A DISCUSSION OF ISSUES AND CONCLUSIONS ON GASOLINE LEAD USE AND HUMAN BLOOD LEAD LEVELS

Ralph A. Bradley*

SUMMARY

This report was written at the request of the Ethyl Corporation, seeking external review from a statistical viewpoint of the use of data from the Second National Health and Nutrition Examination Survey (NHANES II) to contend that there is a causal effect of gasoline lead use on population blood lead levels. Information provided is listed. The NHANES II data were collected to provide population estimates of many health and nutrition characteristics of the U.S. population valid for the midpoint of the survey which was conducted from 1976 to 1980. The survey was designed as a cross-sectional study and not as a study to examine population changes with time. It is the general conclusion of this review that NHANES II data cannot be used for any study in time, including changes in human blood lead levels, because of many sources of bias that arise in attempting to obtain population estimates at times within the time interval of the survey.

*Statistical Consultant, 325 Hickory Hill Drive, Watkinsville, GA, 30677.

Houk presented a statement to a subcommittee of the U.S. House of Representatives implying that similar time trends in human blood lead levels from NHANES II data and in gasoline lead use, both apparently decreasing in time, established a causal relationship. We examine this logic in the second section of this report, note that it is a common error, and indicate that a similar conclusion could be reached with any other variable substituted for gasoline lead use so long as it exhibits a trend over the same time period.

The third section of this report deals with the NHANES II data and time. Possible biases are discussed as is the purpose of the study. It is clear that the study was to be cross-sectional and to provide normative data for comparative use with data from future surveys.

The Ethyl Corporation commented on the Houk presentation and objected to his conclusion and proposed standards for use of lead in gasoline based on that conclusion. Additional analyses of NHANES II data are given and these are discussed in our fourth section. The Ethyl Corporation analyses are initially based on individual data and establish relationships between blood lead level and personal characteristics and location of the individual. We believe that use of such relationships on an individual basis is appropriate and the only way to avoid need for unbiased population estimates by time periods from the NHANES II data. We do not see any major grounds for criticism of these analyses, although the present author would have continued throughout on an individual basis as he has done in a separate section of this report. In that section, it is shown that the effect of gasoline lead use on blood lead is completely non-significent after adjustment for individual personal characteristics, population density, and time. One cannot conclude that gasoline lead use does not affect blood lead level but the analysis does show that the residual effect of gasoline lead use after adjustment for time is negligible.

An ICF Incorporated review of the Ethyl Corporation comments is shown to produce criticisms that are either invalid or irrelevent. Indeed, ICF in further analysis of the NHANES II data uses methods similar to those criticized when used by the Ethyl Corporation. The ICF analysis attempts to answer some issues raised by others and to use some of the independent variables already used by the Ethyl Corporation. However, neither biases in NHANES II data nor time dependence of data are recognized. Consequently, this new analysis is subject to the same criticisms applicable to the original Houk presentation.

No causal effect of use of lead in gasoline and human blood lead levels has been established nor can be established from the NHANES II data. At best such data may suggest relationships, real or spurious, for careful confirmatory experimentation or study. The latter has not been done.

INTRODUCTION AND BACKGROUND

This report has been written at the request of the Ethyl Corporation. The assignment given to the author was to provide an expert and objective review of reports and statements provided and listed below from a statistical point of view on the use of data collected in the Second National Health and Nutrition Examination Survey 1976-80 (NHANES II) associating population blood lead levels with reduction in lead usage in gasoline. The author of this report is a statistician qualified to comment on statistical procedures used and the statistical interpretations made. The author claims no expertise in other technical areas that may be relevant such as nutrition, blood chemistry, air pollution, food packaging, paint, or automotive exhaust emission.

Information provided on the key issues, the apparent association of blood lead levels in the human population and the use of lead in gasoline, was as follows:

[1] Statement by Dr. Vernon N. Houk, Acting Director, Center for Environmental Health, Centers for Disease Control before the Subcommittee on Environment, Energy and Natural Resources, Committee on Government Operations, U.S. House of Representatives, Washington, D.C., April 14, 1982.

[2] Ethyl Corporation's Comments following EPA's Public Hearing on Proposed Rule for Lead in Gasoline, Washington, D.C., September 7-8, 1982 - Submitted to EPA's Docket No. A-81-36, October 8, 1982.

[3] A Review of the Ethyl Corporation's Comments on EPA's Proposed Rule for Lead in Gasoline prepared for the U.S. Environmental Protection Agency, Office of Policy and Resource Management, Office of Policy Analysis, under Contract No. 68-01-5845, November 1982, ICF Incorporated.

[4] The Relationship between Gasoline Lead Emissions and Blood Poisoning in Americans, prepared for the Office of Policy and Resource Management, Office of Policy Analysis, under Contract No. 68-01-5845, October 1982, ICF Incorporated.

Supplementary information available is as follows:

[5] Mahaffey, K.R., Annest, S.L., Barbano, H.E. and Murphy, R.S. (1979), Preliminary Analysis of Blood Lead Concentrations for Children and Adults: HANES II, 1976-78, Trace Substances in Environmental Health 13, 37-51, 1979.

[6] Plan and Operation of the Second National Health and Nutrition Examination Survey 1976-80, DHHS Publ. No. (PHS) 81-1317, National Center for Health Statistics, Hyattsville, Md., July 1981.

[7] Public Use Data Tape Documentation, Tape No. 5701, National Health and Nutrition Survey, 1976-1980, National Center for Health Statistics, Hyattsville, Md., February 1982.

[8] Hematology and Biochemistry Data Tape, Catalog No. 5411, National Health and Nutrition Examination Survey 1976-1980, National Center for Health Statistics, Hyattsville, Md., undated.

[9] Annest, Joseph L. and Pirkle, James (undated), Re: Chronological Trend in Blood Lead Levels in the U.S. Population, NHANES II, 1976-1980. Preliminary draft manuscript.

We shall refer to these references by number in the remainder of this report.

Remaining sections of this report provide review of the main references [1] to [4]. A summary is provided at the beginning.

THE HOUK STATEMENT

Houk [1] concludes his presentation with the statement,

> "With continued reduction in lead use, we could widen the margin of safety for the lead hazards we have already placed in the environment."

His preceding statement,

> "The reduction in the lead being added to our environment has already resulted in a 37% reduction in the mean blood lead level and a 50% reduction in the amount of lead being added to gasoline."

suggests that continued reduction of the use of lead in gasoline is the major means by which continued reduction in lead use is to be achieved, a suggestion confirmed by EPA's "Proposed Rule", FR 47:167, 38078-38089 and the September 7-8, 1982 Public Hearing. However, Houk himself in his statement notes

> "During the last several years the Federal Government has taken a number of actions to limit the exposure of Americans to lead residues. These include limits on the allowable lead content in paint, water, ambient air, processed food, and gasoline."

Several conclusions arise immediately from these quotations.

The first clear conclusion is that Houk is basing his statement on a "confounded" experiment. It was noted that lead in paint, water, processed foods, and ambient air has been reduced as well as lead in gasoline. This suggests immediately that other downward trends over the same time period exist and are comparable to the trend plotted for lead used in gasoline against time in Figure 2 of the Houk presentation. Indeed, the Ethyl Corporation [2] in their Figure 13 presents just such a plot for lead use in solder against time which again agrees well with Houk's plot of average blood lead levels versus time. An incidental conclusion is that the Ethyl Corporation could have adjusted its interrupted scale for lead in solder to produce a better apparent coincidence of the two trends over the years 1978, 1979 and 1980, those years with actual data.

Houk notes that there has been a 50% reduction of use of lead in gasoline and a 37% reduction in blood lead levels over the time period considered. Figure 3 of [1] is a plot of average blood lead levels against lead used in gasoline. The slope of the fitted line is approximately 0.106; the intercept at zero lead use in gasoline is 4.7 micrograms per deciliter of blood lead, demonstrating clearly that other sources of blood lead occur. Further, there has been no discussion of the physiological process relating blood lead to lead exposure from any source. The author has no information on any models that may have been proposed to explain the process. It would not be bizarre to suppose that blood lead level depends on the logarithm of lead exposure concentration. It would not be bizarre to suppose that there is some tolerance or threshold level of lead exposure below which exposure does not lead to blood lead. In this latter situation the blood lead level intercept might be considerably higher than indicated at zero lead use in gasoline. Some clear notion of the physiological process leading to lead in blood is necessary before any data relating blood lead to lead exposure can be interpreted.

Let us suppose for simplicity that there is a perfect linear relationship between blood lead level y and time t, say $y = a_0 + b_0 t$. Let us similarly suppose a perfect linear relationship for lead used in gasoline x and time, say $x = a_1 + b_1 t$. From the second equation, $t = (x-a_1)/b_1$, and substitution in the first equation yields $y = (a_0 - \frac{a_1 b_0}{b_1}) + \frac{b_0}{b_1} x = a_2 + b_2 x$. Thus, in this simplistic deterministic situation, lead in blood may be perfectly predicted by lead use in gasoline. But suppose that x was the average weight of new cars sold or, indeed, the rate of inflation. Then equally good prediction of blood lead level would result. In regression analysis, relationships like those discussed above are considered, but they are imperfect since the variables involved fluctuate subject to some unexplained variation. Regression analysis is useful in prediction. Regression analysis never assures a causal relationship between two related variables. At best, regression analysis may suggest a relationship for further investigation, perhaps through an experiment designed to confirm an hypothesized causal effect if it exists. If similar time trends exist for blood lead, lead use in gasoline, average weight of new automobiles sold, and the rate of inflation, there is no more justification for suggesting a causal effect for one of these variables than another in the absence of additional information.

The literature of statistics abounds in nonsense correlations and misguided interpretations. Classic examples are the high correlations between the number of storks and the birth-rate in Sweden and between mortality in England and Wales and the proportion of marriages in the Church of England, both examples yielding high correlations because all variables were highly time-dependent. Research workers in public health and epidemiology seem particularly prone to interpret time-dependent relationships as cause and effect relationships; perhaps through an intense search for results. Yule* (1926), in his presidential address to the Royal Statistical Society, discussed time induced correlation and Greenwood emphasized the problem in his published comments on the paper. Yule noted a correlation of 0.9511 between standardized mortality per 1000 persons in England and Wales and the proportion of Church of England marriages per 1000 of all such marriages, 1866-1911. He stated:

> "And I can imagine some enthusiast arguing that the fall in the proportion of Church of England marriages is simply due to the spread of scientific thinking since 1866 and the fall in mortality is also clearly to be ascribed to the progress of science; ... But most people, I think, agree with me that the correlation is simply sheer nonsense, that it is absurd to suppose that the two variables in question are in any sort of way, however indirect, causally related to one another."

Greenwald commented:

> "The practical problem involved in the paper was of particular importance to all concerned with medical statistics, because of the tendency to infer conclusions from the apparent concurrence of time-series; ... ."

Current medical research workers are still enthusiasts and the

*G. Udney Yule(1926), "Why do we sometimes get nonsense correlations between time series? - A study in sampling and the nature of time series," Jour. Roy. Statist. Soc. 89, 1-64. Greenwood comments, p.68.

tendency to infer conclusions persists. To note associations between variables may sometimes suggest fruitful directions for research; to use associations between variables as a basis for governmental action is at best poor science.

NHANES II DATA AND TIME:

In the preceding section, we have assumed the validity of the data used by Houk and shown that, even with this assumption, the conclusions reached by Houk cannot be sustained. We now examine the data and find major problems that further deprecate the conclusions reached.

NHANES II was designed to be a cross-sectional study of health and nutrition of the U.S. population. This is amply borne out by a number of statements. NHANES II was not designed as a longitudinal study in time and consequently cannot be used in that way as done by Houk and by Amest and Pirkle [9]. In [6, p.1], it is stated:

> "In 1969 the Department of Health, Education and Welfare established within NCHS a continuing activity to measure the nutritional status of the U.S. population and to monitor changes in status over time."

In [6, p.2], the purpose of the sequence of surveys was made clear:

> "The continuing responsibility for measuring and monitoring the nutritional health status of the U.S. population meant that the first assessment survey, NHANES I, would be followed by later assessment surveys. These would permit comparisons with the NHANES I baseline data and thereby allow measurements of changes over time."

It is clear that changes over time were to be measured by comparisons from successive surveys and not from data internal to a single survey. As we understand it, blood lead data were not collected in

NHANES I and changes over time for blood lead could not be evaluated.

Mahaffey, Annest, Barbano and Murphy [5] wrote a preliminary paper on blood lead analysis before the NHANES II data were available in their entirety. They wrote:

> "Normative data obtained from the total sample after the survey has been completed will provide the basis for assessment of relationships between blood Pb and selected health and nutritional indices measured for persons 6 mo. - 74 yr. in the general population. Comparisons of data obtained in NHANES II with other surveys will hopefully augment our understanding ··· . Results also provide a national data base for assessing changes over time to evaluate intervention efforts on a broad scale."

Data collected between February, 1976 through February, 1978 were used in this preliminary paper. The authors noted:

> "The preliminary NHANES II data presented in this study overrepresent very young, elderly and low-income individuals and should not be used as normative data."

In [7], it is made clear that biases may exist over the time allotted to the survey and that the cross-sectional survey was designed to provide valid estimates for the U.S. population at the midpoint of the survey:

> "Adjusted sampling weights were then computed for 76 age, sex, and race categories in order to inflate the sample in such a manner as to closely reflect the estimated civilian non-institutionalized U.S. population ages 6 months - 74 years at the midpoint of the survey (March 1, 1978)."

It seems clear that biases exist in the sample data when used by time intervals (six-month periods by Houk). Other sources of bias exist- the scheduling of stands (examination locations) for NHANES II was arranged so that the North was avoided in winter [6,p.27]. Much is said in [6,7,8] about use of appropriate weights for population estimates but nothing is noted about changes of responses with time within the time period of the survey. One can only assume that biases such as those indicated by the authors of [7] are not adjusted for possible time effects in producing mid-survey estimates since no comments or information are provided on this issue.

The Ethyl Corporation [2] in Table 3 provides some information on possible biases affecting Figure 2 of Houk's presentation relative to the plot of average blood lead levels against time. Both the % Children less than six years old and the % Residents in urban areas decrease with time, presumably contributing to the apparent decrease in average blood lead levels over time. Summaries in their Tables 4 and 5 show high correlations comparable to those used by Houk to support his conclusions. It is not possible to measure the extent of the biases forming Houk's conclusions but they clearly exist and the summary points made by the Ethyl Corporation on page 10 of [2] cast major doubt on the validity of Houk's Figure 2.

It is clear that the Houk Statement is based on the unpublished work of Annest and Pirkle [9]. They emphasize the quality control on the actual blood lead determinations and the complex survey design necessitating access to specially designed statistical programs leading to "unbiased estimates of standard errors of weighted means." Nothing is said about biases with time noted above and it can be assumed that no adjustments were made. For their Figure 1, on which Houk's Figure 2 was clearly based, unbiased estimates of population blood lead levels are needed for each time point plotted. There is no evidence that such estimates were or could be obtained. It is interesting that Annest in [9] does not heed Annest in [5] where it was stated that normative data were sought and that biases existed in the preliminary data.

In Figure 2 of Houk [1], the scales for both horizontal and vertical axes have been changed from those of Figure 1 of [9] to emphasize the apparent time trend in average blood lead levels. The apparent trend is much less apparent in Figure 1 of [9] and it might easily be concluded that it has been induced by biases noted above. Interruption of the vertical scale by Houk contributes to the apparent effect. In Houk's Figure 2, the vertical scale, again interrupted, for lead use in gasoline, a scale with entirely different units of measurement, has clearly been adjusted to produce as near coincidence of the two time trends as possible. It is clear that dubious statistical graphics have been used to emphasize the apparent association of the two time trends.

In summary, it is clear that NHANES II was designed to be a cross-sectional study providing valid estimates for the survey midpoint. In order that the estimates be valid, it appears that it was necessary to assume that the population was static over the four-year period because there was no "balancing" on age groups, locations and so on by time nor were adjustments for time effects proposed or discussed. Given these assumptions, it is indeed remarkable that the data would be used by Houk [1] and Annest and Pirkle [9] to argue for a reduction in blood lead levels with time, whether or not such a reduction did in fact occur.

ETHYL CORPORATION'S COMMENTS

Ethyl Corporation's Comments [2] follow several directions. The first two sections deal with the effects of standards on gasoline lead usage and the energy penalties involved. Basically, calculations are presented based on EPA's projected gasoline volumes and associated lead usage. The basic contention is that EPA has greatly underestimated the energy penalty. Calculations in these sections have not been questioned in reviews and the issue of the energy penalty is not within the expertise of the writer. Section 4 addresses the need for a regulatory impact analysis. Section 3 addresses the use of NHANES II data as a basis for the proposed rule for lead in gasoline and is the subject of our comments.

Part A of Section 3 relates to the apparent correlation of blood lead and lead use in gasoline. In subsection 1, it is pointed out that NHANES II was a cross-sectional study, that biases in the data on blood lead over time exist, and that similar time trends exist for blood lead, gasoline lead use, % children less than 6 years old, and % non-rural residents. These issues have been discussed above and we concur that the study fails to show that decrease in gasoline lead usage is responsible for the decrease in blood lead. We regard the issue as open; neither proved or disproved by the data, the questionable analysis, or the dubious scientific logic.

In Section III, Subsection A2, the Ethyl Corporation presents its own "in-depth statistical analysis" of the NHANES II data. We comment below on this analysis and show an alternative analysis in a later section. The Ethyl Corporation's analysis is a reasonable one, intended to demonstrate that the NHANES II data may be shown to provide an explanation for the decrease in blood lead levels over the four years of the survey not associated with the parallel reduction in lead use in gasoline. Our general view, however, is that it is not possible to reach any valid conclusions on the causal effect of lead use in gasoline on blood lead from the data from this survey.

The Ethyl Corporation analysis begins on a per individual basis. This has the advantage that unbiased estimates of population blood lead levels are not required for each time period as is the case with the Houk analysis. It was necessary to record the time of sampling for each individual, the population density of the individual's location, and the lead usage (or exposure to lead use in gasoline) of the individual. Such data were difficult to construct, but a serious effort to do so was made. Complete data were constructed for 9402 individuals from the NHANES II sample.

A large regression model was established permitting dependence of the dependent variable y, the logarithm of blood lead level, on 36 personal identifier variables (combinations of age, sex, race, and income level) and 64 location variables. Basically the model was

$$y = \sum_{1}^{35} b_i x_i + \sum_{1}^{64} c_j z_j + e \ , \qquad (1)$$

where $x_1,\ldots,x_{35}$ have values zero or one as indicators of characteristics, while $z_1,\ldots,z_{64}$ are similar indicators of locations, and e is the random error associated with y. We prefer the model

TEH 0532999

DUP050034269

$$y = b_0 + \sum_{1}^{35} b_i x_i + \sum_{1}^{63} c_j z_j + e, \tag{2}$$

where b_0 represents the average log blood level, and use (2) in our analysis even though the two models would be identical for the purpose of the Ethyl Corporation analysis. Note that only 98 independent variables plus a constant or only 99 independent variables without the constant may be used because otherwise functional dependencies arise. The purpose of the analysis based on (1) was to predict blood lead levels for each of the 64 locations for a "standard individual" described by $x_i = 0$, $i=1,\ldots,35$. The predicted blood lead levels by location are given as C_3 in Table 6 of [2]. The data of Table 6 were used for a second set of regression analyses, now with a sample size of 55, one standardized observation for each location, (complete data not being available for 9 locations).

Subsets of the independent variables, time, population density, and lead usage for the locations were considered in the second set of regression analyses as independent variables. The general effect was that time was by far the best independent variable to predict blood lead at a location and lead usage the weakest, with little predictive ability. Note also that the correlation between time and lead usage over the 55 locations is low, 0.215.

We have no major criticisms of the statistical analyses in Section III, Subsection A2 of the Ethyl Corporation Comments. We would have done some things differently. The use of ~~36~~ 35 variables for personal characteristics essentially is equivalent to use of main effects and all order interactions of age, race, sex and income; we might have used six main effect variables and the related two-factor interactions, but sufficient data were available for the use of 3~~6~~5 variables. Similar comments might be made about location variables if locations were classified on a number of location characteristics. Instead of going to the second set of regressions for the 55 locations, the effects of time, gasoline lead use and population density could have been incorporated into the individual regression equation (1) or (2) and we do this in our analyses below.

Table 7 of [2] presents data that essentially removes biases due to sampling for standardized groups over time. Note that there is still a strong time trend for blood lead concentrations.

It is interesting that the small correlation between time and lead usage in gasoline over the 55 locations is positive. This suggests that exposure to lead in gasoline increased rather than decreased over the sequence of locations as sampled in time. This probably should be interpreted as an additional indication of bias for time periods in the sampling sequence of locations in the NHANES II study. It should be noted also that the Ethyl Corporation lead use in gasoline data are location specific and are not necessarily consistent with the national data exhibited by Houk.

ICF REVIEW OF ETHYL CORPORATION'S COMMENTS

The claims in the ICF review of Ethyl Corporation's Comments [2] as given in [3] are fairly devastating if they were valid. It is claimed that techniques used were invalid, the required software was not used, and the method is not capable of partitioning blood lead into personal and environmental contributions. We shall see below that the ICF claims are either incorrect or resulted from a misunderstanding of the Ethyl Corporation's efforts. This discussion again relates to the in-depth analysis of Section III, Subsection A2 of [2] and the related Appendix A.

Let us first address the claim that the most serious flaw in the Ethyl Corporation's approach is their failure to account for the complex survey design of NHANES II and to use required special statistical software in order to obtain unbiased variance estimates [3,pp.10,11]. We have commented on this issue already on pages 8, 9 and 12 of this report. The use of appropriate weights was clearly needed by Houk in presenting population estimates of blood lead levels with time. Proper weights are not available by time periods and not only will there be biases in estimates of variances but also in the estimates of blood lead levels themselves. The special programs or software noted surely can only provide proper weights for estimates of blood lead levels and their variances at the mid-point of the survey because of lack of balance of sampling over the four-year time period. Indeed, we would contend that the Ethyl Corporation's use of individual records is the one way in which association of an individual's characteristics and location can be modeled to predict individual blood lead levels. This does not, however, avoid all problems because time is still a possible factor in blood lead levels and the location variables also.

If it were possible to devise a planned, controlled experiment to directly examine the effect of lead in authmobile emissions on human blood lead, we would not need a sample of the U.S. population but rather homogeneous groups of individuals falling into specified personal characteristic groups subjected to controlled exposure to automobile emissions. Such an experiment seems impossible, but the Ethyl Corporation's examination of individual records with variable exposure because of gasoline lead use may be as close as one can come to the desired experiment. It is not necessary to have population estimates. It is only necessary to examine the relationship of blood lead to a measure of exposure after standardizing in so far as possible for other causes of variation.

The Ethyl Corporation's analysis attempts to standardize blood lead level after allowance for personal characteristics for each of 55 locations - hence Table 6 of [2]. It is perhaps unfortunate that they used the heading, "Blood-lead Contribution", leading to the detailed discussion in [3] on "The stated purpose of the first-stage treatment was to separate the contributions to blood lead into those factors relating to personal attributes and those relating to the environment in which the person lived." What was done, in spite of failures of communication, was to estimate the parameters $b_1,\ldots,b_{35}$ and $c_1,\ldots,c_{64}$ in model (1) and to predict y from the estimated model for each of the 55 locations for a "standard" individual with $x_1,\ldots,x_{35}$ having values zero. C_3 in [2, Table 6] is simply estimated blood lead obtained from the estimate of y, the logarithm of estimated blood lead. We take the use of zeros for $x_1,\ldots,x_{35}$ simply to mean that the "standard" individual was PI No. 36-old, black, female, high income. (See Table A-1 of [2]) In case this "standard" seems an unfortunate choice, note that the effect of any other choice would simply change the predicted values of y for each of the 55 locations by the same constant, thus not affecting the second stage of analyses.

Good

A lengthy discussion of the choice of variables in model (1) begins at the bottom of page 3 of [3]. It seems not to be recognized that values of $x_1,\ldots,x_{35}$ and $z_1,\ldots,z_{64}$ are simply indicator variables that perform a task equivalent to adjusting the constant b_0 in model (2) for each category of person and each location. It is a standard procedure to use discrete indicator variables in regression analysis, the most notable example being in the entire theory of analysis of variance.

The author of [3] makes a valid point in regard to the variables age and income. It seems more appropriate to the present writer to have used actual ages and income since they were available. Consequently, some loss in prediction probably resulted from dividing each of these two variables into three ordered classes, but this does not make the analyses wrong. Indeed, indicator variables for these classes were used in such a way that only the classes were defined and no assumptions about either spacings or even orders were made.

The paragraph beginning on page A-1 of [2] is unfortunate and somewhat misleading. It is this paragraph that led the author of [3] to note the stated purpose quoted above. It is this paragraph that led the author of [3] to claim that the complete model had 100 independent variables and embark on his discussion on pages 5-7 of [3].

which are irrelevent. Only at the top of page 8 of [3] is it recognized that one of the 36 personal identifier variables was dropped. It is then claimed that the whole procedure is invalid when in effect what happens is as described two paragraphs above. That explanation is completely consistent with the changes in regression coefficients of concern on pages 8, 9 and 10 of [3].

The present writer agrees with the author of [3] in his expressed anticipation that time be used in this first-stage regression analysis-see [3,p.2]. We do not see anything invalid in the use of time in the second-stage regression analysis of the Ethyl Corporation, but think that it would have been more effective to have continued all analyses on the basis of individual records rather than on Table 6 of [2]. We show such analyses in a section below.

In summary, the review of Ethyl Corporation's Comments by ICF Incorporated makes points that are either wrong or largely irrelevent. The claim that the analysis is not capable of partitioning blood leads into personal and environmental contributions is irrelevent, because it was a standardization process that was used. The claim that invalid techniques were used is wrong, nor were they used to arrive at arbitrary results. The claimed failure to use special statistical software required for the NHANES II data is irrelevent and it would have been wrong to use that software.

EXTENSIONS OF THE ETHYL CORPORATION ANALYSIS

We have noted above that we would have perferred to extend the initial regression analyses of the Ethyl Corporation to consider the effects of time, population density, and local lead use on individual blood lead levels rather than to go to the second set of regressions based on 55 locations. It appears that this would have been the preference of the author of [3] also.

Extensions and variations of model (2) are considered. First consider the following model with locations omitted but time, population density and local lead use added:

$$y = b_0 + \sum_{1}^{35} b_i x_i + dt + fp + g\ell + e \qquad (3)$$

where t is time, p is population density and ℓ is local lead use, d, f and g being the associated regression coefficients respectively. The analysis of variance in Table 1 below was obtained in four steps: (i) omit terms dt, fp and $g\ell$ of (3); (ii) omit terms fp and $g\ell$ of (3), include dt; (iii) omit term $g\ell$ of (3), include dt and fp; (iv) complete model (3). This method of analysis permits consideration of the significance of added model terms in correspondence with the additional variation in y accounted for by additional terms added to the model; the method seems preferable to consideration of individual regression coefficients and their significances as done in [2].

Care must be used in interpreting Table 1. The effect of local lead use in gasoline is clearly non-significant and variation in blood lead level after adjustment for time, personal identifiers, and population density cannot be explained in terms of local use of lead in gasoline, (possibly interpreted as lead exposure). This interpretation is as defensible as that of Houk in [1]. However, local lead use GASOLINE is also dependent on time and lead exposure may be related to population

Table 1

Analysis of Variance for Regression Model (3)

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F
Personal Identifiers	35	345.62	9.87	82.25 Very Sig.
Population Density (adjusted for Personal Identifiers)	1	9.43	9.43	78.58 Very Sig.
Time (adj. for Pers. Ident. and Pop. Dens.)	1	101.95	101.95	849.58 Very Sig.
Local Lead in Gasoline (adj. for Pers. Ident., Pop. Dens., Time)	1	0.07	0.07	0.56 Non-Sig.
Error	9363	1120.72	0.12	---
Total	9401	1577.79	---	---

density. This being the case, adjustment for time and population density may in a sense be adjustment for lead exposure and we do not contend that we have proved that lead use in gasoline does not affect blood lead levels. As stated at the beginning, we do not believe that the NHANES II data can be used to prove or disprove the alleged causal effect of lead use in gasoline on blood lead levels.

Our analysis could be improved through the use of model (4) with location variables added to model (3):

$$y = b_0 + \sum_{1}^{35} b_i x_i + \sum_{1}^{63} c_j z_j + dt + fp + g\ell + e \qquad (4)$$

The addition of the location variables, strongly correlated with time, will still have some effect of additional explanation of variation in y, reducing the error mean square towards 0.10, but we do not expect that there will be any effect on the non-significance of the F-ratio for lead in the corresponding analysis of variance table.

FURTHER ANALYSES BY ICF

ICF Incorporated [4] provide a further analysis of the NHANES II data and also of Billick's data* [10] in consideration of the relationship between gasoline lead emission and blood poisoning in America. Data used for gasoline lead emission were indeed based on monthly sales of gasoline, assummed related to lead emission. At this point, ICF is sensitive to criticisms made in regard to past analyses and lists a number of issues: valid sampling techniques, changes in subpopulation representation, lead in paint, lead in diet, and seasonality. It is interesting to note that the issue of time is not listed even though, in discussion of subpopulation representation, they state:

> "Therefore, any correlation between time and the fraction of the sample made up of blacks and urban dwellers, for instance, could account for some of the downward trend in overall mean blood lead levels."

Initial sections of [4] discuss the data used and attempt to dispell the issues noted. There is considerable discussion of the independent variables used in a regression model: age, sex, degree of urbanization, residence relative to an SMSA, national gasoline lead consumption, and certain interactions. We are told that the variable MALE was set at 1 for males and 0 for females, showing that ICF is not,

---

*Billick, Irwin H. (1982), "Prediction of Pediatric Blood Lead Levels from Gasoline Consumption," Office of Policy Development and Research, Department of Housing and Urban Development. (unpublished).

after their criticism of the Ethyl Corporation's Comments [2], indisposed to use indicator variables in regression models. It is interesting also that age is given in three categories although [2] was criticized by ICF for not using recorded ages in contrast to age categories by ICF in [3]. Consequently, ICF is inconsistent in two reports issued in consecutive months in 1982. Note that ICF is using a small subset of the regression variables and their interactions used in [3].

We are told almost nothing about the dependent variables used in [4] in the regressions nor are the data provided. We are told that they are NHANES II observations on blood lead levels. We are told that they were run using SAS procedure SURREGR on pages 4 and 5 of [4]. We are told on page 19 of [4] that ordinary least squares tests were used and that SURREGR is very expensive and difficult to use. On page 14, relative to analysis of the Billick data (not given), the response variable was $Y = \log [P/(1-P)]$, the logistic transformation, where P is the proportion, not percentage as stated, of "children with blood poisoning." It appears that this response variable was also used for the subpopulations of black children with the NHANES II data - see page 27 of [4]. We are not told what the response variable was for the main analyses of the NHANES II data; it appears to have been average blood lead level, although the logarithm of average blood lead level may have been more appropriate.

Let the response variable in the regressions be Y, where Y may be the logistic transform of P, average blood lead level or its logarithm. It appears that the response variable had to be measured (estimated from NHANES II data) at monthly intervals for each of 54 categories made up of age, sex, urbanization, and residence variables. Ethyl Corporation [2] avoided the need for unbiased subpopulation estimation through use of data on individuals; ICF requires unbiased estimates for 54 categories for each month! We have already discussed the biases in the NHANES II data with time and those criticisms apply much more severely here.

ICF responds very weakly to the issues that they noted. It is contended that valid sampling techniques were employed since NHANES II is their data source and that study was conducted with attention to accuracy and validity. There may have been attention to accuracy and validity for midsurvey estimates but there can be no doubt that it cannot provide accuracy and validity over time periods within the survey. To suggest accuracy and validity for subpopulations by month is absurd. The issue of subpopulation representation is covered as claimed only if the subpopulation estimates are unbiased with time. The issue of lead in the diet is dismissed on the contention that lead ingested from food cannot explain the magnitude of the change in mean blood lead levels in recent years. This contention is wrong - see our discussion of time trends in regard to the Houk Statement above. The statement on lead in paint is based only on an expectation and we cannot judge the efficacy of the argument.

Considerable attention to the issue of seasonality is given in [4]. Use of the NHANES II data to investigate the issue is rejected because

. "The effects of geography are, therefore, completely confounded in the data with any effects of seasonality."

The confounding of the NHANES II data with time in many other ways is ignored! The author of [4] turns to the Billick data to examine seasonality and does so in detail. He concludes that the hypothesis of no variation in the incidence of pediatric blood poisoning due to seasons can be accepted. Since he shows that gasoline lead consumption is a significent variable he concludes that the seasonal variation in the consumption of gasoline lead accounts for the seasonal variation in the incidence of pediatric blood poisoning. These conclusions depend on single regression functions and the significances of their estimated regression coefficients and careful analysis of variance techniques for regression as illustrated in the section above were not used. We do not have information on the quality of the Billick data. We are told in [4] that these data were gathered as part of various screening programs designed to detect lead poisoning due to exposure to leaded paint. We are told in [2] that the Billick Chicago data were collected over the 1967-80 period. We have no information on sampling techniques nor the defined population to be represented. It appears likely that the data are subject to biases with time again.

The new analyses of [4] may explain some variation in blood lead levels through the use of a limited number of independent variables, thus producing a smoothing effect. The new analyses do nothing to circumvent criticisms of the Houk presentation which apply equally well here. Basically, both blood lead levels and gasoline lead use are subject to trends with time and are consequently highly correlated.

The facts that such trends are not exhibited in [4], not discussed in any way, and hidden in the details of supposedly sophisticated regression analyses do not alter the conclusion that the NHANES II data cannot be used to infer a causal effect of gasoline lead use on blood lead levels.