973; Period C covers March through September 1973.

(adjusted) rate during month i for the particular series. Here we let $t_{95,55}$ represent the cutoff value for a one-sided t-test at the 95% significance level with 55 degrees of freedom -- $t_{95,55} \approx 1.7$. The pooled estimate of the standard deviation of the observations is used so that

$$V_p = \sqrt{(49s_x^2 + 6s_y^2)/55}, \quad \text{where } s_x^2 = (1/49)\sum_{i=1}^{50}(X_i - \bar{X}_i)^2$$

and

$$s_y^2 = (1/6)\sum_{i=51}^{57}(X_i - \bar{Y})^2 .$$

For those series whose correlograms show signigicant autocorrelation, the analysis becomes very difficult and less discriminating since correlated but unexplained fluctuations can give rise to large shifts in a time series in the absence of any other changes.[1,9] Removing further trends is hard since the effort may create further autocorrelation in the adjusted series.[9] Earlier studies of Massachusetts fatalities have indicated little correlation and we hoped that enough series would appear well behaved to enable the above t-test to be used. As it turned out this was the case.

In addition to the correlograms, a further check was employed to test if a rise in the accident rate may have started in the months just previous to the policy change, but somehow did not affect the correlogram. The mean of months 44 through 50 was checked against the mean for months 1 through 50 to see if that period had a significant difference as well. If it did, it would indicate that the change in level may have occured before the policy change.

Part B: The Adjusted Well-behaved Time Series

In brief review, two methods were used to adjust the data for increased mileage and seasonal considerations. A correlogram was used to check which adjusted data were amenable to the t-statistic hypothesis test for a significant difference between two means. This test was then performed and results obtained for that available accident data. As indicated in the main body of the this report, series 1, 3, 5, 8, 10, 19, 20 and 21 were adquately behaved and amenable to the t-statistic hypothesis test. In this appendix, some of the details of applying the procedures are discussed.

Correlograms were computed using the computer at the M.I.T. Information Processing Center. As a check, a psuedo-random normally distributed series of fifty terms and several series of linear and periodic data were correlogrammed and produced the expected results. Figures A-2 through A-5 graph the correlograms of the normal samples and several well-behaved and correlated series of accident data. Note the apparent periodicity of the driving under the influence violations.

For the available accident data, the mileage and seasonality corrections accounted for only a small part of the variations from month to month. This result is to be expected since the numbers of accidents occurring per month are small for many categories we considered. The seasonal adjustment did reduce the total variance of the series as was expected, and the mileage adjustment did remove an increasing trend of about 5% per year. The random fluctuations were so large, however, that correlograms of some of the acceptably-behaved series <u>before</u> mileage and seasonal adjustments, were slightly correlated but otherwise not much different from the correlogram for the adjusted data. Our interpretation of this result is that it shouldn't affect the analysis substantially if the seasonal or mileage adjustments are generally reasonable though perhaps not very precise.

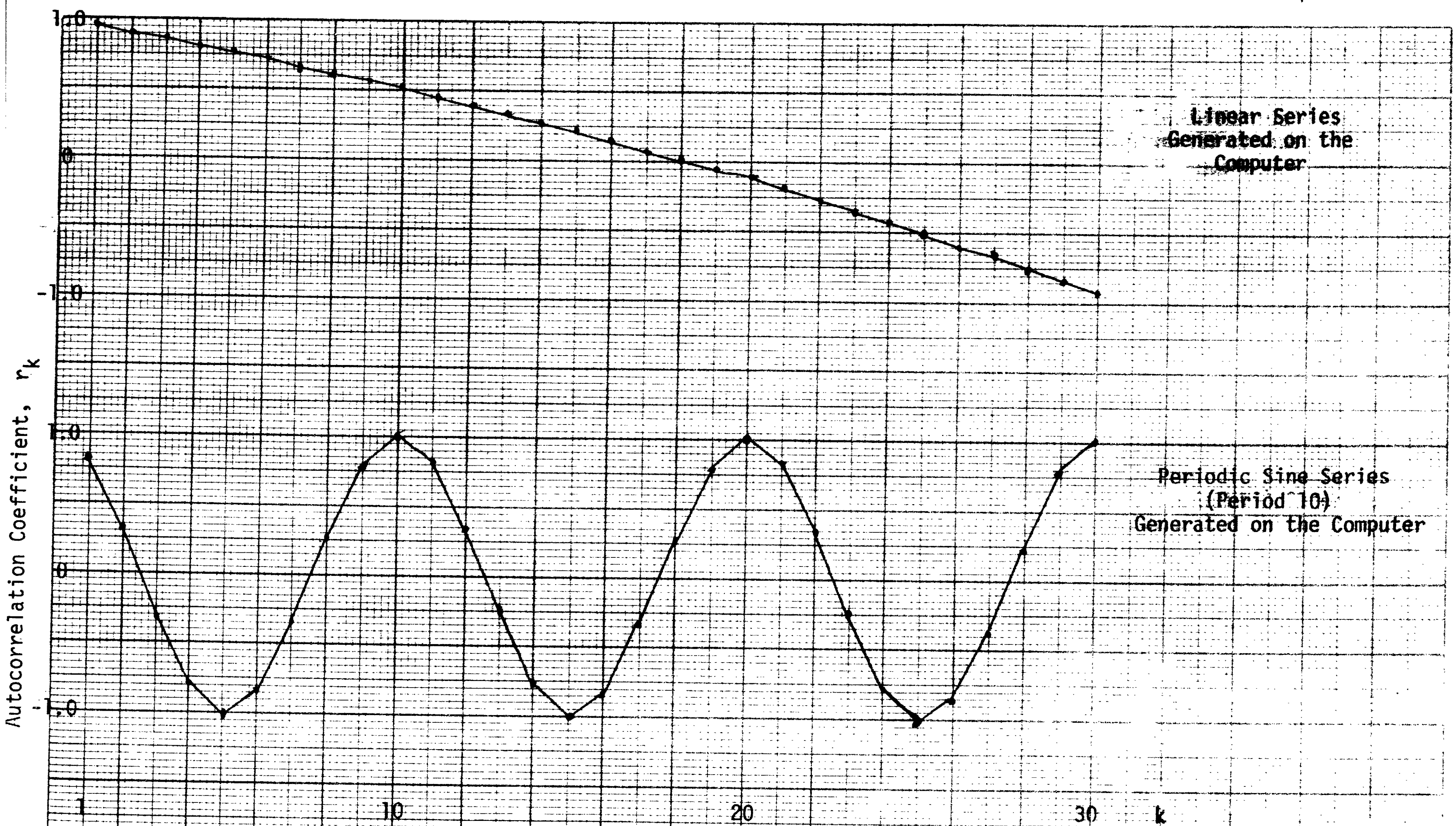
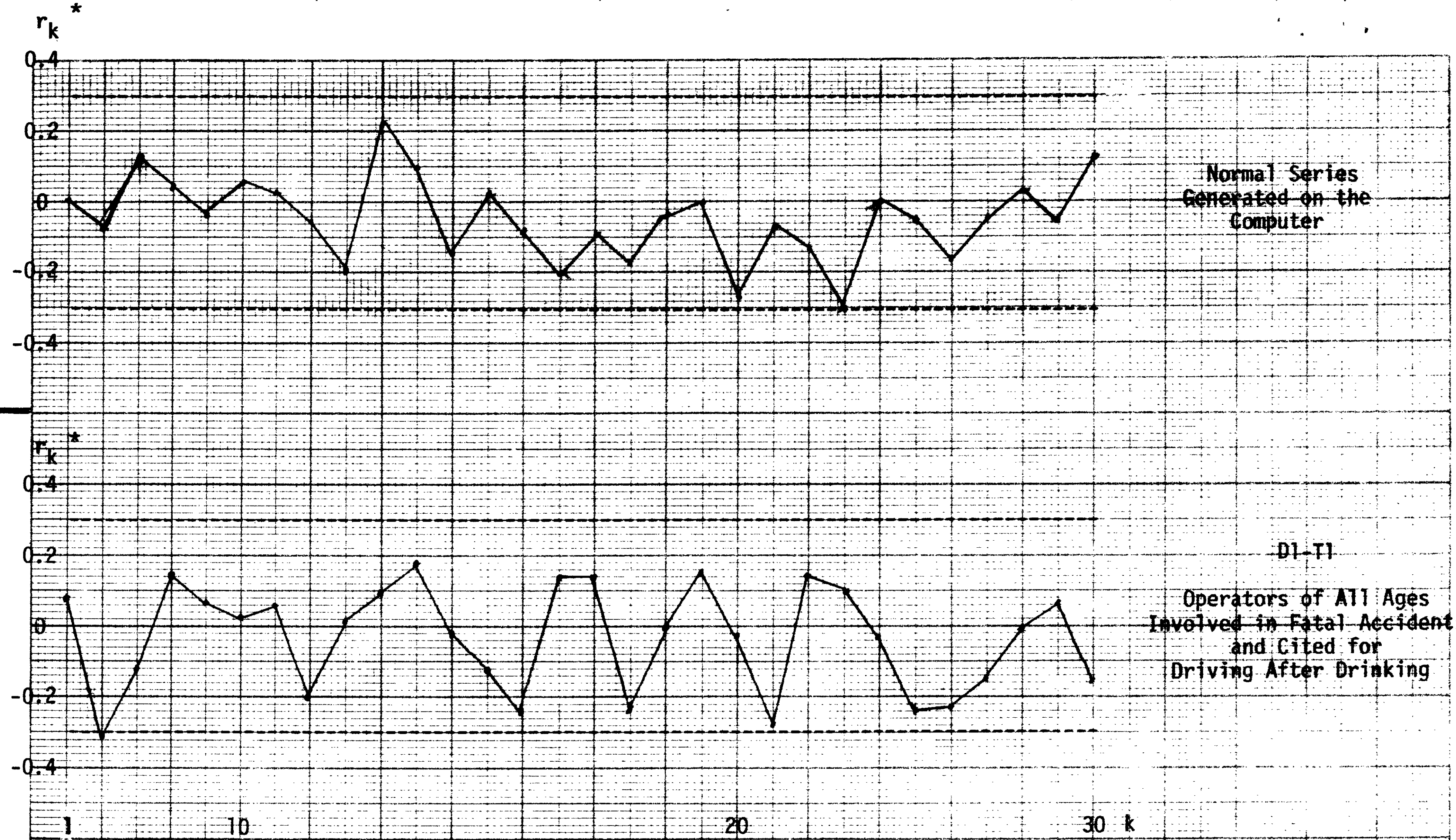


Figure A-2: Sample Correlograms for Simulated Data



---------- The dotted lines represent the range between which 95% of the r_k values should be if the series is a sequence of independent, identically distributed normal random variables.

* r_k is the autocorrelation coefficient of order k .

Figure A-3: Sample Correlograms for Selected Series (Part 1)

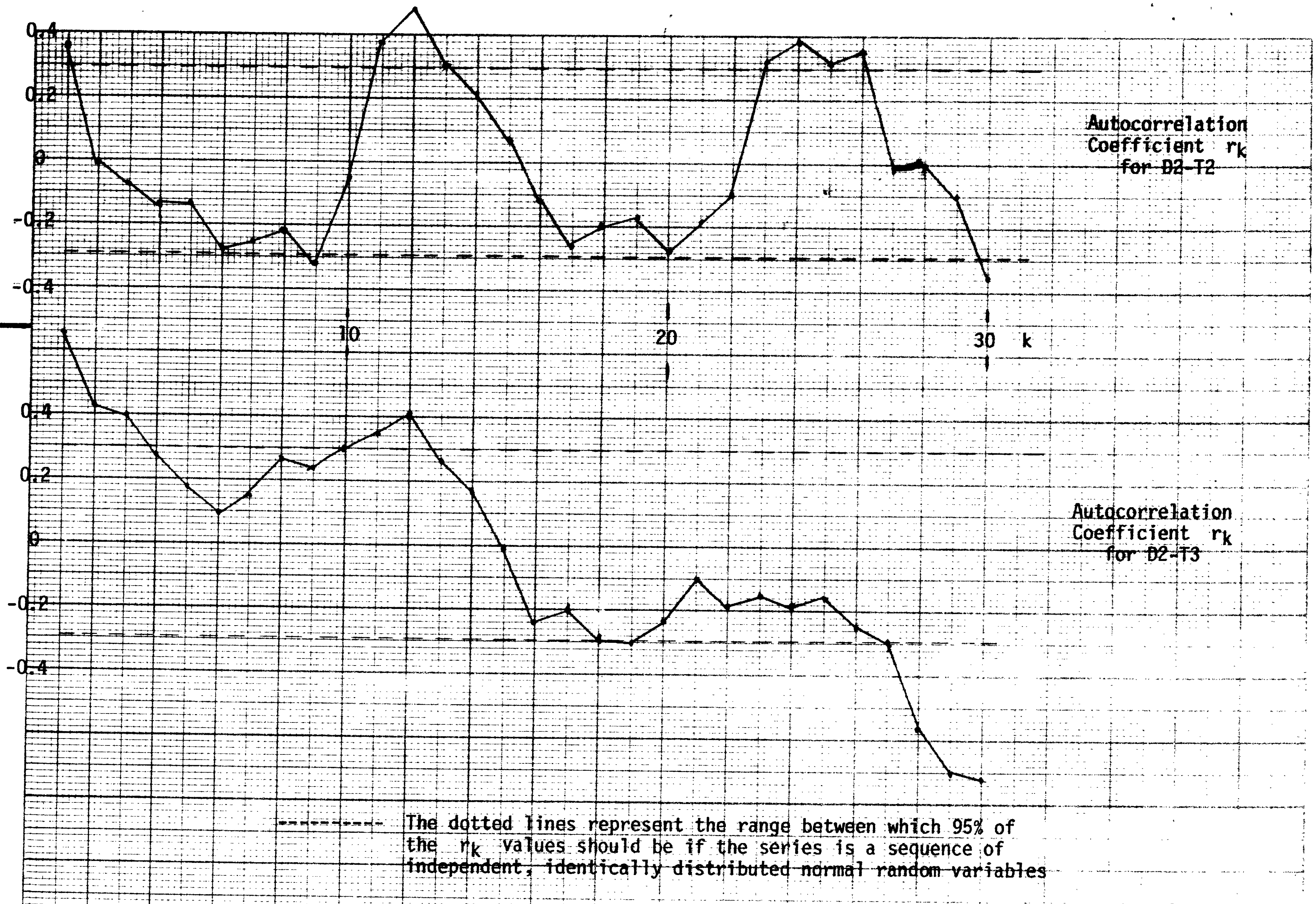


Figure A-4: Sample Correlograms for Selcted Series (Part 2)

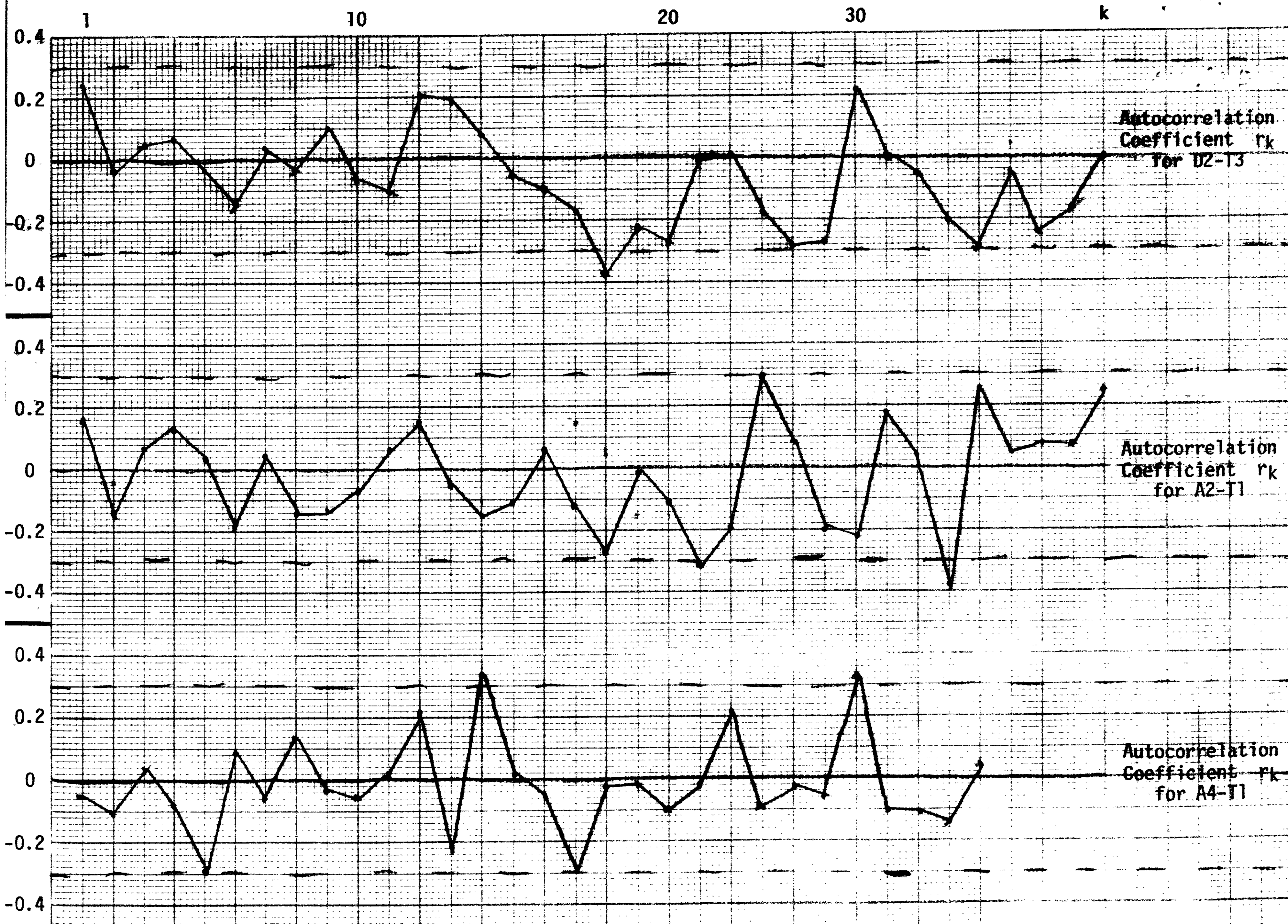


Figure A-5: Sample Correlograms for Selected Series (Part 3)

Appendix III -- DATA DIFFICULTIES AND FUTURE RESEARCH

In addition to the introduction of no-fault insurance (which affected the consistency with which accidents are reported), other data difficulties are possible. In trying to develop inferences from the data concerning the effect of the policy change, other historical events occuring near the time of the policy change could effect the accident statistics. The gasoline shortage beginning in October of 1973 is just such an event. This study used data only through September 1973. Use of data beyond October 1973 would have to take into account possible effects of reduced speed and driving on the general accident rate.

Another source of difficulty may be in the differences in reporting for different age groups and the completeness of reporting for certain types of accidents. For fatal accidents, reporting is generally regarded as consistent. However, this may not be the case for drinking-related violations. If most violations for one age group were reported while other groups were less frequently cited, then a rise in that one group's rate could produce a sharp rise in the total. This might explain why the impact of the increase in 18-21 year old fatalities did not register on the total number of accident fatalities (all of which are assumed to be reported) but did register on the total number of the after drinking violations in fatal accidents. Few of Massachusetts' fatal accidents involve drinking violations (around 20%) compared with the 50% figures developed by controlled studies of alcohol's role on fatalities. Hence, the above reasoning is a likely explanation.

Massachusetts State Police lab test results are subject to other biases but are closer to the 50% figures and are consistent with the section 4 results.

Medical Examiners are required by law to send the lab blood specimens of all individuals killed in auto accidents.* The lab then determines the alcohol content of the specimens. However, the results can provide only a crude indication of the percentage of accidents that are alcohol related since operators involved in fatalities are tested only if they themselves die within hours of the accidents. Thus, the percentage of specimens from operators that have significant alcohol content may be biased estimates and are subject to large fluctuations since the total number of operators of all ages tested each year is around 140. During 1972, 57% of the 21-and-over operators and 47% of the under 21 operators tested had blood alcohol contents above 0.05%. In 1973 the corresponding percents were 59% and 55%. The 47% and 55% figures for 18-20 year olds during 1972 and 1973 roughly consistent with the (more reliable) findings for the 18-20 year old fatality data. To see this assume that 47% (or 6.4 per month) of all 1972 fatalities involving 18-20 year olds were alcohol related. Assume also that 56.6% (or 10.8 per month) of all post March 1973 fatalities involving 18-20 year olds were alcohol related. The difference--i.e., 10.8-6.4 = 4.4 fatalities per month--is not too far from 5.3 per month increase in alcohol-related fatalities among 18-20 year olds.

Future studies may make use of more sophisticated techniques to analyze data such as the property damage and non-fatal accident figures that were too poorly behaved to be studied here. The Box-Tiao method of correcting for autocorrelation is one possibly--a computer program is available for using this model on a time series [5]. Other methods to analyze data with increasing trends could include regression analysis with or without provision

* Here, killed is defined as dead within 4 hours of the accident.

for correlated error. However, most such analyses are complicated since few data points following the policy change are available.

More data would be helpful in analyzing the series to check for transient effects after March 1973. Using such data is complicated by the gasoline shortage effect. Using more data before 1969 may not help significantly since the real accuracy of the tests depend on the availability of the post policy change data.

Further breakdowns of the alcohol-related violations by age might help illuminate the effect of reporting biases. These further breakdowns, however, would involve very small numbers and exhibit large fluctuations relative to the mean. In addition, they would be subject to the same biases and underreporting discussed earlier for the fatal number of operating-after-drinking violations.