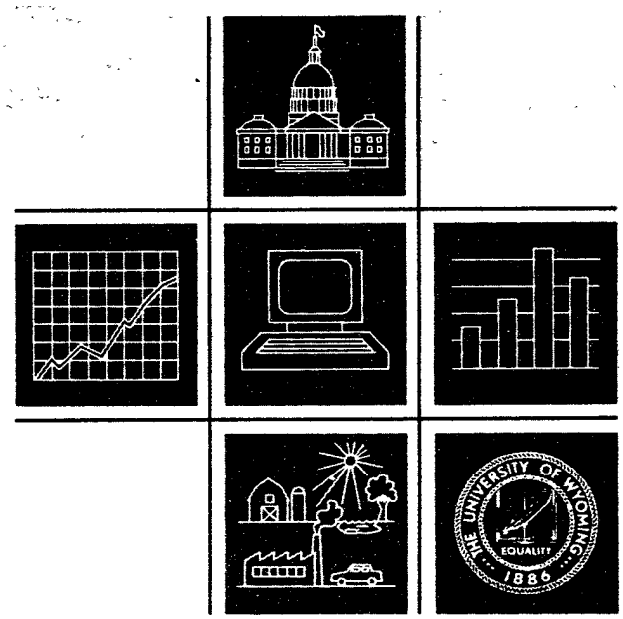
The Institute for Policy Research



. . . addressing public policy issues for the ‘80’s.

Volume VI

AN ECONOMIC ANALYSIS OF AIR POLLUTION AND HEALTH: THE CASE OF ST. LOUIS

METHODS DEVELOPMENT IN MEASURING BENEFITS OF ENVIRONMENTAL IMPROVEMENTS

Volume VI

AN ECONOMIC ANALYSIS OF AIR POLLUTION AND HEALTH: THE CASE OF ST. LOUIS

by

Shelby D. Gerking
Linda R. Stanley
William N. Weirick

Department of Economics and
Institute for Policy Research
University of Wyoming
Laramie, Wyoming 82071

USEPA Assistance Agreement CR808893010

Project Officer

Alan Carlin
Office of Policy Analysis
Office of Policy, Planning and Evaluation
U.S. Environmental Protection Agency
Washington, D.C. 20460

OFFICE OF POLICY ANALYSIS
OFFICE OF POLICY, PLANNING AND EVALUATION
U.S. ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

DISCLAIMER

This report has been reviewed by the Office of Policy Analysis, Office of Policy, Planning and Evaluation, U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the U.S. Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

CONTENTS

FIGURES

TABLES

ACKNOWLEDGMENT

We wish to thank Dr. Alan Carlin for his patience and assistance in aiding our efforts to complete this research project. Dr. Carlin was instrumental in providing us with the St. Louis health data, which we received in October, 1981, and made numerous helpful suggestions concerning how the data might be analyzed. Additionally, we deeply appreciate Dr. Carlin's efforts to aid us in securing the supplemental funding we received in September, 1982.

Also, we wish to thank without implicating, William Schulze, Michael Grossman, Paul Portney, Jon Harford, Michael Koontz, and August Auer for helpful comments and suggestions as well as Ema Shenefelt and Beth Daniels for expert clerical assistance.

S.D.G.
L.R.S.
W.N.W.

CHAPTER 1

INTRODUCTION

One of the principle tenets underlying the Clean Air Act and its subsequent amendments is that decreases in air pollution result in reduced human mortality and morbidity rates. Empirical work aimed at demonstrating the existence and strength of this dose-response relationship, however, have produced uneven results. Lave and Seskin (1973, 1977), for example, have reported regression relationships based on aggregate data that show a strong positive association between levels of sulfur dioxide and mortality rates. In contrast, Gerking and Schulze (1981) present an illustration, based on an alternative aggregate data set, indicating that when attention is paid to intervening variables such as medical care, the above conclusion may be reversed. This apparent conflict in results has a number of possible sources including: (1) lack of knowledge concerning the physiological process by which air pollutants affect health and the resulting uncertainties about the choice of covariates used to explain health status as well as the functional form of the relationship, (2) problems in obtaining accurate measurements on the desired variables, (3) multicollinearity between the covariates used, (4) the use of inappropriate statistical methods in estimating the relationships hypothesized, (5) an incomplete understanding of how individuals may alter their consumption patterns in order to defend against possible negative effects of air pollution.

The purpose of this report certainly is not so ambitious as to attempt to "solve" these five problems which have proved to be quite vexing in previous epidemiological investigations. Nevertheless, each of these five problems is addressed at least indirectly in the process of formulating and empirically testing a theoretical health model This model, which adopts an approach similar to that used by Grossman (1972), Cropper (1981), and Rosenzweig and Schultz (1982a, 1982b), and Harrington and Portney (1983) views individuals as producers of health capital in a utility maximizing framework. Hence, the individual is able to adjust his behavior so as to compensate for changes in air quality. Specific adjustments considered involve alterations in the consumption of medical care, exercise, and cigarettes as well as changes in dietary patterns and time spent working. The model also is used to derive a remarkably simple contingent valuation measure of the marginal willingness to pay for improvements in air quality.

The theoretical model is subjected to extensive empirical testing using survey data on adult workers drawn from households in the St. Louis SMSA. Those tests are of interest in one respect because the data used are of high quality, particularly in comparison with their counterparts used in other studies. The health, socioeconomic, and demographic measures are

very rich and were collected for the purpose of analyzing the relationship between air pollution and health, rather than for another primary purpose. Additionally, the air pollution measures obtained from the Regional Air Pollution Study (RAPS) are both more detailed and more spatially disaggregated than those available in alternative data sets. Most importantly, however, the empirical results are of interest because they support the hypothesis that reductions in air quality, measured principally as the concentration of ozone at 19 sites in the St. Louis SMSA, negatively affect health. Marginal benefit estimates for various percentage reductions in ozone concentrations are computed based upon these empirical results.

The remainder of this report is divided into five chapters. Chapter 2: (1) reviews the portion of the applied econometric health epidemiology literature that is based on individual observations (microdata) and is concerned with morbidity and (2) describes the methodology typically used to estimate the dollar benefits associated with improvements in air quality. Chapter 3, then, presents in detail the theoretical health model that was briefly described earlier and contrasts the approach taken with the methods discussed in Chapter 2. Chapter 3 also demonstrates how a contingent valuation measure of the marginal willingness to pay for improved air quality can be calculated. Chapter 4 begins the empirical portion of this report by highlighting the key features of both the St. Louis data set and the air quality data drawn from the RAPS and Chapter 5 compares the RAPS data with that from three alternative sources for the St. Louis SMSA in an effort to justify the use of the former. An important by-product of this analysis, is the surprising conclusion that air quality data drawn from RAPS stations frequently are linearly unrelated to corresponding measures obtained from physically adjacent stations in locally operated air monitoring networks. The material presented in the preceding four chapters then is used in Chapter 6 to empirically estimate: (1) the equations of the theoretical model and (2) the marginal willingness to pay for air quality improvements.

Finally, three appendices provide supplementary material. Since this report focuses on the willingness to pay to avoid exposure to ozone and lead, appendices 1 and 2 summarize the controlled experimental and epidemiological results on the human health effects of those pollutants. Appendix 3, reproduces selected portions of a volume by Koontz (1981) which provides documentation for the data base used in the empirical work in the present study. This appendix also explains how a computer tape containing these data may be obtained from the authors.

CHAPTER 2

LITERATURE REVIEW

1. Introduction

During the past several decades, economists have sought to estimate the economic costs associated with the adverse health effects of air pollution. The first step in this estimation has typically been to quantify the relationship between air pollution and human health. Human health is hypothesized to depend on several variables measuring socioeconomic and demographic factors, personal or lifestyle habits, as well as ambient concentrations of atmospheric pollutants and other environmental hazards. By specifying a physical damage or dose-response function describing the relationships between these various factors and human health, the association between air pollution and human health can, at least in principle, be isolated. From this relationship, the change in morbidity or mortality rates resulting from a given alteration in air pollution levels can be computed and an economic value attached. To arrive at this economic value, some measure of direct and indirect costs of illnesses usually is utilized in morbidity studies or, for mortality studies, a measure of the value of life.

This chapter contains a brief and highly selective review of the literature dealing with the benefits to human health resulting from improvements in air quality. Emphasis is placed on empirical studies that use microdata sets to estimate the dollar benefits of improved air quality that arise from reduced morbidity. There are two reasons why this very narrow focus is justifiable in the present setting. First, the purpose in undertaking this review is to assess the methodology and results obtained from data similar to that used in the present study. The St. Louis data set contains cross&sectional measures of the health status of living individuals; there is no information given regarding patterns of mortality. Second, broader surveys of the econometric health epidemiology literature which recently have been conducted by Burness et al. (1983) and Atkinson (1983), are contained in other documents submitted pursuant to USEPA Assistance Agreement CR808893010. The remainder of this chapter is organized into two sections. Section 2 summarizes the approaches taken and results found in several selected studies and Section 3 reviews alternative strategies for computing the benefits of air quality improvements. A brief summary is given in Section 4.

2. The Air Pollution-Health Relationship

As indicated, this section summarizes and critically evaluates

selected microdata based studies of the air pollution-health relationship. These studies can be divided into two classifications, depending upon the definition of the health measure used. In the first category, which is comprised of the work by Jaksch and Stoevener (1974), Bhagia and Stoevener (1978), and Seskin (1977, 1979a, 1979b), health (or illness) is measured in terms of contact with the medical system. As a consequence, the results of these studies bear most directly on the question of how alterations in air quality affect short-term or acute illness. The second category of studies, made up of contributions by Crocker et al. (1979), Ostro and Anderson (1981), and Cropper (1981) also focus primarily on acute illness; but examine measures such as days absent from work. Crocker et al. also consider data on disabilities in an attempt to gain in sight into the relationship between air pollution and chronic illness.

A. Air Quality and Medical System Contacts

Jaksch and Stoevener (1974) attempted to quantify in monetary terms the effects of air pollution on the consumption of outpatient medical services. They hypothesized that air pollution can aggravate one's health resulting in increased consumption of medical services for certain respiratory and cardiovascular diseases and any other diseases that may be aggravated by air pollution. Using the Portland, Oregon SMSA as the study area, they obtained their data from several air pollution control agencies, the National Weather Service stations, and the Kaiser Foundation Health Plan.

Their first hypothesis was that deterioration in air quality can increase the consumption of medical services per outpatient contact with the medical system. The model designed to test this hypothesis was specified as follows:

$$I_{ijk} = h(A_{jk}, W_{jk}, S_{ijk}) \qquad (1)$$

where I_{ijk} is an index of the dollar value of outpatient medical services consumed for treatment of the ith disease on the jth day for the kth person, A_{jk} represents a measure of air quality (expressed as suspended particulates, micrograms per cubic meter per day) on the jth day for the kth person, W_{jk} is a measurement of meteorological conditions on day j for the kth person, and S_{ijk} represents the socioeconomic-demographic characteristics of the kth person on the jth day of exposure for the ith disease. The index of medical services, though in monetary terms, was designed to reflect the quantity of medical services and not just dollar expenditures which would be influenced by variations in fee charges. Moreover, besides controlling for age, sex, race, and income, the socioeconomic-demographic variables also included measures of smoking, occupational factors, and physical fitness.

The second hypothesis was that deterioration in air quality increases the number of contacts with the medical system per disease category. The model used to test this hypothesis was specified as follows:

$$Y_{icj} = g(A_{cj}, W_{cj}, S_{cj}) \qquad (2)$$

where Y_{icj} is computed as the ratio

$$Y_{icj} = \Sigma_k^K y_{icjk}/N_{cj}, \quad (3)$$

$\Sigma_k^K Y_{icjk}$ is the summation of the index of consumed outpatient medical services for disease i for all K Kaiser health plan members in census tract c on day j, converted to dollars and N_{cj} is total number of Kaiser respondents residing in census tract c on day j. Also, A_{cj} represents an average measure of ambient air pollution (again expressed as suspended particulates, micrograms per cubic meter) in census tract c on day j, W_{cj} represents an average measure of meteorological conditions in c on day j, and S_{cj} is an average of socioeconomic-demographic characteristics of Kaiser health plan members in c who sought outpatient medical services on day j.

The air pollution exposure data used in this study were obtained from readings on ambient concentrations of suspended particulates available from 18 monitoring stations in the Portland, Oregon area. Although several of these stations were not used in the empirical work performed, location specific exposure levels still could be assigned to individuals, since for each respondent the places of residence and employment were known.

Various forms of these models were estimated using ordinary least squares regression methods. With respect to the first hypothesis, a deterioration in air quality appeared to result in the increased consumption of medical services per outpatient contact for respiratory diseases, but not for circulatory-respiratory diseases. With respect to the second hypothesis, however, variations in air quality did not appear to affect the number of contacts with the medical system per disease category. These results led Jaksch and Stoevener to conclude that economic costs of deteriorating air quality, measured as the frequency and intensity of contacts with the medical system, are positive but not particularly large. Another finding of interest was that there appears to be a time delay between exposure to comparatively high levels of air pollution and contact with the medical system.

Bhagia and Stoevener (1978) followed the Jaksch and Stoevener formulation in a parallel study of the use of medical services in the Portland area. This model was specified as follows:

$$M_{ijk} = f(A_{jk}, W_{jk}, S_k) \quad (4)$$

where M_{ijk} is a dollar index of inpatient medical services (as compared with outpatient services studied by Jaksch and Stoevener) consumed for treatment of the ith disease episode on the jth day for the kth person, A_{jk} is a measure of air quality on the jth day for the kth person, W_{jk} is a measure of meteorological conditions on day j for the kth person, and S_k represents socioeconomic characteristics of the kth person.

Included in the socioeconomic variables were smoking habits, consumption of alcoholic beverages, education, and number of hospital visits in the last three years. Suspended particulates was chosen as the air pollution measure and readings were assigned on the basis of the proximity of monitoring stations to the home and employment locations of respondents. An index of medical costs associated with M_{ijk} was then regressed on all the variables discussed above using least-squares metho None of the coefficients on air pollution were statistically significant. Age of the patient, family income, drinking habits, and the number of visits by the patient were significant variables in all the regressions run. Bhagia and Stoevener concluded that given the specification of the model, no relationship between suspended particulates and the consumption of inpatient medical services existed. They feel that this lack of apparent influence of air pollution on health may be due to a distributed lag relationship of greater length than could be analyzed effectively in their data set.

Seskin (1977, 1979a, 1979b) also examined data that were similar to those used by Jaksch and Stoevener (1974). In particular, Seskin drew a sample from members of the Group Health Association in the Washington, D.C. metropolitan area in an effort to uncover a connection between their use of outpatient medical services and day to day variations in air quality levels. Although this study began with data on individual health plan members, all of the empirical work reported focused on aggregates of those observational units. The time series model examined was:

$$V_i = V_i(AP, W, D) \qquad (5)$$

where V_i denotes the number of unscheduled visits per day over the period 1973-74 to department i in one of the Association's medical centers; i = urgent visit (from which patients often are referred to another department), internal medicine, pediatric, optometry and opthalmology; AP denotes air quality, W denotes meteorological conditions, and D denotes dummy variables for Saturday and Sunday. Air quality was measured as the maximum one hour average oxidant reading taken from one station in downtown Washington, D.C. for each day in the time period considered. Because Seskin examined only the time series variation in the data and ignored the cross-sectional variation, personal, socioeconomic, and demographic characteristics of the study population were thought to be unimportant.

The only air pollution-morbidity incidence relationship which Seskin found to be significant in both years 1973 and 1974 was that between daily unscheduled visits to the ophthalmology department and oxidant levels. The best results were obtained when the oxidant measure included in the equation was contemporaneous with the visit data. Lagged relationships between visits end air quality, episodic effects created by poor air quality lasting for several days, and synergistic effects between oxidants and NO_2, SO_2, and CO also were examined; but without much success in detecting consistently significant relationships. These results support the contention that apart from comparatively minor effects such as eye irritation, short-term or acute ill health is largely unrelated to the

mobile source emissions which are primarily responsible for ambient oxidant concentrations. However, the aggregate nature of the observational units used easily could have masked potentially important relationships between air quality and unscheduled visits by special populations of individuals such as asthmatics or the elderly. The use of readings from only one monitoring station to represent oxidant exposure levels for all sample members in the Washington, D.C. area may also have been a problem; although as Seskin indicated, levels reported from other stations did not differ greatly from those used.

B. Air Pollution, Work Loss Days, and Disability

Crocker et al. (1979) used a detailed, highly disaggregated data set gathered from interviews conducted by the University of Michigan Survey Research Center's Panel Study in Income Dynamics (PSID) from 1968 through 1976 to analyze acute and chronic illness attributed to air pollution. The independent variables used in their model can be divided into biological and social endowment variables (age, education, income, race, etc.), life-style variables (exercise and smoking habits, food expenditures, etc.), pecuniary variables (medical insurance, savings, the wage rate, etc.) and environmental variables (air pollution, weather, occupational exposure). Two dependent variables were employed: (1) work days lost due to illness, and (2) whether the respondent had a disability limiting the type or amount of work that could be performed. The former variable was thought to measure incidence of acute illness and the latter was thought to measure chronic illness. Observations on each of these variables pertained to household heads; consequently, the focus of this study was on the health of adults.

The air quality data used in this study measured the concentration of five pollutants: (1) nitrogen dioxide, (2) ozone, (3) total oxidants, (4) total suspended particulate, and (5) sulfur dioxide. Annual geometric means as well as 30th and 90th percentile values for each pollutant were obtained. However, for the ozone and total oxidant data, the number of monitoring stations and the monitoring time intervals allowed for only minor variations in exposure levels between individuals. As a result, the empirical estimates reported do not consider these two pollutants.

A potentially serious difficulty with the data set arose because the air pollution data had to be matched with the measurements on the household heads for the PSID data. Only the county of residence for each of those individuals was known. For some counties, no air quality data were available, while for other counties data were available from two or more monitoring stations. To circumvent this problem, individuals who did not live in a county where air quality data were available were dropped from consideration, thus reducing the sample size by 40-50 percent. Individuals living in counties with multiple monitoring stations were assigned the readings from the single station that had operated for the greatest proportion of the period 1968-1976. That strategy, which probably resulted in a tendency to assign downtown pollution measures for all who lived in a county, may have represented an important source of measurement error since there frequently are substantial variations in air quality between

location, particularly in larger counties. Moreover, the current year air quality measures assigned may not be indicative of the long term exposures that individuals have faced.

Using this data set, Crocker et al. estimated several dose-response functions for both acute and chronic illness. All equations were estimated by ordinary least squares. Although the researchers took care to assure that only those variables that are outside a household's control or had been established prior to the period being considered were included as regressors, they admitted that some of the variables (cigarette consumption, exercise, and dietary habits) could be adapted quickly to changing circumstances. However, they argue that these habits are insensitive to changing circumstances and are likely to persist. In any event, most of the regressions run had statistically significant coefficients on the air pollution measures employed. More specifically, in each of the seven different unpartitioned samples used to estimate acute illness dose-response functions, significant and correctly signed air pollution coefficients were present.[1] Additionally, significant and correctly signed coefficients of air pollution occurred 9 9 of 12 regressions run where the chronic illness measure was used as the dependent variable. On the basis of these results, Crocker et al. concluded that air quality, in addition to other factors, play a significant role in determining patterns of both acute and chronic illness.[2]

Ostro and Anderson (1982) performed an analysis of data from the 1976 Health Interview Survey (HIS) conducted by the National Center for Health Statistics. Similar to the work of Crocker et al., these authors used days absent from work as the health measure to be explained; however, the sample of individuals considered was drawn more restrictively. Ostro and Anderson included only nonsmoking males in their data set in order to reduce the possibility that: (1) the negative health effects of air pollution would be confounded with those of cigarette smoking and (2) the work loss days observed may have occurred for non-health reasons. In any case, work loss days were hypothesized to be a function of various demographic and socioeconomic variables, the existence of chronic disease, climate conditions, measures of "urbanness," in addition to air quality. Two air quality variables actually were employed, total suspended particulates and sulfates. The estimated relationship was specified in linear form and estimated alternatively by ordinary least squares and limited dependent variable methods (Tobit and logit).

As the authors indicated, the results obtained from their estimated equations generally were consistent with a priori expectations. Air pollution, as measured by total suspended particulates, frequently was positively and significantly (at the 5 percent level) related to the incidence of work loss days (WLD). Interestingly, however, particulate readings did not appear to influence the number of days lost given that a WLD episode has occurred. The coefficient of the sulfate variable, on the other hand, never was positive or significantly related to WLD. This poor performance was blamed, at least partially, on measurement errors in the data. Unfortunately, Ostro and Anderson do not discuss in any depth the important issue of how air pollution readings were

assigned to the individual sample members. The assignment methodology, as has been indicated previously, also could have influenced the performance of the pollution variables.

The work of Cropper in explaining cross-sectional variations in the incidence of work loss days represents a refreshing change from the approaches taken in the five studies just reviewed. In those five studies, the main focus was on estimating damage or dose-response functions in an effort to identify the health-air pollution relationship. A key assumption underlying this approach is that individuals are unaware of the health effects of air pollution and, as a consequence, do not take them into account in making decisions (Lave, 1972, and Lave and Seskin, 1977). That view has been disputed by many health economists (see, for example, Grossman, 19721, 1972b, 1976) who instead have hypothesized that individuals do indeed take defensive action against ill health. More specifically, this action is assumed to be taken in a utility maximization framework in which diet, exercise, and medical services are but a subset of the available consumption goods. Cropper's model adopts this perspective; however, a detailed description of that model is not presented here for two reasons: (1) the main focus of this section is on summarizing empirical results and (2) a similar model is presented in Chapter 3.

In any case, Cropper's theoretical analysis leads to the following statistical formulation of the equation to be estimated as a Tobit model

$$\begin{aligned} \ln TL_{it} &= \text{undefined} \quad \text{if} \quad X_{it}^{T}\beta + U_{it} \leq 0 \\ \ln TL_{it} &= X_{it}^{T}\beta + U_{it} \quad \text{if} \quad X_{it}^{T}\beta + U_{it} > 0 \end{aligned} \tag{5}$$

where TL_{it} denotes time lost from work by the i^{th} individual in year t, X_{it}^{T} is a row vector of observations on individual i in year t measuring characteristics including factors affecting investment in health (education, marital status, presence of a chronic condition), the wage rate, and variables affecting the rate of decay in the health stock (stress, race, air pollution exposures), β is a vector of coefficients, and U_{it} is a random disturbance term. Cropper obtained health related data from the Michigan Survey Panel Study in Income Dynamics which were then matched with county level observations on air quality, measured as the natural logarithm of the annual geometric mean SO_2 reading. Thus, Cropper's data set bears similarity to that used in the Crocker et al. study.

Cropper reports three almost identically specified regression equations for the PSID interview years of 1970, 1974, and 1976. In each of those regressions, the sign of the coefficient of the air pollution variable was positive and barely significantly different from zero using conventional tests. The ratio of the estimated coefficients to their standard errors were in the range of 1.38 to 1.52. These results corroborate epidemiological studies in which exposure to sulfur dioxide has been linked to acute illness (which in turn would be linked to work loss days). However, additional pollutants were not explicitly

considered since collinearity between them produced insignificant coefficients. The SO_2 variable was therefore regarded as an index of all air pollution exposures found.

3. Benefit Estimation

As indicated previously, the main purpose of the six studies reviewed in the previous section was to quantify the dollar benefits of air quality improvements using the estimated air pollution health relationships. This section provides an overview of the procedures used in those studies, as well as in others, in order to assess the benefits or reduced morbidity. The overview presented again will be selective. Burness et al. provide a somewhat broader evaluation of benefit assessment methods which examines both morbidity and mortality.

The most common method of estimating the benefits of reduced morbidity involves computing the costs of air pollution induced illness. Jaksch and Stoevener and Bhagia and Stoevener, for example, attempted to estimate these costs directly by defining the dependent variables in their regression equations as indices of the dollar value of medical services consumed. In other words, their regression equations showed the extent to which medical expenses might be expected to fall in the face of an improvement in air quality. Nevertheless, neither Jaksch and Stoevener nor Bhagia and Stoevener made explicit benefit calculations.

Seskin, on the other hand, did report explicit benefit calculations in his study of the association between unscheduled visits to a medical care facility and ambient concentrations of photochemical oxidants. Because only visits due to an ophthalmologic reason were found to be significantly associated with the level of oxidants, only benefits pertaining to this relationship were calculated. From the estimated damage function, Seskin determined that a 10 percent reduction in the level of photochemical oxidants would bring a decrease of 1.1 percent in 1973 and 4.3 percent in 1974 of unscheduled ophthalmologic visits. He then determined that in order for Washington, D.C., oxidant levels to comply with the 1971-1978 national standard for these pollutants, oxidant levels would have to have been reduced 55.6 percent in 1973 and 42.9 percent in 1974. These estimates were then applied to the oxidant-opthalmology elasticities derived from the damage function. Assuming linearity, this calculation indicated that reductions in oxidant levels to meet national standards would produce a 6.1 percent reduction in unscheduled opthalmologic visits during 1973 and an 18.4 percent reduction during 1974. These figures represent approximately 136 unscheduled visits in 1973 and 367 in 1974.

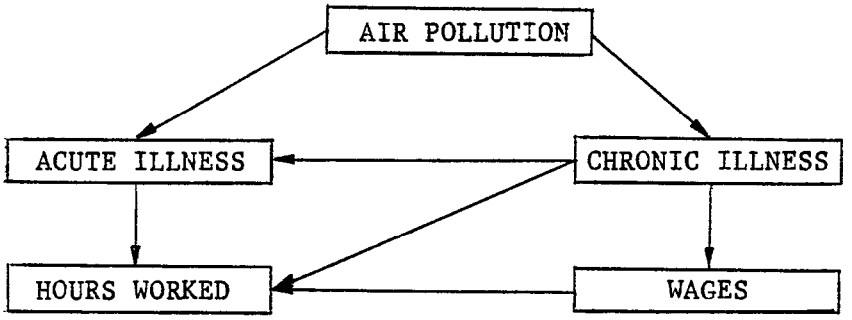
The next step was to attach monetary values to these visits. Using a value representing average medical costs associated with visits for simple eye problems in Maryland, the direct medical costs were calculated to be $20 per visit or approximately $2700 in 1973 and $7300 in 1974. Using data tabulated by Cooper and Rice (1976) which showed that indirect morbidity costs of diseases of the central nervous system and sense organs to be 66.3 percent as great as direct medical costs, Seskin calculated indirect costs

of these visits to be $1790 for 1973 and $4840 for 1974. Summing direct and indirect medical costs gives the benefits accruing to Group Health Association members of $4490 in 1973 and $12,140 in 1974. Assuming that those members are representative of the resident population of the Washington, D.C. area, total benefits can be calculated by multiplying this figure by the ratio that population bears to the number of Group Health Association members (20:1). Thus, the 1973 total benefit estimates would be $89,800 for 1973 and $242,800 in 1974.

Crocker et al. took an alternative approach to assessing the benefits of reduced morbidity associated with an improvement in air quality. They incorporated their previously discussed acute and chronic illness dose-response equations into a larger system that also determined an individual's wage and hours of work. Figure 1 provides a schematic illustration of the empirical model employed. As shown, increases in air pollution are hypothesized

Figure 1

Effect of Air Pollution on Labor Productivity



to increase both acute and chronic illness, and those illnesses, in turn, force reductions in hours worked as well as in the wage received, To implement this model, Crocker et al. estimated a system of four recursive equations using a sample drawn from the 1969, 1970 and 1971 University of Michigan Survey Research interview data. The four equations consisted of expressions for chronic illness, acute illness, the wage an individual receives, and the labor supply. From this system of equations, estimates of the direct and indirect effects upon the labor supply, measured in annual hours worked, of air pollution induced acute and/or chronic illness were calculated according to the following expression:

$$\frac{\Delta(\text{WORK HOURS}\cdot\text{WAGE})}{\Delta(\text{AIR POLLUTION})}=\frac{\Delta(\text{WORK HOURS})}{\Delta(\text{AIR POLLUTION})}\cdot\text{WAGE}+\frac{\Delta(\text{WAGE})}{\Delta(\text{AIR POLLUTION})}\cdot(\text{WORK HOURS}) \quad (6)$$

Using the equations estimated with the 1970 PSID data, they found that a 60 percent reduction in air pollution nationally would result in a per capita labor productivity gain of $288 (in 1978 $).

Unfortunately, neither the Seskin study nor the Crocker et al. study offers a persuasive theoretical rationale justifying the approach taken to compute estimated benefits. As indicated above, Seskin focuses on the savings on medical expenses resulting from reductions in air pollution, while Crocker et al. examine increased labor productivity. These measures, however, may not accurately reflect what individuals are willing to pay for improved air quality. In particular, in the absence of an explicit choice model, it is difficult to see: (1) why individuals would choose to value air quality changes in terms of either of these media,[2] or (2) how the benefits computed relate to an explicit welfare measure. In regard to this second point, a widely used and intuitively appealing welfare measure involves calculating the maximum amount that an individual would be willing to pay for an improvement in air quality subject to the constraint that his utility level must remain unchanged. While other measurements of welfare change certainly could be used, the studies cited make little effort to determine the extent to which an individual ends up better off when air quality improves.

In fact, Cropper is apparently the only investigator to have calculated the willingness to pay for improved health using methods derived from a theoretical model of consumer choice. However, she adopts the assumption that health is a pure investment good. That is, good health produces no direct utility; it only serves to reduce time lost from work, thereby augmenting the capacity to earn income. In this case, the decision to invest in health can be separated from the decision to produce other goods and a health investment point is chosen so as to maximize the present value of full income (R) net of the cost of investment. Utility then is maximized subject to a given level of R. Because of the separability between the decisions to invest in health and to consume other goods, and since air quality (A) has no direct effect on the individual's utility level, the maximum willingness to pay for a small percentage rise in A is $(dR/dA) \cdot (A)$. In other words, this framework predicts that an individual would, at most, be willing to give up his net income increase occasioned by reduced sick time when air quality improves. This result would appear to lend at least partial support to the general approaches used by Crocker et al. and also Ostro and Anderson where the income foregone due to air pollution induced illness was of primary concern.

4. Summary

This chapter has directed attention to a key problem in the econometric air pollution-health epidemiology literature. In five of the six studies reviewed, neither the health expressions estimated nor the benefit calculations made were based on an explicit theoretical model. As a consequence, although the findings reported are interesting, they are difficult to apply in a policy setting. In fact, the lack of a sound theoretical approach may have at least partially led to the wide variations in estimated impact of air pollution on health noted by Burness et al. Additionally, the benefit estimates reported by Seskin and Crocker et al. do not correspond to a precise measure of welfare change. The sixth study, conducted by Cropper represents a substantial improvement over this

situation, as her benefit estimates were constructed on the basis of a theoretical health model first proposed by Grossman (1972). Her model justifies the use of net income changes induced by the positive effect of air quality on health as a benefit measure. Thus, her approach also lends some support to the Crocker et al. and Ostro and Anderson studies, which focused on the work loss days as a measure of the impacts of air pollution.

The chapter to follow generalizes a version of the Cropper model in order to treat health as a consumption good as well as an investment good. That is, in the model presented, individuals are assumed to derive direct utility from good health as well as increased levels of income. As demonstrated there, that alteration in the model also implies a change in the method used to compute willingness to pay for air quality.

FOOTNOTES

1. Other regressions were run where the sample was partitioned according to income, smoking habits, and severity of disability.

2. Crocker et al. does present an explicit model which has similarities to the work of Grossman (1972); however, that model was not employed in the benefit estimation exercise.

CHAPTER 3

A SIMPLE HEALTH MODEL

1. Introduction

The model to be developed in this chapter, as indicated previously, represents a generalization in one respect of the approach taken by Cropper (1981). Health is treated as a consumption good from which an individual derives direct utility as well as an investment good which contributes to increased income. Decision-making, however, is examined by focusing on only one time period rather than on the multiperiod framework used by Cropper (1981) and Grossman (1972a, 1972b, 1976). The model, which is presented in Section 2, has strong parallels with the work of Rosenzweig and Schultz (1982a, 1982b) and Harrington and Portney (1983). The purpose of the Rosenzweig and Schultz papers was to analyze how the behavior of expectant mothers affects the birthweight of their children. The objective of this chapter, on the other hand, which is similar to that purused by Harrington and Portney, is to determine how to calculate the maximum dollar amount that an adult worker would bid in order to obtain an improvement in air quality levels. Methodology for computing that bid is considered in Section 3.

2. The Model

In the model to be applied, as shown in equation (1), individuals derive utility from the consumption of two classes of goods: (1) their own stock of health capital (H) and (2) goods that yield direct satisfaction, but do not affect health (X).

$$U = U(X, H) \tag{1}$$

Thus; apart from its one period focus, the utility function in equation (1) is quite similar to that used by Grossman in his treatments of the consumption aspects of health. In contrast, Rosenzweig and Schultz use a somewhat more complete specification of the utility function by including a good, Y, which yields direct satisfaction and also affects health. As a consequence, the model proposed by those authors is able to explain behavior such as smoking, dietary patterns, alcohol consumption, and exercise. The added richness resulting from incorporating the Y good is not pursued here, however, since the expression giving the willingness to pay for improved air quality would be unaffected. Two other possible refinements in the model would be to: (1) replace the health stock variable in the utility function with the flow of services it generates and (2) allow for home production of the

X-commodity. Both of these features are reflected in the Grossman (1972a) modeland the first is reflects in the Herrington and Portney analysis. However, they are not adopted here since no substantive alteration in the analysis would result.

The health stock is treated in this model as an endogenous variable, whose value is determined by the production function

$$H = H(M; \alpha) \tag{2}$$

where M denotes medical care (from which the individual derives no direct utility),, α denotes a set of exogenous variables, such as air quality, that affect the efficiency with which an individual can produce H using a given level of M, $H_M > 0$, and $H_\alpha > 0$.

Utility then is maximized subject to equation (2) as well as the money and time constraints shown in equations (3), (4), and (5)

$$XP_X + MP_M = I \tag{3}$$

$$XT_X + MT_M + T_W + T_L = T \tag{4}$$

$$WT_W = I \tag{5}$$

In the above equations P_i denotes the money price of commodity i (i = X, M), I denotes money income, T_i denotes the time required to consume one unit of commodity i (i = X, M), T_W denotes time spent working , and T_L denotes time lost from market and non-market activities due to illness. T_L, in turn, is related to the health stock according to

$$T_L = G(H) \tag{6}$$

where $G_H < 0$ and $G_{HH} > 0$ reflecting the assumption that an improvement in health reduces time lost from market and non-market activities, but at a decreasing rate. Equations (3), (4), (5), and (6) easily can be combined into the "full income" budget constraint shown in equation (7)

$$Xq_X + Mq_M + WG(H) = WT \tag{7}$$

where $q_i = (P_i + WT_i)$, i = X, M.

The approach adopted in this model stands in contrast to the dose-response or damage function approach that is widely used in the econometric health epidemiology literature. In the model described above, the consumer determines the amounts of goods and services to consume by examining all relevant exogenous information including full prices (the q_i), the wage rate paid for his labor services, and the air quality levels faced. In other words, health is treated as just one good in the model. Moreover, the health stock is produced using another consumption good as an input, namely M. As a consequence, the model illustrates that by regressing measures of morbidity on variables such as the consumption of medical care in addition to air quality levels,

previous investigators may have been regressing one choice variable on others.[1] Since single equation estination methods have been used almost exclusively, simultaneous equation bias easily could have been a significant problem in measuring the effect of air quality changes on health. In summary, therefore, any estimate of the health production function should account for the fact that it is embedded in a larger set of optimization equations.

The model just specified can be analyzed further by examining: (1) the first order conditions for a constrained utility maximum and (2) certain comparative statics results. The choice problem confronting the individual is to maximize equation (1) subject to equations (2) and (7). Formally, this problem can be re-expressed by substituting the health production function, equation (2) into both (1) and (7) and then writing the Lagrangian

$$L = U(X, H(M; \alpha)) - \lambda(Xq_X + Mq_M + WG(H(M; \alpha)) - WT) \quad (8)$$

First order conditions for a maximum are

$$\frac{\partial L}{\partial X} = U_X - q_X\lambda = 0 \quad (9)$$

$$\frac{\partial L}{\partial M} = U_H H_M - \lambda(q_M + WG_H H_M) = 0 \quad (10)$$

$$\frac{\partial L}{\partial \lambda} = WT - Xq_X - Mq_M - WG(H(M; \alpha)) = 0 \quad (11)$$

As these equations show, the utility maximizing individual will equate the ratio of marginal utilities of the goods consumed to the corresponding ratio of prices. An interesting aspect of this calculation, as noted by Rosenzweig and Schultz (1982a), is that the individual derives utility from the consumption medical care, M, only indirectly through its effect on health. Additionally, the price term shown in equations (10) account for not only the full, time inclusive, costs of M; but also the effect its consumption has on the time available for market and non-market activities $(T - T_L)$. As shown, an incremental change in available "healthy" time is valued at the individual's wage rate.

The comparative static properties of the model also are of interest. More specifically, the effects of an alteration in air quality on the. consumption of X and M can be expressed as weighted averages of pure price and income changes as shown in equations (12) and (13)

$$\partial X/\partial \alpha = C_X(\partial X/\partial q_X) + C_M(\partial X/\partial q_M) - G_H H_\alpha(\partial X/\partial T) \quad (12)$$

$$\partial M/\partial \alpha = C_X(\partial M/\partial q_X) + C_M(\partial M/\partial q_M) - G_H H_\alpha(\partial M/\partial T) \quad (13)$$

In these equations

$$C_X = -\lambda^{-1} U_{XH} H_\alpha \quad (14)$$

$$C_M = -\lambda^{-1}(U_{HH} - \lambda WG_{HH})H_M H_\alpha \qquad (15)$$

use is made of the simplifying assumption that $H_{M\alpha} = 0$, and the price effects shown reflect pure substitution in the sense that they ignore any associated income effects. Furthermore, if $U_{XH} > 0$, then $C_X < 0$ and if $U_{HH} < 0$, then $C_M > 0$. Therefore: (1) since $\partial X/\partial q_X < 0$, $\partial X/\partial q_M > 0$, and if $\partial M/\partial T > 0$, an improvement in air quality leads unambiguously to an increase in the consumption of the X-good and (2) since $\partial M/\partial q > 0$, $\partial M/\partial q < 0$, and if $\partial M/\partial T > 0$, that same improvement in air quality is "likely" to lead to a decrease in consumption of medical services. In other words, the consumption of medical services would fall if the pure substitution effects reflected in the first two terms of equation (13) are larger in absolute value than the pure income effect captured by the third term.

Intuitively, then, an improvement in air quality has two effects. First, ignoring income effects, it operates in exactly the same way as a wedge driven between q_M and q_X. That is, an increase in a is equivalent to an increase in the ratio q_M/q_X which would encourage consumption of X and discourage consumption of M. Second, the improvement in air quality has an income effect which results from the associated increase in time available for market and non-market activities. That positive income effect would encourage the consumption of both goods. Hence, consumption of X must rise and consumption of M falls if pure substitution effects dominate the income effect.

This situation, where changes in air quality have both substitution and income effects, contrasts with the implications of the model specified by Cropper. As indicated in Chapter 2, the only effect of health on utility in the Cropper model is through the budget constraint. Hence, the decision to invest in health can be separated from the decisions to purchase other goods. This distinction has a direct bearing on the appropriate method for estimating the willingness to pay for air quality. In the Cropper model, the willingness to pay estimate was derived via a simple manipulation of the budget constraint, whereas in the model presented here, the entire equation set must be utilized.

3. Calculating the Marginal Willingness to Pay

In this section, the marginal willingness to pay for air quality is derived from the model previously presented using the method of compensating variations (CV). This approach to calculating the marginal willingness to pay is chosen because it explicitly holds the individual's utility level constant in determining how much money an individual would give up in order to consume at a new price set. The method of equivalent variation (EV), which asks what an individual would willingly give up in order to forego consuming at the new price set, also would hold utility constant. However, since there may be only a minor difference in the numerical values of the bids produced by the two methods, the choice between them may not be important (Freeman, 1979). Moreover, note that either method would be appropriate because, as demonstrated in Section 2, the comparative static changes in X, and M given an alteration in air

quality all can be re-expressed as weighted averages of pure substitution effects and income changes.

Before deriving the expression for the CV marginal willingness to pay for air quality, a minor adjustment must be made to the model. In particular, the budget constraint is respecified as

$$Xq_X + Mq_M + B - W(T - G(H)) = 0 \tag{16}$$

where B denotes the bid, or amount paid, for air quality. Equation (16), then, requires that the individual actually must spend part of his full income, WT, to obtain an improvement in air quality. As a consequence, the individual's maximum willingness to pay for an incremental change in air quality would be found by computing $dB/d\alpha$ while: (1) holding utility constant and (2) ensuring that the individual's equilibrium conditions are obeyed.

One way to find a suitable expression for $dB/d\alpha$ is to totally differentiate both the utility function (equation (1)) and the health production function (equation (2)), substitute the latter into the former, and set dU = 0, in order to obtain

$$0 = U_X dX + U_H H_M dM + U_H H_\alpha d\alpha \tag{17}$$

Next, totally differentiate the budget constraint, as shown in equation (18), holding $dq_i = dW = dT = 0$ for i = X and M.

$$0 = q_X dX + (q_M + WG_M H_M)dM + dB + WG_H H_\alpha d\alpha \tag{18}$$

Equation (17) can be solved for dX, rewritten using the first order equations as

$$dX = -((q_M + WG_M H_M)/q_X)dM - (U_H H_\alpha / U_X)d\alpha \tag{19}$$

and then substituted into equation (18). After rearranging terms, those operations yield:

$$dB/d\alpha = [(U_H q_X/U_X) - WG_H]H_\alpha \tag{20}$$

which can be further simplified using the first order equations (9) and (10) as shown in equation (21)

$$dB/d\alpha = H_\alpha q_M / H_M \tag{21}$$

Note that this expression for estimating the marginal willingness to pay for air quality is relatively straightforward to implement empirically, since the utility terms have been eliminated. Additionally, this expression for $dB/d\alpha$ would be unaltered if the Y good were appropriately introduced into the utility function, the health production function, and the budget restraint. Hence, that expression is robust in the face of at least some alterations in model specification.

Equation (21) suggests that the individual would be willing to pay more for a given air quality improvement, the greater the associated improvement in health. Also, the bid would be higher, the lower the productivity of medical services and the higher their cost. As a consequence, if medical services are an expensive and ineffective means of producing good health, then quite naturally, the individual would be willing to pay more for air quality improvements. In that situation, air quality improvements simply become a more attractive mechanism through which to augment the health stock.

The marginal air quality bid shown in equation (21) easily can be compared to the bid that would result if health was treated as a pure investment commodity; i.e., if H was eliminated from the utility function. In that event, the first term in brackets in equation (20) would equal zero and this alternate marginal bid $dB'/d\alpha$ would be calculated according to

$$dB'/d\alpha = - WG_H H_\alpha \tag{22}$$

which simply values the reduction in time lost from market and non-market activities caused by the improvement in air quality at the wage rate. This bid is remarkably similar to the damage function approach used by Lave and Seskin (1977) to estimate the gains from reducing air pollution induced illness. Those authors obtained their benefit estimates by adding the value of the decrease in lost work time to the decrease in required medical services. The bid in equation (22) ignores any changes in medical expenses; however for minor illnesses these are likely to be small. In that case, therefore, the CV benefit estimate (based on the situation where health is treated as a pure investment good) and its counterpart using the damage function approach would be very similar. Note, however, that the willingness to pay estimate from equation (21) exceeds the one from equation (22) by the amount shown in equation (23)

$$(dB/d\alpha) - (dB'/d\alpha) = (q_M + WG_H H_M)(H_\alpha/H_M) > 0 \tag{23}$$

which is positive in light of the first order equation (10). That result should be expected since the model analyzed in this chapter treats health as a commodity with both consumption and investment attributes.

4. Conclusions

This chapter has developed a health oriented choice theoretic framework for the purpose of determining an individual's marginal willingness to pay for improved air quality. The willingness to pay expression, which was derived using the compensating variation approach, is quite simple in that it involves only one price (that of medical care) and two partial derivatives from the health production function (those for air pollution and medical care). Moreover, the willingness to pay expression does not involve any utility terms so that empirical estimation of it appears to be relatively straightforward. Chapter 4,

which follows, begins the empirical portion of this report by describing the St. Louis health and air quality data used in making the willingness to pay estimates. Those estimates themselves are presented in Chapter 6.

FOOTNOTES

1. If the Y good were explicitly included, the model also would illustrate the general problem of regressing a morbidity measure on variables reflecting cigarette and alcohol consumption, exercise patterns, or dietary habits.

2. Another method of deriving $dB/d\alpha$ would make use of an indirect utility function. That procedure is used by Harrington and Portney to obtain a result that is similar to the one presented here. In fact, the two results would be identical if Harrington and Portney had defined their D variable as units of a good rather than the dollar amount of defensive expenditures. If D were defined in units, then the price per unit would appear in their equation (17) (compare with equation (21) below in this report).

CHAPTER 4

THE ST. LOUIS HEALTH AND AIR QUALITY DATA

1. Introduction

The preceding chapters developed a framework for the methodology to be used in the empirical estimation of benefits accruing to individuals from improvements in air quality. The remainder of this report deals with the application of this framework to the St. Louis pollution-morbidity data set, the collection of which was supported by the USEPA. This chapter is organized into three additional sections. Section 2 provides an overview of the St. Louis health data and Section 3 discusses the air quality data used. A brief conclusion is contained in Section 4.

2. The St. Louis Health Data Set

A. Background

In April, 1974, the Environmental Protection Agency (EPA) awarded a contract to Geomet, Inc. (now a subsidiary of Geomet Technologies, Inc. [GTI]) to perform a household interview survey of the St. Louis SMSA. The survey was designed to gather data relevant for quantifying morbidity costs. Furthermore, it was to be administered during the years 1975-1977 in an effort to take advantage of comprehensive air quality data gathered from a network of 25 monitoring stations throughout the St. Louis SMSA which were operated in connection with the Regional Air Pollution Study (RAP). However, final clearance of the questionnaire was not obtained from the Office of Management and Budget (OMB) until January, 1978. By this time, the RAPS network was no longer in operation.

In any case, the questionnaire was administered over the period June 1978-July 1979. Households were enrolled in the survey on a weekly basis at the rate of 80 per week for the 52 weeks. The households selected for the survey were systematically allocated over the 52 weeks so that at any point in time the study sample would still represent the whole. Each household was asked to participate in the survey for eight weeks. A background interview was administered at the beginning of the eight weeks followed by four biweekly follow-up interviews. Later, a follow-up interview designed to collect supplemental data was conducted to complete the data base. Of the 4160 solicited households, 3063 or 73.63 percent participated in the basic survey to completion. Eighty-five percent of those households provided supplemental information. This represents 62.36 percent of the 4160 households originally selected to participate.

Three forms originally were utilized to gather health, socioeconomic, and demographic data from each household: a "Household Background Questionnaire," and "Individual Background Interview," and a "Telephone Interview Form." The first two instruments were to be completed during the initial interview with each participating household. The third instrument was administered during each of the four biweekly follow-up periods. The Household Background Questionnaire was designed to identify all members of the household and their basic demographic characteristics. The demographic characteristics included age, race, sex, and education. In addition, in order to ascertain the socioeconomic status of the household, a categorical question was asked regarding the income of the entire household. Information on the medical insurance of each member of the household also was gathered.

A major purpose of the Individual Background Interview was primarily to establish health status. Any pre-existing chronic conditions that were reported by the respondents were recorded according to a detailed code developed by Schneider, Appleton, and McLemrie (1979) in their paper entitled "A Reason for Visit Classification for Ambulatory Care." Other questions concerning the subjective appraisal of each individual's health status and the number of contacts with the medical care system also were posed in order to measure baseline health levels. The remainder of the Individual Background Interview dealt with identifying each individual's regular activity patterns. An individual's major activity and the occurrence of this activity were obtained to determine the typical location, days, and hours away from the home. Individuals who worked either part or full time were asked for information on the nature of their job. Also included in this survey were questions concerning: (1) commuting routines associated with the major activity and other various activities, (2) commuting and waiting times associated with doctor visits, and (3) smoking habits of each individual who was 14 years or older. The extent and duration of cigarette smoking as well as cigar and pipe smoking, was obtained for each individual who was smoking at the time of the survey.

At the conclusion of both background interviews the respondents were asked to participate in the eight week follow-up survey. Respondents who agreed to take part in this survey provided daily diary type information on how their health or other events altered their normal routine. These data, provided via the Telephone Interview Form, were intended for use in determining the relationship between air pollution and acute illness. The information obtained dealt with the following health related areas:

1) Activity restrictions

   Any activity restriction was categorized according to its degree of restriction, i.e., days confined to a hospital, to a nursing home, to bed, to the home or simply a reduction in usual activities.

2) Absenteeism

   The days of work and/or school absenteeism were recorded for each follow-up period.

3) Physician contacts

Physician contacts were categorized according to the means by which the contact took place, i.e., a visit to the doctor's office, a visit to the outpatient service in a hospital, etc.

4) Receipt of Ancillary Services

This question concerned whether shots, x-rays, or lab tests were required when a physician's care was sought. Prescriptions filled also were recorded.

The reasons associated with these health events were categorized and recorded in the same manner for the chronic illness data. Also provided by the respondents were changes in the normal routines of each individual. Included was information on changes in job status, days out of town, and days at home for vacation from work or school.

The EPA decided after the basic survey was completed to collect supplemental data to enhance the analytic potential of the basic data base. These data were collected between April and August 1980 from only those households completing both the background interview and the four biweekly telephone interviews. This supplemental information included more detailed data concerning an individual's personal habits, work habits, and workplace characteristics. More specifically, questions concerned:

1) The place of residence

- Type of heating/cooking fuel used
- Presence of air conditioning
- Residential history

2) The place of work

- Length of employment, salary, exposure to irritants, membership in a labor union, vacation/sick leave availability for individuals who were working either full-time or part-time at the time of the survey
- Occupation, nature of the job, length of employment, exposure to irritants for retired individuals

3) Income

- Proportion and source of income not related to jobs

4) General Health

- Height, weight, frequency of dieting
- Exercise patterns, time spent outdoors/watching TV
- Length of long term health problems
- Historical tobacco use

- Consumption of cured meats, sweets, salty snacks, and caffeine and the number of hot meals eaten per day
- Alcohol consumption (including frequency and type of alcohol consumed).

Data from both the basic survey and the supplemental survey were edited, coded, and merged into a single data file by GTI. The complete documentation for it is contained in Koontz (1981).

B. Survey Design

The sampling frame from which the 4160 households were drawn was taken from the 238 census tracts within the urbanized area of the St. Louis SMSA which were bounded by a network of continuous air quality monitoring stations. These 238 census tracts contained 432,162 households, or approximately 60 percent of those within the entire St. Louis SMSA. The 4160 households surveyed, then, represents somewhat less than 1 percent of the total in the sampling frame.

A stratified random sampling procedure was used to determine the exact choice of individual households to include in the survey. More specifically, the probability-proportional to size (PPS) technique was used in order to assure equal selection probabilities for all households. Equal selection probabilities, guarantee that the observations do not require weighting in any subsequent analysis. Counts of dwelling units by census tracts in 1970 were used as the basis for implementing the PPS technique. Obviously, due primarily to new home construction the actual dwelling unit count for 1978 would have differed, perhaps substantially from its 1970 counterpart. Consequently, the probabilities associated with each households inclusion in the sample are only approximately equal.

There are two additional problems with the sample to which attention also should be drawn. First, no specific attempts apparently were made to include observations from the institutionalized subpopulation. In other words, because the sampling design was based on dwelling unit counts, individuals subject to long term health care in nursing homes or hospitals may not have been adequately represented. To illustrate the impact of this problem, incidence of chronic conditions in the population of the United States can be compared to incidence in the original sample. Statistics compiled by the U.S. Department of Health, Education, and Welfare, National Center for Health Statistics (1973) indicate prevalence of chronic bronchitis, emphysema, and asthma (with or without hay fever) over the entire population which can be compared directly to the sample. Table 1 shows this comparison. As is evident, chronic illness incidence in the St. Louis health survey appears to be biased downward relative to national averages.

Second, of the 4160 households in the original sample, 1097 (26.4 percent) did not, for various reasons, either complete the Individual Background Interview or provide the diary data requested or both. For these individuals, only a portion of the information requested is

TABLE 1

INCIDENCE OF SELECTED DISEASES

	INCIDENCE PER 1,000 PERSONS	
	St. Louis Survey	National Survey
Emphysema	5.1	6.6
Bronchitis, chronic	3.8	32.7
Asthma	12.7	30.2

available. Also, for a household *to* be eligible for the supplemental or follow-up interview that household must have completed the initial phase. Of those eligible (3063), 84.7 percent then completed the follow-up interview; thus, there were 2594 households providing an entire data record. The major reasons for non-response or incomplete response during the initial portions of the survey were outright refusal to participate and relocation outside the St. Louis area. Corresponding primary reasons for non-response during the follow-up portion were outright refusal to participate and failure on the part of the interviewer to locate or contact.

Despite these difficulties, the sample information would appear to have three redeeming features. The first and most obvious is the care that was used in collecting the acute and chronic illness data. The illness classification scheme referred to earlier is very detailed; consequently users of the St. Louis data have a great deal of information concerning the respondent's chronic medical problems. Moreover, the diary type data reflecting acute illness is potentially useful particularly since they were recorded day-by-day rather than on the basis of recall by respondents. Second, data were collected for each respondent on a broad range of variables that are potentially useful predictions of illness patterns. The exact specifications for these data are described by Koontz; however, the availability of both current and historical measures concerning each respondent's income, education, smoking, exercise, exposures to on-the-job hazards and indoor air pollutants, leisure time activities, and diet is worth mentioning. Additional details concerning certain of these data are given in Chapter 6 of this report in which the empirical results are presented. Third, the data set also contains comprehensive information concerning the physical locations of the respondent's residence, employment, and other activities. Those data are of critical importance in matching air pollution exposures to the individuals surveyed, The air pollution exposure measurements are discussed in the Section 3 to follow and the matching procedure is described in Chapter 6.

3. St. Louis Air Quality Data

Air quality data are available from four sources for the St. Louis

area: (1) the EPA financed Regional Air Pollution Study (RAPS), (2) the county of St. Louis, (3) the city of St. Louis, and (4) the Illinois Environmental Protection Agency. As previously indicated, the RAPS had been concluded prior to the time the St. Louis health data were collected. However, the air quality data based upon daily averages obtained from RAPS appear to be of substantially better quality as compared with those from the other three monitoring systems. As a consequence, they are used exclusively in this study in spite of the unfortunate time sequencing problem. In fact, Chapter 5, which presents a detailed comparison of all four air monitoring systems, concludes that the three locally operated networks simply do not provide daily data that are accurate to a useful degree of approximation. This conclusion implies that health-air pollution relationships estimated by matching the daily diary data with daily air quality readings from one of the three local monitoring systems are to be viewed with considerable suspicion. In fact, no such relations are estimated in this report and the daily diary data are not used. However, if local readings were averaged over a longer time interval, say a year, their correspondence to those obtained from the RAPS system are considerably closer. Thus, the implicit condemnation of the quality of the readings from the three local monitoring networks applies only when those data are used in daily average form.

A possible extension of the analysis presented in Chapter 6 would involve making comparisons between the estimated health effects of air pollutants by holding the regression specification constant and varying only the source of the air quality data. In such an analysis, the air quality data would be obtained from long term averages of readings from stations in each of the four monitoring systems. Due to the expense of matching the pollution exposures to the individuals surveyed and to the fact that the results would provide only a sidelight to the present study, the extension is not pursued. In Chapter 6, only the RAPS data are utilized, and as a consequence, the remainder of this section provides an overview of the data produced by the RAPS together with a correlation analysis designed to highlight certain features of importance.

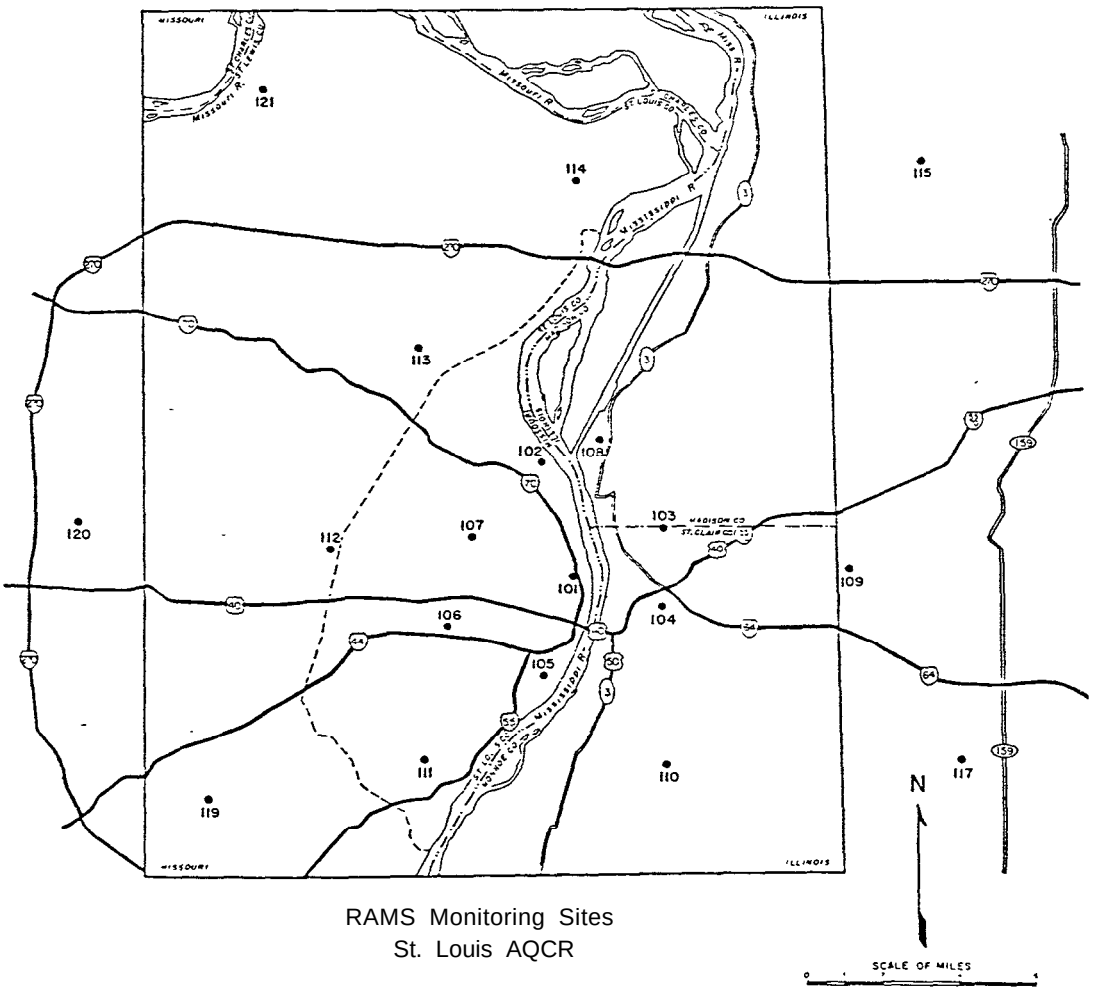
A. Background

Between the years 1974 and 1977, the Environmental Science Research Laboratory of the U.S. Environmental Protection Agency contracted with the Air Monitoring Center of Rockwell International to set up and operate an extensive monitoring system for meteorological and pollutant variables in the St. Louis, MO-IL air shed. This network, called the Regional Air Monitoring System (RAMS) was part of a larger Regional Air Pollution Study (RAPS) investigating many facets of the effects of emitted air pollutants on the scale of an entire Air Quality Control Region (AQCR). In addition to studies directly supplementing the RAMS surface network by helicopter and radiosonde measurements, RAPS included additional endeavors such as air-parcel trajectory studies extensive point source emission inventories, and analysis of micro-amounts of material in suspended particulates. The RAMS network was designed to provide state-of-the-art measurement and monitoring mechanisms.

The actual RAMS network consisted of 25 stations arranged in very roughly laid out concentric circles from a central urban station at radii 5, 11, 20, and 44 km (see Figure 1). The network was designed to generate an accurate and retrievable data base for all criteria and certain selected non-criteria pollutants to be used in air quality simulation models. Each station included pollutant analyzers, meteorological sensors, and test/control systems. Not every station collected data on all pollutants measured by the RAMS system. Table 2 lists the distribution of pollutant measurements by station. Table 3 displays the equipment used and the variables measured by the instruments.

One of the characteristics which distinguished the RAMS system from other monitoring networks existing in the St. Louis AQCR is the extensive attention paid to calibration and quality control. This is important since available instrumentation for gaseous pollutant measurements seem to manifest a tradeoff between stability and reliability of measurement and sensitivity of response to ambient changes and to low concentration levels. Since the focus of the RAMS network was on a complete concentration record, as evidenced by the production of complete records on minute intervals (see below), the instrument choice in the RAMS network tended to fall on the side of sensitivity and rapidity of response time which necessitated a comprehensive set of calibration and maintenance procedures. Measurements at the level of tenths and hundreths of parts per million (ppm) require sensitive instruments which are notoriously subject to internal drift and instability. The gaseous pollutant analyzers were automatically calibrated daily between the hours of 2000 and 2400 C.S.T. by an operating sub-system of the central computer which collected the data. Each gaseous pollutant measurement except hydrogen sulfide was so calibrated. Calibration was performed by a zero and one upscale point in the normal operating range of the instrument. If the instrument could change ranges, calibration was performed on the low range. Zero levels were obtained by introducing zero air into the calibration manifold. Calibrated source concentrations for the upscale point were diluted with air scrubbed of water and all pollutants and then introduced into the calibration manifold. Multipoint calibration was automatically performed at 10 week intervals. The hi-vol particulate samples were checked every 6 months, the dichotomous samplers were checked periodically with a flow meter calibrated at Lawrence Berkeley Labs. Technical details of these procedures can be found on pages 33-42 of the "Documentation of the Regional Air Pollution Study . . . ." cited below. Extensive preventive maintenance procedures also were performed on the RAMS hardware. In addition, quality audits were performed independently of RAMS operating, personnel. Table 4 is a reproduction of the summary of audit results on the RAMS network. Despite the apparently high average measurement errors and the wide range of errors encountered, the RAMS network is still considered to be one of the most reliable monitoring systems ever put together. The general conclusions of the audit procedures involving pollutant measurements can be paraphrased into the following list: (Documentation, p. 55).

1) Given proper calibration and maintenance procedures, most of the gaseous pollutants measurements were reasonably accurate.



RAMS Monitoring Sites
St. Louis AQCR

Figure 1

TABLE 2

RAMS REMOTE STATIONS INSTRUMENT DISTRIBUTION

	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125
O_3 - Monitor Labs 8410	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
NO - NO_x Monitor Labs 8440	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
CO - CH_4 THC Beckman 6800	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
TS - SO_2 - H_2S Tracor 270HA	X		X	X	X	X							X	X	X	X				X	X	X			
TS - Meloy SZ 185		X					X		X	X	X	X					X	X	X				X	X	
Visibility - MRJ 1561		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Wind Speed - MRI 1022 S		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Wind Direction - MRI 1022D		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Temperature - MRI 840-1		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Den Point - Cambridge 880	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Temp. Gradient - MRI 840-2	X	X		X	X	X	X		X		X	X	X									X	X		
Barometer - MRJ 751/YSI 2014	X								X			X										X	X	X	
solar Pyranometer			X	X				X						X				X				X			
Radiation Pyrheliometer			X					X						X				X				X			
(Eppley) Pyrogeometer			X					X						X				X				X			
Turbulence - R. M. Young 27002					X		X		X		X		X												
Gas Bags - Xonics	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Hi - Vol, Sierra 305			X		X	X		X				X			X			X		X		X			
LBL Dichotomous Sampler			X		X	X		X				X			X			X		X		X		X	
10 Meter Tower								X		X				X	X	X	X	X				X			
30 Meter Tower	X	X	X	X	X	X	X		X		X	X	X						X	X		X	X	X	

TABLE 3

INSTRUMENTATION USED IN POLLUTANT DATA COLLECTION

POLLUTANT	RAMS
Ozone	Monitor Labs 8410A chemiluminescent analyzer, Range: O-.20 ppm or O-.5 ppm, span value: min det: .005 ppm, ht: 4 meters
Oxides of Nitrogen	Monitor Labs 8440 chemiluminescent analyzer. NO, NO_x continuously - NO_2 by differencing Range: O-.50 ppm Span: .10 ppm Min det: .005 ppm ht: 4 meters
Hydrocarbons, carbon monoxide	Beckman Instrument, Model 6800 gas chromatograph, Ranges: CO: 0-10 or C-50 ppm ht: 4 meters
Sulfur dioxide & total sulfur	Tracor, Inc., Model 270 HA chemiluminescent analyzer. Total sulfur SO_2 simultaneous ranges 0-.2 ppm, O-1.0 ppm. Hydrogen sulfide: simultaneous ranges: O-.1 ppm, O-1.0 ppm. Min det: .005 ppm ht: 4 meters
Total Sulfur	Meloy Labs. Model SA-185 chemiluminescent analyzer range automatically selected: O-.2 ppm or O-1.0 ppm. Min det: .005 ppm ht: 4 meters
Wind speed	Meteorology Research, Inc. Model 10225 3-cup anemometer: 10 or 30 meter ht. range: .22-22.35 M/S 1 percent accuracy
Wind Direction	Meteorology Research, Model 1022D wind vane 10 or 30 meters. Range: 0-540° threshold: .3. Damping ratio: .4 at 10° angle
Temperature	Meteorology Research Model 840-1 dual thermistor resistor network: ht: 5 meters. Range: -20° C - +50° C
Barometric Pressure	Meteorology Research, 751/YST. Model 2014 transducer. ht: 2 meters Range: 27.0-31.5 inches (947-1080mb)
Solar Radiation	Eppley Labs precision spectral pyranometer, normal incidence pyrheliometer: ometer and pyrgeometer. 5 meters. Range: O-4 cal cm^2/min
Gas Bags	Xonic Inc. bag sampler, 100 liter teflon & tedlar bags 4 meters - ht flow rate: O-1200 cc/min.
T.S.P.'s	Sierra Instruments, Model 305. High vol. sampler. 20 x 25 cm. glass fiber filter - flow rate .02 m^3/sec. - by model 320 flow controller .4 meters
TSP's Fractionated	Lawrence Berkley Labs, dichotomous air sampler. Two size ranges, greater than and less than 2 micrometers

TABLE 4

SUMMARY OF AUDIT RESULTS

	Parameter Measured	Method	Average Error Found in Field Stations Calibration (Computed as slope error)	Range of Errors for the 8 Stations Audited
1.	Nitric Oxide	Chemiluminescence	12.5%	-5.8% to +29.2%
2.	Nitrogen Dioxide	Chemiluminescence	11.7%	-6.7% to +29.3%
3.	Total Oxides of Nitrogen	Chemiluminescence	12.3%	-5.7% to +27.4%
4.	Sulfur Dioxide	Gas Chromatograph and Flame Photometric	34.8%	-55.2% to -5.3%
5.	Hydrogen Sulfide	Gas Chromatograph		
6.	Total Sulfur	Gas Chromatograph		
7.	Ozone	Chemiluminescence	15.6%	-37.9% to +14.8%
8.	Total Hydrocarbons	Gas Chromatograph	12.8%	-23.9% to -3.6%
9.	Methane	Gas Chromatograph	22.1%	-31.3% to -10.5%
10.	Carbon Monoxide	Gas Chromatograph	17.2%	-32.0% to -1.5%
11.	Nitric Oxide Cylinder Concentration	HSL Calibrator-CPT	24.4%	+49.1% to +6.6%
12.	Permation Tube Water Temperature	Thermistor	0.4°C	-0.7°C to -0.3°C
13.	Dew Point	Thermistor	1.8°C	-2.3°C to +5°C
14.	Ambient Temperature	Thermistor	0.6°C	-1.6°C to +1.1°C
15.	NO_2 Converter Efficiency	15 of 16 tested 1 of 16 tested	greater than 90% efficiency 82% efficiency	

2) Ambient carbon monoxide levels less than 2 ppm were consistently measured low network-wide. This was because the Beckman 6800 analyzer had a breakthrough CO concentration for measurement of .3186 ppm. The actual minimum value was "not consumed by the over-concentrated stripper column." (Documentation, p. 67). The understated error decreases as the span pt. is approached.

3) The sulfur analyzers considerably improved in the later part of network operation, after 1975.

4) Details of the audit results conducted throughout RAMS' operation can be found in publication listed on p. 57 of "Documentation . . . ."

Data were collected for all variables except the high volume gas bag and automatic dichotomous samples on a minute by minute basis except for automatic calibration periods. Each station included a mini-computer on the premises which fed to 2 central computer data management system. Back-up storage was included at the RAMS site. The minute values represent averages over 120 1/2 second readings. A complete data record was produced every minute and the high volume samples obtained a 24-hour sample every three days. The gas bag samples took a sample on an as-required basis. The automatic dichotomous samples collected a sample every 2 to 12 hours depending on study requirements. The body of the RAMS data base dates between January 1, 1975 and March 31, 1977. Table 5 gives the exact dates of operation for each of the RAMS network stations.

For the purposes of this health study, not every RAMS station was utilized. In particular, six of the stations, those numbered 116, 118 and 122-125 on Figure 1, were completely outside the health survey study area. Each respondent's reported home location always was closer to one of the other 19 stations. The measurements from those stations were those utilized in assigning air pollution exposures to the survey respondents. Table 6 gives a characterization of the extent of missing data in the data base. Cells with an asterisk indicate variables not measured by a particular station, so that a value of 438 or 439 is expected. Also shown in Table 6 are the mean days with no data across stations by variable. This is computed using only the 19 stations in the health study, and only those stations which were supposed to be measuring a parameter. Such an aggregate view of the missing data characteristics is perhaps only relevant when judging the quality of the data base in its use studying chronic health effects. System-wide, equipment measuring ozone and nitric oxide seemed to be particularly reliable, with the equipment measuring hydrogen sulfide being particularly problematical. It is somewhat puzzling that although one instrument, the Monitor Labs 8440, was responsible for measurements of nitric oxide, nitrogen dioxide, and oxides of nitrogen; the mean days missing over the last two measurements was more than twice that of the first.

B. A Correlation Analysis of the Air Quality Data

The correlation analysis of the RAMS air quality data is designed

TABLE 5

OPERATION DATES OF RAMS STATIONS

OPERATION DATES:

STATION	START	STOP
101	8/15/74	3/31/77
102	6/28/74	3/31/77
103	6/24/74	3/31/77
104	8/20/74	6/30/77
105	8/02/74	3/31/77
106	4/22/74	6/30/77
107	8/28/74	6/30/77
108	8/02/74	3/31/77
109	6/19/74	3/31/77
110	8/21/74	3/31/77
111	8/26/74	6/30/77
113	6/24/74	3/31/77
114	8/20/74	3/31/77
115	6/02/74	6/30/77
116	7/21/74	3/31/77
117	7/09/74	3/31/77
118	7/26/74	3/31/77
119	9/04/74	3/31/77
120	8/27/74	3/31/77
121	7/26/74	6/30/77
122	8/09/74	3/31/77
123	8/07/74	2/12/77
124	7/24/74	2/12/77
125	7/14/74	6/30/77

TABLE 6

DAYS WITH NO DATA

STN	TEMP	BARO	OZONE	CRBMON	METH	HYDRO	NITOX	NITDIO	OXNIT	SULFUR	HYSULF	SULDIO
1	15	98	19	90	45	73	26	119	126	43	14	30
2	0	439*	1	20	28	29	6	17	19	55	439*	439*
3	1	439*	11	29	64	54	22	63	65	15	26	13
4	2	439*	15	53	76	74	178	27	24	22	73	22
5	1	439*	8	83	54	74	7	10	10	26	52	25
6	2	439*	18	30	65	47	18	26	21	117	162	64
7	0	439*	5	40	59	73	4	29	26	66	439*	438*
8	0	439*	6	22	19	78	10	12	15	42	114	37
9	2	0	19	71	25	15	62	74	48	82	439*	438*
10	0	439*	12	32	57	47	18	55	49	46	438*	439*
11	11	439*	11	65	88	81	11	102	102	60	439*	438*
12	13	11	11	77	55	172	14	52	49	70	438*	439*
13	0	439*	7	88	84	76	12	30	23	22	140	21
14	1	439*	18	93	53	66	15	25	20	20	70	29
15	4	439*	20	96	71	66	15	18	18	51	352	46
**16	9	439*	43	106	98	87	31	39	36	53	77	50
17	7	439*	21	44	49	52	50	03	48	20	439*	438*
**18	9	439*	23	12	24	22	25	27	27	55	438*	439*
19	1	439*	17	19	66	40	20	59	54	59	439*	438*
20	51	439*	37	70	68	70	26	40	36	58	116	63
21	0	439*	17	80	33	56	15	42	37	40	135	27
**22	1	3	24	52	93	65	22	26	25	30	175	22
**23	2	2	70	56	70	70	26	34	29	71	439*	438*
**24	72	2	53	79	75	72	21	45	46	24	438*	439
**21	12	85	47	110	112	146	36	45	37	53	439*	438*
Mean days w/o data	5.8		14.4	58.0	55.7	65.4	19.4	45.4	41.6	48.1	114	34.3

*: indicates station not equipped to measure this variable
**: indicates station not used in health-impact estimation

to obtain information of use in specifying the regression models estimated in Chapter 6. Two types of correlations are calculated: (1) for a given aerometric measure between stations and (2) for different aerometric measures at a given station. The main interest in making these calculations lies in discerning which readings tend to be highly correlated, thereby duplicating exposure information. More specifically, the first type of correlation coefficient shows the extent to which individuals living at different locations in St. Louis are subject to different exposure patterns. The second type of correlation coefficient, then, shows which air quality measures have 2 high linear association. That knowledge could help to avoid a multicollinearity problem in later empirical work and to define better proxies for missing variables. For the following correlation estimates, only data from the times 0600 to 2000 were used. As mentioned before, calibrations were performed between 2000 and 2400 daily and from 2400 to 0600 the bulk of the population is inactive. All coefficients cited are Pearson correlation coefficients, all test of significance are two-tailed.

1) Correlation Between Stations for Given Measures

Beginning with temperature, as would be expected, all stations are highly correlated, for the most part at values of .98 and up. The outstanding exception is Station 120 which appears to consistently read lower temperatures than the rest of the stations in the network. For this one station, correlation with other stations are in the .60-.75 range.

Ozone measurements also stand out as implying a relatively homogenous exposure field over the entire region. There are no uncomputable correlations between stations, and no negative correlations. There also are no insignificant relationships between stations, the only pollutant for which this was true. The Pearson coefficients range from a maximum of .8988 between Stations 102 and 111 to a minimum of .3317 between 109 and 115. Stations 102 and 111 are both located in St. Louis city along the Mississippi River although 102 is in the north part of town and 111 is at the southern end of the city limits. Both 109 and 115 are in Illinois at the eastern edge of the study region with 109 being adjacent to East St. Louis and 115 being approximately 20 miles north adjacent to Southern Illinois University at Edwardsville. Despite these comparatively high correlations, however, there was substantial variation in the mean values of the ozone observations between stations. The lowest mean reading, for Station 104 was about 54 percent of the highest mean reading, which occurred at Station 121.

Carbon monoxide revealed wide variation in correlation coefficients between stations. The maximum coefficient was between 120 and 101; .8926 with the minimum being between 105 and 114; .0005. The exposure field, thus, appears to vary considerably for carbon monoxide. There are no negative correlations or uncomputable coefficients. Out of a potential 171 relationships in the exposure field, the null hypothesis of a discontinuity (e.g., no relationship) in the exposure field along a particular gradient between two stations could not be rejected at a 5 percent level for only 18

cases. All low correlations (insignificant coefficients) were geographically dispersed. Stations 101 and 120, although approximately 15 miles from each other (east to west), are both enclosed by the intersection of major transportation arteries. Stations 105 and 114 are widely separated, 105 being downtown next to the intersection of I-55 and I-44, with 114 being in a rural area north of St. Louis near the intersection of the Missouri and Mississippi Rivers. The picture formed is one of sub-regions with a uniform exposure field, each sub-region still well connected to other sub-regions rather than the whole region being relatively homogeneous as with ozone.

The methane field exhibits, in general, a much weaker set of relationships among stations than all other pollutants except sulfur dioxide and hydrogen sulfide (see below). The maximum is between stations 101 and 102 with a coefficient of .3983. Of the 171 relationships between stations, 23 are negative, and in approximately 40 cases the null hypothesis of an exposure field discontinuity is accepted at the 10 percent level. The lowest correlation coefficient in this set was -.0723; the lowest positive coefficient was .0025. Both of those were not significantly different from zero at the 10 percent level. The pattern of relationships, as well as appearing weak, also appears to be quite complicated. For example, the methane readings from Stations 106 and 107, both located in St. Louis about three miles apart, arc basically uncorrelated. Station 106 is sandwiched quite closely between two freeways, while 107 is not. However, Station 108, which is across the Mississippi River in Illinois about seven miles away has 2 significant relationship with Station 106 at the 5 percent level. Station 108 is located in an industrial area subject to a fair amount of ship, barge, and rail traffic. For the most part, however, widely dispersed stations exhibit low correlations. Assuming no measurement problems, the picture which emerges in a very localized pollutant exposure field, not well connected into one large urban field, heavily dependent on specific site characteristics.

The total hydrocarbon field is more akin to the carbon monoxide field than to the methane field. The positive range is from a maximum of .6268 between Stations 102 and 107 to a minimum of .0338 between 115 and 105. There is only one negative correlation, between 119 and 108, but that coefficient was not significantly different from zero at the 5 percent level. By way of contrast to methane, only 12 out of 171 Pearson correlation coefficients insignificantly different from zero, 141 relationships exhibited a coefficient greater than .2 (significant at the 1 percent level), showing a well-connected relatively homogenous field. Insignificant relationships appear well-behaved in terns of expected geographical dispersion.

The oxides of nitrogen variables all present a similar exposure field image among one another. The nitric oxide and total oxides of nitrogen coefficients, which are based on direct readings; are particularly similar. The nitrogen dioxide correlation matrix, although very similar, is not as directly comparable to the former as the former are to each other. For nitric oxide, the coefficient range is between .7986 between Stations 107

and 111, and .0959 between Stations 106 and 117. At a 10 percent level, the null hypothesis of no relationship can be rejected for all coefficients. There were no negative correlations. Most of the coefficients are in the .4 to .7 range.

The total oxides of nitrogen correlation matrix shows a range of .7493 between Stations 114 and 108 down to .0547 between 105 and 121. Relationships are overwhelmingly significant except for two, the minimum cited above and .0650 between Stations 106 and 109. Both of these coefficients came from stations that were quite distant from each other. Although the total oxides of nitrogen and nitric oxide fields appear similar in structure, they are not identical. For example, some relationships between stations over the two pollutants are almost identical others are not. Between Stations 108 and 101, the nitric oxide correlation coefficient is .6822, the oxides of nitrogen coefficient .6911. However, between Stations 106 and 110, the nitric oxide correlation is .6116, the oxides of nitrogen correlation is .2739.

The nitrogen dioxide matrix is obtained from measurements derived by differencing the nitric oxide and oxides of nitrogen readings. The positive range, similar to the above, is from a maximum of .6210 between Stations 114 and 108 down to .0106 between Stations 117 and 106. The maximum is between the same two stations as the total oxides of nitrogen maximum, the minimum between the same two stations as the nitric oxide minimum. There were five relationships with negative coefficients, of which one could be considered non-zero, between Stations 121 and 110, which are widely dispersed. In general, the relationships between stations are significant but of lower magnitude than the direct measurements. Only further study can determine whether this situation is due to the measurement method for nitrogen dioxide or whether there is a qualitative difference between the exposure field for nitrogen dioxide and either the fields of nitric oxide or total oxides of nitrogen. In any case, one obtains an image of a well-diffused pollutant class for oxides of nitrogen, not as homogenous as ozone, but certainly more homogenous than the hydrocarbon group.

Characteristics change dramatically when the sulfurous group of pollutants are considered. Turning first to the total sulfur readings, the positive coefficients range from .4877, between Stations 102 and 111, down to .0028 between Stations 101 and 105. There are many negative coefficients, representing relationships between 15 sets of stations. Among these negative coefficients, the null hypothesis of no relationship was never rejected at the 5 percent level. Among the 171 relationships, approximately 65 involved coefficients where the null hypothesis of no relationship could not be rejected at the 5 percent level. Total sulfur was measured at all stations, but not by the same instrument at every location. Two instruments were used to cover the total sulfur field. Whether this had an impact on consistency of measurement between stations is not known, although the structure of the coefficient matrix is similar to that of the related pollutants sulfur dioxide and hydrogen sulfide, which were both generated by a single instrument.

Hydrogen sulfide measurements generate the most unusual pattern of correlations of all the inter-station readings. Measurement equipment was missing at several stations, although there were no unexpected uncomputable correlations. Among computable correlations, 68 are negative and 51 are positive. The positive coefficients range from .6407 between Stations 115 and 104 down to .0029 between Stations 114 and 103. The negative values range from -.0043 between Stations 105 and 108 down to -.1134 between Stations 121 and 113. Overall, these correlations exhibit low significance; only seven relationships could be characterized as rejecting the null hypothesis. The maximum coefficient cited above is a true outlier. The two stations involved are widely geographically dispersed, making for some suspicion as to the validity of the coefficient. The related coefficients between 115 and 104 for total sulfur and sulfur dioxide are .1151 and .1277, respectively. If there is one exposure field generated by the RAMS system of which to be suspicious, the hydrogen sulfide field would be a prime candidate.

Finally, sulfur dioxide's relationship matrix is similar in structure to hydrogen sulfide's. There were no unexpected uncomputable correlations, with the total number of relationships computed matching that of hydrogen sulfide's field. There are 93 positive and 16 negative computed relationships. The positive range was from .2662 between Stations 114 and 115 down to .0008 between Stations 104 and 105. The negative range was from -.0017 between 119 and 140 to -.0479 between 121 and 108. Stations 104 and 105 are right across the river from each other in high density industrial and transport areas while 114 and 115 are in rural areas north of St. Louis across the river from each other, but further apart than 104 and 105. Stations 121 and 108 are quite widely separated, while 119 and 120 are fairly close together in what appear to be similar residential communities. Not much rhyme or reason can be made from the geographic characteristics of these extreme values. There are about double the significant relationships among sulfur dioxide measurements as compared with the hydrogen sulfide measurements; however, sulfur dioxide is still a very disconnected, heterogeneous exposure field. Among the sulfurous measurements, the exposure field can be treated as either extremely localized (the most localized of any pollutant class) or as unreliable. There is not enough information to differentiate rigorously between these possibilities.

2) Correlations Between Measures for Given Stations

Of greatest interest among the single station, between measure-estimates is the possible use of total sulfur as a proxy for hydrogen sulfide and/or sulfur dioxide, since the former was measured at every station, while the latter two were only measured at eleven out of the nineteen stations in the study zone. Between SO_2 and total sulfur, the average correlation over eleven cases was .77 with a standard deviation of .10. Two stations exhibited extremely high correlations between SO_2 and total sulfur: station 113: .9368, and station 101: .9136. Between total sulfur and hydrogen sulfide, over eleven stations, the average correlation was .33 with a standard derivation of .28. The range in this set was from

.0542 to .8396. Perhaps total sulfur can be used as a proxy for SO_2, however, this is probably not the case for hydrogen sulfide.

Another feature of the single station-between pollutant correlations is the preponderance of negative correlations between ozone and all other pollutant measurements. Many of these negative correlations were significant at least at the 5 percent level. Of the 156 ozone--other pollutant correlations by station, 131 were negative and 25 were positive. Station 113 had the highest incidence (6) of positive and positive significant correlations. The rest of the positive correlations were fairly widely distributed over stations and pollutants. Temperature was highly positively correlated with ozone, as might be expected. Mean correlation was .55 with a standard deviation of .09. The incidence of significant negative correlations (5 percent level) not involving barometric pressure, temperature, or ozone measurements in the relationship was very rare. There are only five: (1) Station 106 between SO_2 and methane, (2) Station 106 between CO and methane, (3) Station 110 between CO and methane, (4) Station 117 between sulfur and CO; and, (5) Station 120 between hydrogen sulfide and methane. In a very general sense, then, the data suggest that exposure to ozone and exposure to other pollutant groups are inversely related.

The data further suggest that total suspended particulates (TSP's) share this characteristic of being inversely related to other pollutant groups. Because of the nature of TSP measurement techniques, correlations between TSP's and other variables were performed differently than correlations between other pollutants. Of the ten stations taking TSP measurements, only eight operated throughout the study period. Of these, two took samples of six hours instead of twelve hours. TSP measurements consist of blowing atmospheric air through a filter for a specified period and measuring the accumulated particulate mass. For the six remaining stations, the twelve hour duration TSP measurement starting at noon was matched to the daily mean for ozone, total hydrocarbons, oxides of nitrogen, and total sulfur. A Pearson zero-order correlation coefficient matrix was then calculated using the small and large TSP measurements separately. Focusing on results for Station 108 suffices to illustrate the resulting pattern. At station 108, the correlation between small TSP particles (TSPS) and ozone was .1589, while the coefficient between large TSP particles (TSPL) and ozone was .1951. The correlation between total hydrocarbons and TSPS was -.0113, between total hydrocarbons and TSPL the correlation was -.2000. Between oxides of nitrogen and TSPS and TSPL, the coefficients were -.1288 and -.1602 respectively. For total sulfur the respective coefficients were .0690 and .0203. In other words the pattern can be described as a small but significant positive correlation between ozone and TSP with roughly equivalent negative correlation between TSP and the groups of hydrocarbons and oxides of nitrogen. Between TSP's and the sulfur group the correlations are weaker and of mixed sign. The pattern is somewhat more pronounced for large particles than for small particles.

Among the oxides of nitrogen group, there would perhaps be some interest in using nitric oxide as a proxy for NO_2, or total oxides of

nitrogen, given the relatively lesser incidences of missing data as described by Table 6. Among stations in the study area, the mean correlation between oxides of nitrogen and nitric oxide was .77 with a standard derivation of .14. Between nitric oxide and NO_2 the mean correlation was .49 with a standard derivation of -16. Parenthetically, the mean correlation between NO_2 and oxides of nitrogen was .77 with a standard derivation of .11, almost identical to the nitric oxide, oxides of nitrogen correlation above. Unfortunately, oxides of nitrogen was not the variable with the substantially fewer missing values!

Among the hydrocarbon group, only methane and total hydrocarbons even approached being modestly correlated. Again, among study area stations, this calculation set had a mean of .70 with a standard deviation of .10 comparing favorably with other highly correlated sets. Carbon monoxide and methane, and carbon monoxide and total hydrocarbons, were consistently significantly related, but at much lower levels (less than .3 average). The existence of the "automobile effect" is evident by the average correlation of approximately .45 between total hydrocarbons and both nitric oxide and total oxides of nitrogen.

4. Summary

This chapter has provided a description of the health and air quality data to be used in empirically implementing the theoretical model presented in Chapter 3. That description focused on the types of variables measured as well as how the data were collected. The results of a correlation analysis performed on the air quality data obtained from the RAMS also were presented in order to assess the degree of association between: (1) measurements of the same pollutant at different locations and (2) measurements of different pollutants at the same locations. One key question not adequately addressed in this chapter, however, concerns the rationale for using air quality data from the RAMS system. Chapter 5 pursues this question together with the implications of choosing the RAMS data for the empirical estimates to be reported later.

CHAPTER 5

A COMPARISON OF ALTERNATIVE SOURCES OF AIR QUALITY DATA

1. Introduction

As previously indicated, information from the "St. Louis' Pollution - Morbidity" study, performed by Geomet Technologies, Inc., was to have been collected concurrently with aerometric data generated by the Regional Air Monitoring System (RAMS). However, this planned matching of extensive health and daily activity information with detailed air quality measurements did not occur when delays in approval of the survey design prevented implementation of the first phase of data collection until June, 1978. The RAMS system had entirely shut down by March 31, 1977. One of the reasons for continuing with the Geomet study in spite of this problem was the existence of locally operated pollution monitoring networks in the study area including those administered by the city and county of St. Louis, and the Illinois EPA. The purpose of this chapter is to evaluate the quality of the aerometric data produced by these three sources. The results of this evaluation, which are contained in Section 2, suggest that the three non-RAMS data sources are quite unreliable and should not be used in an econometric health epidemiology study. Section 3, then, discusses the constraints placed by this finding on the use of the St. Louis health data set.

2. Quality of the Non-RAMS Aerometric Data

Aerometric monitoring systems generally are plagued with measurement drift problems if highly sensitive devices are used. Correspondingly, if less sensitive measurement devices are used, changes in ambient pollutant concentrations may not be accurately detected. The audit documentation presented in Chapter 4, together with informal opinion related by scholars in the atmospheric science field indicated that the RAMS was distinguished by the use of sensitive instruments, the attention paid to the calibration of those instruments in particular, and overall quality control in general. Most monitoring devices measuring gaseous pollutants were automatically recalibrated daily as part of many routine on-site physical inspections and maintenance procedures. From that evidence, the conclusion was drawn that the RAMS network data could be taken as the best available benchmark for measuring actual pollutant levels. Decisions concerning how to best use the St. Louis health survey data in a particular study design hinge in large part on the quality of the data from non-RAMS, locally administered pollution monitoring systems. As a step in judging this quality, a correlation study was performed which compared the locally-administered

stations' pollutant time series with the corresponding time series from RAMS Stations.

As a first step, pollution data from the appropriate regional EPA administrations were collected for all county, city and Illinois EPA monitoring stations in the St. Louis Air Quality Control Region (AQCR) for the decade 1970-1980, Preliminary analysis of these sources revealed a relatively small degree of overlap between operation dates and a large incidence of missing data for entire days and/or fractions of days. Based upon this preliminary analysis, a procedure was devised for a RAMS vs. non-RAMS pollutant time series comparison. The best quality RAMS data, after making the system operational, were in the latter portion of the system's operation. This distinct improvement in RAMS system quality is a conclusion based both on the audits performed on the RAMS system[1] and discussion with atmospheric scientists familiar with the system. Accordingly, the dates of 1/1/76 - 3/14/77 were chosen as the dates for generating the RAMS benchmark time series, by pollutant. Within these dates, daily averages were computed for the hours of 0600-2000, inclusive. This time frame was chosen because of the importance of activity in exascerbating the negative effects of pollution exposure, the generally higher pollution levels during the day, and because RAMS automatic calibration procedures were performed between the hours of 2000-2400 daily. For each pollutant and station, two time series were generated: the daily means over all valid hourly readings and daily means over the 90 percent decile determined by the daily peak.

Given these benchmark time series, the next step was to generate a time series that matched pollution measurements from non-RAMS and RAMS stations for the purpose of calculating a set of Pearson correlation coefficients. Because of the variation in operation periods and valid data days, the analysis proceeded on a non-RAMS station-pollutant basis. In other words, given a non-RAMS station that operated at least intermittently over the period 1/1/76 and 3/14/77, a data matrix was constructed in order to generate a set of pollutant specific correlations between the non-RAMS station chosen and each of the RAMS stations in the St. Louis area. Between station correlation coefficients were calculated for the following representatives of the major pollutant groups; ozone, sulfur dioxide, nitric oxide, and carbon monoxide.

The actual program generating the time series data matrix by station for a given pollutant proceeded as follows. A date was selected in the non-RAMS time series. If more than twelve hours were missing from that date's hourly measurements, that date was ruled invalid. If a date had measurements for at least the minimum number of valid hours then an average reading over the hours of 0600-2000 was computed. In order to be identical to the RAMS daily average, readings below a threshold of .0025 ppm were classified as missing and did not enter into the average computed for all pollutants except carbon monoxide. If all hours were classified as missing, the date was ruled invalid. When a valid date had been found and a valid average computed, the matching date from the 439-day RAMS time series was read and the nineteen RAMS daily averages and the one non-RAMS daily average were output on an external file. After repeating this

procedure over all candidate dates in the non-RAMS time series, the matrix of valid dates by stations was then input into a standard SPSS Pearson zero-order correlation procedure which generated the correlation coefficients, number of days in the generated time series, and the means and standard deviation of measurements at each station over the time series. Thus, correlations were run between a given non-RAMS station, in turn, with all RAMS stations over identical days with identically-computed daily averages over all readings above a given threshold and over all readings in the daily 90 percent decile.

Results from this data manipulation and analysis are presented in Tables 1-8. St. Louis County operated stations have a CO prefix, St. Louis City stations a CI prefix, and Illinois EPA stations an IL prefix in these tables. The overlays in Figures 1 and 2 compare the location of the stations in each of the four monitoring systems.

For each of the four pollutants there are two tables, one for daily means and one for 90 percent decile averages. The tables of daily means give the correlations of each non-RAMS station with the two or three closest RAMS stations, as well as the correlations between these RAMS stations. In all tables, n.a., indicates that data were not available for a particular pollutant or a particular RAMS station. The tables of daily means also include the mean readings for all stations explicitly considered; the high correlation or correlations among the set of all RAMS stations with the non-RAMS station, and the number of days in the time series generating the above data. The tables showing the 90 percent decile averages do not include correlations between the two or three closest RAMS stations. Each table also contains the critical value of the correlation coefficient, ρ_c, to test the null hypothesis $\rho = 0$ using a two-tailed test at the 5 percent level of significance. Assuming that the daily mean and 90 percent decile values are normally distributed, then

$$\hat{\rho}[(n - 2)/(1 - \hat{\rho}^2)]^{\frac{1}{2}} \sim t(n - 2) \qquad (1)$$

The number of observations used in computing the values for $\hat{\rho}$ exceeded 40 in all cases but two. Consequently, in that overwhelming majority of cases, values for $\hat{\rho}_c$ were computed by substituting the number 2 as an approximation to the critical value $t_c(n - 2)$ into equation (2). ρ_c then was computed as

$$\hat{\rho}_c \approx t_c(n - 2)/(n + 2)^{\frac{1}{2}} \qquad (2)$$

In both of the two cases where $n < 40$, $n = 16$ and the value $t_c(n - 2) = 2.145$ was used.

Tables 1 and 2 present the results for ozone daily means and 90 percent decile averages, respectively. Among the county and Illinois EPA stations and their spatially close RAMS counterparts, the means for both daily means and 90 percent decile measurements agree quite closely except for station CO6 in Table 1. Among the city stations, however, the average daily means are considerably less than the average among close RAMS

stations, although the 90 percent decile means agree fairly closely. An important pattern to notice, however, is the consistently lower correlations between each non-RAMS station and close RAMS stations as compared to correlations between these RAMS stations. In Table 1, the correlations between non-RAMS and RAMS stations range from .10 to .57, while correlations between RAMS stations range from .44 to .88 for daily means. Given the RAMS data as a benchmark of quality, a distinct lack of quality in the non-RAMS systems would be indicated. Particularly striking are the low correlations between non-Rams and RAMS stations which were essentially operating at the same location. For example, the correlations between the Illinois EPA Station IL7 and RAMS Station 104 are -36 for daily means and .31 for 90 percent decile averages. City Station CI5 was located very close to RAMS Station 101 in downtown St. Louis. Their daily means correlation is -15 while their 90 percent decile correlation is .25. In the county system, Station CO4 and RAMS Station 112 are in close proximity. Their daily means correlation is .26 while the 90 percent decile average correlation is .27. In all six of these cases, $\hat{\rho} > \hat{\rho}_c$; thus the null hypothesis $\rho = 0$ should be rejected in favor of the alternative hypothesis, $\rho \neq 0$. However, even though there appears to be some linear association between the ozone readings from adjacent stations, that association simply is too weak to support the use of the non-RAMS in this study. This conclusion obviously is based upon an implicit (and arbitrary) lower bound of "acceptability" for the values of ρ. However that lower bound for adjacent stations surely must exceed .40. In fact, that correlation implies that a linear regression of the RAMS ozone readings on non-RAMS ozone readings would produce an $R^2 = .16$!

Moreover, another anomaly seen in Tables 1 and 2 is the high incidence of highest correlation between a non-RAMS station and the set of nineteen RAMS stations being with RAMS Station 102. Stations as geographically disparate as C06, IL2, and IL6 all were most highly correlated with Station 102 in the 90 percent decile average time series. Although no reasons are given here for this phenomena, it is suggestive of the problems involved in using the non-RAMS data as a source for exposure data. There appear to be anomalous patterns which need to be explained before use of the data can be justified.

As inconsistent as the ozone RAMS and non-RAMS data appear, the other three pollutants checked showed dramatically lower correlations. Tables 3 and 4 display correlations for nitric oxide (NO) computed from daily means and daily 90 percent decile averages, respectively. The daily mean correlations between RAMS and non-RAMS readings range from -.0034 to .21, while the RAMS intercorrelations range from .35 to .76. For the daily 90 percent decile correlations, the RAMS--non-RAMS values ranged from -.04 to .26, with the latter figure being something of an outlier. Choosing stations in close proximity again does not appear to increase the correlation between the NO readings from RAMS and non-RAMS stations. Between Stations IL6 and 109 there is essentially nothing but a random relationship between the two daily means NO time series; indicated by a correlation coefficient of -.007. Between Stations CO4 and 112, the NO daily means time series comparison shows a correlation of .026. For the city system, the same daily mean time series comparison between Station CI5

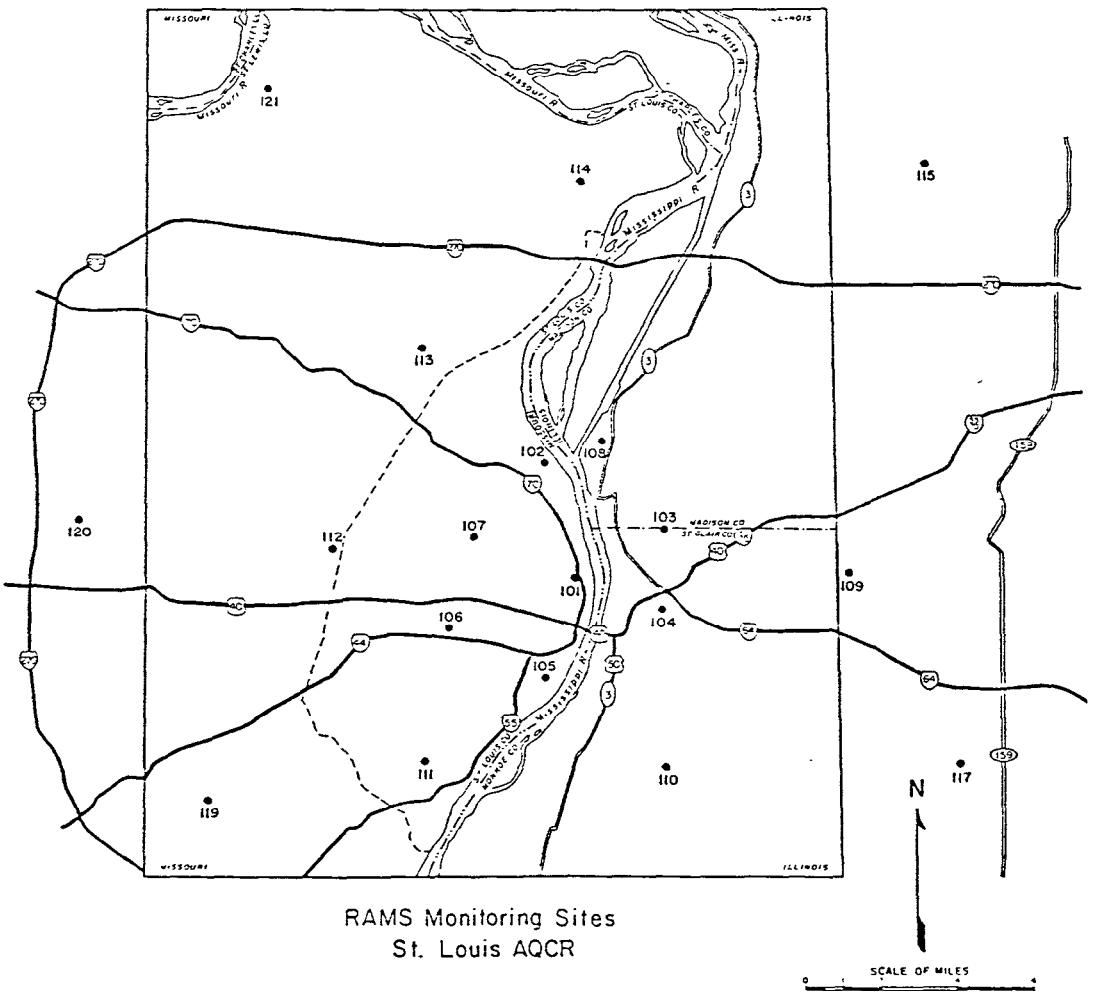
and 101 generates a correlation of .02. Furthermore, Tables 3 and 4 again show the anomalous pattern of highest correlation being with RAMS stations quite far away geographically combined with low correlation with RAMS stations in close proximity.

Tables 5 through 8 continue to point this same dismal picture of data quality from the non-RAMS systems. Correlations reported for daily means and 90 percent decile readings for the pollutants, carbon monoxide and sulfur dioxide frequently are negative and seldom exceed the critical values shown for $\hat{\rho}_c$; a situation that persists even when immediately adjacent stations are considered. For example, the correlations for carbon monoxide readings on both daily means and daily 90 percent deciles are negative for IL7 and 104, CI5 and 101, and IL6 and 109. However, the correlation between daily mean carbon monoxide readings at CO4 and 112 was .15, which was just slightly greater than the critical value of .1421. Finally, with respect to sulfur dioxide, there was only one case of immediately adjacent monitoring station. The correlations between daily mean and daily 90 percent decile readings for IL7 and 104 were .06 and -.003 respectively.

3. Constraints on Using the St. Louis Health Data

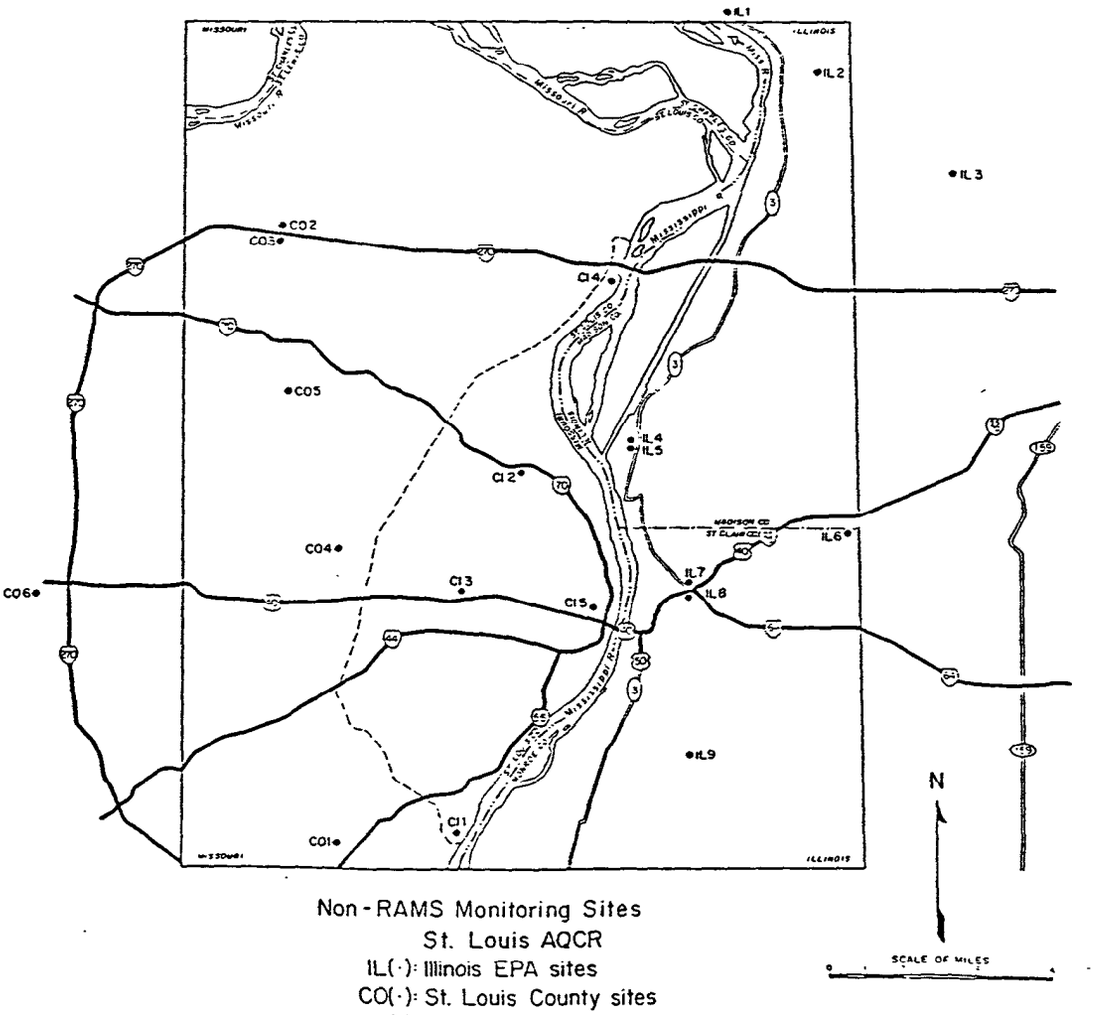
Among the four pollutants examined, only the ozone data obtained from the non-RAMS monitoring networks appears to bear a significant positive association with concurrent ozone readings from the RAMS network. However, that association is weak; for immediately adjacent stations it was never higher than .36. Telephone interviews with officials in charge of operating the St. Louis City and St. Louis County monitoring systems reinforced a suspicion as to the likely cause of the poor-quality data: the lack of a full-scale calibration and maintenance for the instruments in these systems when the time series were generated.

In any case, because of these rather extraordinary problems with the non-RAMS sources of air quality data, the full potential of the health survey data cannot be realized. More specifically, the diary data simply are unusable in a study designed to assess the consequences of air quality changes on short-term or acute illness. Such a study would require reasonably accurate air quality measures matched on a day-by-day basis to the diary health data. If those air quality measures do not exist, there would appear to little that can be done to salvage the situation, regardless of the care that may have been used in collecting the health data. Of course, the correlation study just reported covered the period 1/1/76-3/14/77 whereas the health survey was conducted between June, 1978 and July, 1979. However, there is no evidence that improved monitoring procedures were instituted by the non-RAMS systems in the interim. As a consequence, the empirical work reported in Chapter 8 makes no use of the health diary information at all. Only data obtained from the Household Background Questionnaire, the Individual Background Interview, and the Supplemental Interview were examined. Prior to use, they were matched with air quality data from the RAMS using a procedure described more fully in Chapter 6. The RAMS data, then, are interpreted as reflecting historical or long-term exposure patterns faced by residents of the St. Louis area.



RAMS Monitoring Sites
St. Louis AQCR

Figure I



Non-RAMS Monitoring Sites
St. Louis AQCR
IL(·): Illinois EPA sites
CO(·): St. Louis County sites
CI(·): St. Louis City sites

Figure 2

TABLE 1

CORRELATIONS BETWEEN DAILY MEANS: OZONE

		119	111	110	
C01 (30001)		.35	.39	.33	
119			.74	.68	
111				.83	
					C01
Means:		.026	.023	.029	.028
High:	.453, with 108				
Days:	104				
$\hat{\rho}_c$:	.1943				

		113	120	121	
C02 (20002)		.42	.28	.44	
113			.68	.79	
120				.62	
					C02
Means:		.030	.028	.035	.027
High:	.494, with 111				
Days:	118				
$\hat{\rho}_c$:	.1826				

		112	120		
C04 (1040001)		.26	.12		
112			.59		
					C04
Means:		.023	.023		.026
High:	.419, with 102				
Days:	221				
$\hat{\rho}_c$:	.1339				

		104	103	105	
IL7 (2120008)		.36	.33	.40	
104			.54	.55	
103				.59	
					IL7
Means:		.018	.021	.020	.017
High:	.535, with 102				
Days:	398				
$\hat{\rho}_c$:	.1000				

(Table 1, Continued)

Table 1, continued

		109	103	104	
IL6 (2120009)		.26	.20	.27	
109			.64	.55	
103				.54	
					IL6
Means:		.031	.020	.018	.0015
High:	.386, with 102				
Days:	397				
$\hat{\rho}_c$:	.1001				

		112	120	113	
C05 (4120001)		.17	.10	.29	
112			.59	.72	
120				.68	
					C05
Means:		.024	.024	.027	.027
High:	.33, with 102				
Days:	159				
$\hat{\rho}_c$:	.1579				

		111	105	106	
CI1 (4280007)		.27	.20	.19	
111			.71	.77	
105				.66	
					CI1
Means:		.023	.019	.021	.024
High:	.369, with 103				
Days:	214				
$\hat{\rho}_c$:	.1361				

		102	107	113	
CI2 (4280061)		.39	.30	.36	
102			.73	.88	
109				.67	
					CI2
Means:		.024	.019	.025	.003
High:	.39, with 102				
Days:	336				
$\hat{\rho}_c$:	.1088				

(Table 1, Continued)

Table 1, continued

		106	111	112	
CI3 (4280062)		.35	.36	.26	
106			.77	.65	
111				.76	
					CI3
Means:		.025	.027	.025	.003
High:	.41, with 102				
Days:	297				
$\hat{\rho}_c$:	.1157				

		113	114	102	
CI4 (4280063)		.47	.42	.57	
113			.80	.87	
114				.82	
					CI4
Means:		.019	.026	.018	.003
High:	.57 with 102				
Days:	249				
$\hat{\rho}_c$:	.1262				

		101	105	106	
CI5 (4280064)		.15	.16	.17	
101			.65	.73	
105				.66	
					CI5
Means:		.026	.022	.025	.002
High:	.26 with 102				
Days :	294				
$\hat{\rho}_c$:	.1162				

		120	119	
CO6 (4300006)		.22	.46	
120			.54	
				CO6
Means :		.024	.027	.0025
Highs:	.62, with 102.			
Days :	226			
$\hat{\rho}_c$:	.1325			

(Table 1, Continued)

Table 1, continued

		115	114	121	
IL2 (8520007)		.31	.46	.45	
115			.46	.44	
114				.74	
					IL2
Means:		.026	.031	.031	.0025
Highs:	.57, with 102				
Days:	402				
$\hat{p}_c$:	.0995				

Table 2, continued

		101	105	106	
CI5 (4260064)		.25	.24	.21	
					CI5
Means:		.046	.036	.042	.028
High:	.31 with 102				
Days :	291				
$\hat{\rho}_c$:	.1168				

		120	119		
CO6 (43000061		.22	.41		
					C06
Means:		.040	.044		.036
High:	.56, with 102				
Days:	224				
$\hat{\rho}_c$:	.1330				

		115	114	121	
IL2 (8520007)		.30	.38	.40	
					IL2
Means :		.041	.049	.049	.044
High:	.47, with 102				
Days:	402				
$\hat{\rho}_c$:	.0995				

TABLE 3

CORRELATIONS BETWEEN DAILY MEANS: NITRIC OXIDE

		119	111	110	
CO1 (30001)		-.0034	-.024	-.0345	
119			.76	.68	
111				.71	
					CO1
Means :		.012	.015	.007	.032
Highs:	-0356, with 112				
Days:	200				
$\hat{\rho}_c$:	.1407				

		113	120	121	
CO2 (200002)		.027	-.011	.26	
113			.57	.52	
120				.35	
					CO2
Means:		.0162	.0084	.0082	.023
High:	.26, with 121				
Days:	116				
$\hat{\rho}_c$:	.1841				

		112	120		
CO4 (1040001)		.026	.097		
112			.62		
					CO4
Means:		.021	.012		.038
High:	.0972, with 120				
Days:	198				
$\hat{\rho}_c$:	.1414				

		109	103	104	
IL6 (2120009)		-.007	.11	.094	
109			.53	.41	
103				.54	
					IL6
Means :		.0072	.011	.023	.036
Highs:	.23, with 107				
Days:	343				
$\hat{\rho}_c$:	.1077				

(Table 3, Continued)

Table 3, continued

		112	120	113	
co5 (4120001)		-.04	.06	-.06	
112			.57	.64	
120				.60	
					CI1
Means:		.02	.01	.02	.024
High:	.06, with 120				
Days :	195				
$\hat{\rho}_c$:	.1425				

		102	107	113	
CI2 (4280061)		.19	.21	.09	
102			.60	.62	
107				.70	
					CI2
Means:		.024	.022	.014	.008
High:	.22, with 111				
Days:	206				
$\hat{\rho}_c$:	.1387				

		106	111	112	
CI3 (4280062)		.12	.13.	.19	
106			.69	.59	
111				.73	
					CI3
Means:		.019	.013	.016	.008
High:	.23, with 113				
Cays:	111				
$\hat{\rho}_c$:	.1881				

		101	105	106	
CI5 (4280064)		.02	.028	-.026	
101			.45	.61	
105				.59	
					CI5
Means:		.02	.026	.022	.024
High:	.19, with 115				
Days:	54				
$\hat{\rho}_c$:	.2673				

(Table 3, Continued)

Table 3, continued

		120	119	
CO6 (4300006)		.23	.14	
120			.62	
				CO6
Means:		.011	.012	.023
High:	.32, with 103			
Days:	203			
$\hat{\rho}_c$:	.1397			

TABLE 5

CORRELATIONS BETWEEN DAILY MEANS: CARBON MONOXIDE

		119	111	110	
CO1 (30001)		.0036	.0256	-.077	
119			.59	.79	
111				.58	
					CO1
Means:		.74	.88	.60	.36
High:	.0545 with 120				
Days:	224				
$\hat{\rho}_c$:	.1325				

		113	120	121	
CO2 (200002)		-.065	-.12	-.023	
113			.55	.36	
120				.42	
					CO2
Means:		1.14	1.56	.26	.37
High:	.1053, with 104				
Days:	110				
$\hat{\rho}_c$:	.1890				

		112	120	
CO4 (1040001)		.15	-.12	
112			.15	
				CO4
Means:		.91	1.03	.34
High:	.15 with 112			
Days:	196			
$\hat{\rho}_c$:	.1421			

		104	103	105	
IL7 (2120008)		-.18	-.08	.075	
104			.13	.32	
103				.24	
					IL7
Means:		.80	2.04	5.16	.19
High:	.075, with 105				
Days:	16				
$\hat{\rho}_c$:	.4714				

(Table 5. Continued)

Table 5, continued

		109	103	104	
IL6 (2120009)		-.09	-.05	-.06	
109			.18	.33	
103				.19	
					IL6
Means:		.17	.68	.74	.15
High:	.076, with 108				
Days:	290				
$\hat{\rho}_c$:	.1170				

		112	120	113	
CO5 (412001)		.26	-.05	-.015	
112			.15	.12	
120				.55	
					CO5
Means:		.86	1.06	1.05	.55
High:	.30, with 102				
Days:	156				
$\hat{\rho}_c$:	.1591				

		111	105	106	
CI1 (4280007)		-.07	-.06	-.10	
111			.38	.58	
105				.41	
					CI1
Means:		.71	.61	1.05	.22
High:	.09, with 104				
Days:	247				
$\hat{\rho}_c$:	.1267				

		102	107	113	
CI2 (4280061)		.13	-.05	.022	
102			.18	.28	
107				.26	
					CI2
Meals:		.44	1.83	.99	.29
High:	.135, with 102				
Days:	321				
$\hat{\rho}_c$:	.1113				

(Table 5, Continued)

Table 5, continued

		106	111	112	
CI3 (4250062)		n.a.	n.a.	n.a.	
106			n.a.	n.a.	
111				n.a.	
					CI3
Means:		1.19	.79	.77	.19
High:	n.a.				
Days:	141				
$\hat{\rho}_c$:	.1672				

		113	114	101	
CI4 (4280063)		.08	.04	.13	
113			n.a.	n.a.	
114				n.a.	
					CI4
Means:		1.006	.23	.45	.17
High:	.18, with 105				
Days:	318				
$\hat{\rho}_c$:	.1118				

		101	105	106	
CI5 (4280064)		-.23	-.095	-.26	
101			.56	.70	
105				.41	
					CI5
Means:		.97	.65	.97	.57
High:	.03, with 112				
Days:	319				
$\hat{\rho}_c$:	.1116				

		120	119		
CO6 (4300006)		-.23	-.27		
120			.88		
					CO6
Means:		.70	.56		.57
High:	.03, with 112				
Days :	152				
$\hat{\rho}_c$:	.1612				

(Table 5, Continued)

Table 5, continued

		115	114	121	
IL2 (8520007)		.0042	-.02	.019	
115			.46	.27	
114					.46
					IL2
Means:		.26	.24	.36	.11
High:	.1329, with 101				
Days:	289				
$\hat{\rho}_C$:	.1172				

TABLE 6

CORRELATIONS BETWEEN DAILY 90% DECILES: CARBON MONOXIDE

		119	111	110	
CO1 (30001)		.034	-.004	-.089	
					CO1
Means:		1.89	2.50	1.25	.486
High:	.034, with 119				
Days:	224				
$\hat{\rho}_c$:	.1330				

		113	120	121	
CO2 (200002)		-.06	-.12	.085	
					CO2
Means :		3.03	3.29	.90	.55
High :	.085, with 121				
Days:	110				
$\hat{\rho}_c$:	.1890				

	112	120		
CO4 (1040001)	.076	-.049		
				CO4
Means:	.277	2.40		.52
High:	.13, with 103			
Days:	197			
$\hat{\rho}_c$:	.1418			

		104	103	105	
IL7 (2120008)		-.14	.04	-.03	
					IL7
Means:		1.71	5.31	9.72	.26
High:	.13, with 107				
Days :	16				
$\hat{\rho}_c$:	.5055				

		109	103	104	
IL6 (2120009)		-.055	-.009	-.046	
					IL6
Means:		.69	1.46	1.99	.18
High:	.15, with 108				
Days:	318				
$\hat{\rho}_c$:	.1118				

(Table 6, Continued)

Table 6, continued

		112	120	113	
CO5 (4120001)		.07	-.047	-.042	
					CO5
Means:		2.56	2.45	3.14	.87
High:	.27, with 102				
Days:	156				
$\hat{\rho}_c$:	.1591				
		111	105	106	
CI1 (4280007)		-.013	-.03	-.05	
					CI1
Means:		2.09	1.96	2.21	.41
High:	.06, with 121				
Days:	248				
$\hat{\rho}_c$:	.1265				
		102	107	113	
CI2 (4280061)		.15	-.014	.04	
					CI2
Means:		1.37	4.09	3.01	.51
High:	.15, with 102				
Days:	323				
$\hat{\rho}_c$:	.1109				
		106	111	112	
CI3 (4280062)		-.023	-.014	.02	
					CI3
Means:		2.17	2.16	2.52	.35
High:	.066, with 102				
Days:	255				
$\hat{\rho}_c$:	.1248				
		113	114	102	
CI4 (4280063)		-.OO1	-.005	.08	
					CI4
Means:		3.08	.81	1.39	.28
High:	.098, with 121				
Days:	326				
$\hat{\rho}_c$:	.1104				

(Table 6, Continued)

Table 6, continued

		101	105	106	
CI5 (4280064)		-.11	-.006	-.11	
					CI5
Means:		2.57	1.99	2.20	.78
High:	.095, with 103				
Days:	319				
$\hat{\rho}_c$:	.1116				

		120	119		
CO6 (430006)		.014	.019		
					CO6
Means:		1.82	1.64		.37
High:	.11, with 102				
Days:	154				
$\hat{\rho}_c$:	.1601				

		115	114	121	
IL2 (852000)		.008	-.006	-.07	
					IL2
Means:		.67	.83	1.04	.15
High:	.10, with 119				
Days:	317				
$\hat{\rho}_c$:	.1120				

TABLE 7

CORRELATIONS BETWEEN DAILY MEANS: SULFUR DIOXIDE

		119	111	110	
CO1 (30004)		n.a.	n.a.	n.a.	
119			n.a.	n.a.	
111				n.a.	
					CO1
Means:		n.a.	n.a.	n.a.	.019
High:	.09, with 114/.05 with 106/.07 with 101				
Days:	190				
$\hat{\rho}_c$:	.1443				

		114	121	115	
IL1 (160006)		.0017	.078	-.14	
111			.16	n.a.	
121				.10	
					IL1
Means:		.012	.008	.010	.029
High:	.08, with 121				
Days:	53				
$\hat{\rho}_c$:	.2697				

		113	120	121	
CO2 (200002)		-.05	.11	-.006	
113			.14	.12	
120				.09	
					CO2
Means:		.013	.018	.015	.014
High:	.11, with 120				
Days:	120				
$\hat{\rho}_c$:	.1181				

		112	120	
CO4 (1060001)		n.a.	.018	
112			n.a.	
				CO4
Means:		n.a.	.007	.010
High:	.21, with 113			
Days:	190			
$\hat{\rho}_c$:	.1443			

(Table 7, Continued)

Table 7, continued

		104	103	105	
IL7 (2120008)		.06	.024	.010	
104			.04	.0008	
103				.016	
					IL7
Means:		.032	.019	.020	.024
High:	.06, with 104				
Days:	396				
$\hat{\rho}_c$:	.1003				

		109	103	104	
IL6 (2120009)		n.a.	-.19	-.12	
109			n.a.	n.a.	
103				.04	
					IL6
Means:		n.a.	.009	.014	.025
High:	.176, with 114				
Days:	63				
$\hat{\rho}_c$:	.2481				

		108	103	102	
IL4 (29600012)		.12	-.14	n.a.	
108			.010	n.a.	
103				n.a.	
					IL4
Means:		.011	.009	n.a.	.024
High:	.13, with 114				
Days:	67				
$\hat{\rho}_c$:	.2408				

		112	120	113	
CO5 (4120001)		n.a.	-.045	-.064	
112			n.a.	n.a.	
120				.20	
					CO5
Means:		n.a.	.013	.019	.017
High:	.13, with 103				
Days:	194				
$\hat{\rho}_c$:	.1479				

(Table 7, Continued)

Table 7, continued

		102	107	113	
CI2 (4280061)		n.a.	n.a.	.12	
102			n.a.	n.a.	
107				n.a.	
					CI2
Means:		n.a.	n.a.	.009	.030
High:	.32, with 114				
Days:	41				
$\hat{\rho}_c$:	.3050				

		113	114	102	
CI4 (4280063)		-.027	-.077	n.a.	
113			.11	n.a.	
116				n.a.	
					CI4
Means:		.0085	.014	n.a.	.024
High:	.043, with 121				
Days:	43				
$\hat{\rho}_c$:	.2981				

		115	114	121	
IL2 (8520007)		-.048	-.105	.035	
115			.26	.10	
114				.16	
					IL2
Means:		.0096	.012	.0073	.042
High:	.17, with 106				
Days:	65				
$\hat{\rho}_c$:	.2443				

TABLE 8

CORRELATIONS BETWEEN 90% DECILES: SULFUR DIOXIDE

		119	111	110	
CO1 (30001)		n.a.	n.a.	n.a.	
					C01
Means:		n.a.	n.a.	n.a.	.40
High:	.14, with 105/.06, with 108/.087 with 114				
Days:	190				
$\hat{\rho}_c$:	.1443				
		114	121	115	
IL1 (160006)		-.027	-.06	-.12	
					IL1
Means:		.04	.021	.036	.072
High:	.025, with 108				
Days:	53				
$\hat{\rho}_c$:	.2697				
		113	120	121	
CO2 (200002)		-.09	.033	-.075	
					CO2
Means:		.043	.037	.055	.030
High:	.033, with 120				
Days:	120				
$\hat{\rho}_c$:	.1811				
		112	120		
CO4 (1040001)		n.a.	.059		
					CO4
Means:			.025		.025
High:	. 17, with 113				
Days:	190				
$\hat{\rho}_c$:	.1443				
		104	103	105	
IL7 (2120008)		-.003	-.OO2	.06	
					IL7
Means:		.086	.045	.049	.062
High:	.066, with 114				
Days:	396				
$\hat{\rho}_c$:	.1003				

(Table 8, Continued)

Table 8, continued

		109	103	104	
IL6 (2120009)		n.a.	-.22	-.068	
					I16
Means:		n.a.	.025	.035	.057
High:	.27, with 120				
Days:	63				
$\hat{\rho}_c$:	.2481				

		108	103	102	
IL4 (2960012)		.21	-.18	n.a.	
					IL4
Means:		.036	.025	n.a.	.051
High:	.21, with 108				
Days:	67				
$\hat{\rho}_c$:	.2408				

		112	120	113	
CO5 (4120001)		n.a.	.041	.04	
					CO5
Means:		n.a.	.032	.063	.035
High:	.19, with 103				
Days:	194				
$\hat{\rho}_c$:	.1429				

		102	107	113	
CI2 (4280061)		n.a.	n.a.	-.01	
					CI2
Means:		n.a.	n.a.	.028	.069
High:	.57, with 114				
Days:	41				
$\hat{\rho}_c$:	.3050				

(Table 8, Continued)

Table 8, continued

		113	114	102	
CI4 (4280063)		-.08	-.07	n.a.	
					CI4
Means:		.028	.045	n.a.	.061
High:	-.0015, with 104				
Days:	43				
$\hat{\rho}_c$:	.2981				

		115	114	121	
IL2 (8520007)		.041	-.13	.086	
					IL2
Means:		.032	.038	.017	.10
High:	.21, with 106				
Days:	65				
$\hat{\rho}_c$:	.2443				

FOOTNOTES

1. In particular, discussions with August Auer, Professor of Atmospheric Science, University of Wyoming, Laramie.

CHAPTER 6

EMPIRICAL ESTIMATION OF THE MODEL

1. Introduction

This chapter presents estimates of the willingness to pay for reductions in ozone levels in St. Louis. The methodology used in making these calculations focuses on estimating the marginal bid expression derived in Chapter 3

$$dB/d\alpha = H_\alpha q_M / H_M \tag{1}$$

where H_α denotes the marginal effect of a change in air pollution on the stock of health, H_M denotes the marginal effect of a change in the consumption of medical services on the health stock, and q_M denotes the full price of medical care. Section 2 discusses two alternative approaches that could be used to estimate equation (1) and selects one of these for empirical implementation. Section 3 shows how the variables used in the empirical analysis were constructed from the information given in the St. Louis health survey and the RAMS data. (The St. Louis health survey also is described in detail in Appendix 3.) Section 4, then, presents the results. More specifically, this section gives estimates of the health production function as well as calculations of the marginal willingness to pay for reduced ozone levels. Section 5 offers a brief summary.

2. Estimation of the Willingness to Pay

As indicated, this section surveys two possible approaches to estimating equation (1). The first approach seeks to obtain separate estimates of H_α and H_M from the health production function based on a procedure similar to that used by Rosenzweig and Schultz (1982a, 1982b). Rosenzweig and Schultz, who examined the health of children rather than adults, used birthweight as an indicator of the heaith stock. Unfortunately, from the St. Louis health data set, no single comprehensive parallel measure of the size of an adult's health stock can be constructed. Information concerning subjectively reported health status (excellent, good, fair, poor), existence of specific types of chronic conditions, and length of suffering from chronic conditions is available for each individual; however, these variables alone may not accurately measure the theoretical concept of the stock of health. This point will be discussed more fully momentarily. Nevertheless, the conclusion drawn is that the stock of health may be best treated as a multi-dimensional, rather than a single dimensional variable. That perspective underlies the second approach to obtaining $dB/d\alpha$, in which only the ration H_α/H_M is estimated.

A. The Rosenzweig and Schultz Approach

The Rosenzweig and Schultz method for estimating the health production uses a two stage procedure. First, a reduced form demand equation for the health input, M,is derived from the model as shown in equation (2) and empirically estimated

$$M = M(q_X, q_M, W, \alpha) \qquad (2)$$

All variables in those equations are defined as in Chapter 3. The fitted values of M then are used in the second stage in order to estimate the health production function

$$H = H(M; \alpha) \qquad (3)$$

The idea behind using this procedure is to explicitly account for the interaction between the choice variables H and M and to obtain statistically consistent estimates of the health production function. In other words, the health of an individual is determined by the level of the health inputs used in the production of health and other exogenous factors. The levels of the health inputs are, in turn, determined by the individual on the basis of all relevant information known to him. Because of the interaction of the demand for health inputs and the production of health, simultaneity exists in this system of equations, and thus the whole system of equations must be estimated in order that the resulting coefficients are consistent. A two stage procedure which utilizes the estimated input demand equations to calculate fitted values for the health inputs and then uses these fitted values in place of the actual values when estimating the health production function would eliminate this inconsistency.

In order to estimate a production function for health in this manner, a variable that measures the stock of heaith must be available. This presents a problem since the stock of health is a theoretical concept and not easily quantified. Rosenzweig and Schultz used birthweight as a proxy for the health stock of children; however, as previously indicated, the available health stock measures pertaining to adults in the St. Louis data probably are not as good. In any case, the following subsection provides a critical evaluation of the health information that the St. Louis data set contains.

B. The St. Louis Health Data

Three measures are available from the St. Louis data set that directly pertain to an individual's stock of health. These are the categorical self-evaluation of health; i.e., reported health status is excellent, good, fair, or poor; the number and types of chronic conditions or illnesses; and the length of time the individual has suffered from these conditions on illnesses. Each of these measures have been used in other empirical health related analyses; although none of them are without limitations For example, the self-evaluation of health is a highly subjective measure. What is considered excellent health to one individual may only be good health to another. Moreover, someone with a chronic

illness may consider himself to be in good health as does someone who gets several colds a year. But do both of these individuals actually have the same health stock or is there an important difference in their health status? Another problem with this measure is how to quantify the four possible responses. How much more health stock does an individual in excellent health possess as opposed to an individual in poor health? This measure could be expressed as a categorical variable but much information on the actual health stock would be lost.

Additionally, expressing the health stock of an individual as a function of the number and types of existing chronic illnesses also presents significant problems. One problem was mentioned in the previous paragraph. How can comparisons be made between the health stock of an individual with a chronic illness and the health stock of an individual with no chronic illness? Is it true that the presence of a chronic illness implies that the individual possesses a smaller health stock than those without any chronic illnesses? Compare, for example, the person who is rundown, gets frequent colds, and feels tired much of the time to the person who has asthma but takes very good care of his health. In this case, the absence of a chronic illness does not necessarily imply a larger health stock.

Finally, the duration of chronic illnesses is most meaningful when used in conjunction with information on the number and types of such health problems. Without knowing the condition from which the individual is suffering, information on the length of time it has been present probably tells little about the magnitude of the health stock.

A further problem with each of these variables taken separately is that they each may be measuring different dimensions of the health stock. To illustrate, if the available three variables were comprehensive measures of the health stock, then they should be highly correlated. That situation, however, does not materialize in the St. Louis data set as shown in Table 1. The Pearson correlations between possible health stock variables does have a plausible sign pattern in that the incidence of chronic conditions is negatively associated with the excellent health variables and positively associated with the poor health variables. Nevertheless, the linear associations between these variables are not particularly strong. CHRO and POOR have a Pearson correlation of .313 and CHRO and EXCELLENT have a corresponding correlation of -.295. The correlations between the subjective health evaluation variables and the duration of chronic illness measures are quite similar to those shown in Table 1 since that latter variable is either zero or non-zero whenever the incidence of chronic illness variable is zero or non-zero.

On the basis of this discussion, assigning great faith in the available measures of the health stock probably is unwarranted. Nevertheless, these measures easily may provide greater information about the health stock when used together, rather than when used separately. In other words, if a choice must be made as to which single variable is best suited to serve as a proxy for the health stock in a regression equation, there would be little on which to base it. As a

TABLE 1

PEARSON CORRELATIONS BETWEEN POSSIBLE HEALTH STOCK VARIABLES

	CHRO	RESPCIRC	OTHERCH
POOR	.3125	.1633	.256
FAIR	.2894	.2401	.1804
GOOD	.0380	.0561	-.0001
EXCELLENT	-.295	-.213	-. 197

* Variable definitions are:

CHRO: Indicates presence of a chronic illness
RESPCIRC: Indicates presence of a respiratory or circulatory illness
OTHERCH: Indicates presence of a chronic illness other than of a respiratory or circulatory nature
POOR: Indicates respondent is in poor health
FAIR: Indicates respondent is in fair health
GOOD: Indicates respondent is in good health
EXCELLENT: Indicates respondents is in excellent health

consequence, the design of the estimation strategy presented below shows how multiple indications of the health stock can usefully be incorporated into the equation to be estimated.

C. An Alternative Approach

A suitable alternative approach to that proposed by Rosenzweig and Schultz for estimating the quantity dB/da easily can be obtained under the assumption that the health production function is approximately linear. The assumption of linearity in logarithms or some other simple function also could be applied with out affecting the nature of the approach considered. Let

$$H_i = \beta_0 + \beta_1 M_i + \beta_2 \alpha_i + U_i \qquad (4)$$

where the subscript i denotes the ith individual and U_i denotes a random disturbance capturing unmeasured variables affecting H_i. To more easily allow for the fact that H_i may be best measured by a set of health indications rather than a single variable rewrite equation (4) as

$$M_i = -(\beta_0/\beta_1) + (1/\beta_1)H_i - (\beta_2/\beta_1)\alpha_i + V_i \qquad (5)$$

where the new disturbance V_i can be expressed as

$$V_i = -U_i/\beta_1 \qquad (6)$$

Two features of equations (4), (5), and (6) are worth elaborating. First, in equation (5), the coefficient of the air quality variable α_i is just the negative of the analogue to H_α/H_M from the theoretical model? Therefore, when equation (5) is empirically estimated, the negative of the coefficient on air quality need only be multiplied by q_M, the full price of medical care, in order to find $dB/d\alpha$. Second, (4) is likely to be an overidentified equation from the structural system to which it belongs. The other three structural equations for the theoretical model determining M_i, X_i, and the value of the Lagrange multiplier, λ_i, would contain the predetermined variables, q_{Xi}, q_{Mi}, and W_i. Those prices would not enter the production function. Moreover, equation (4) contains an air quality measure which would not appear in the other three structural equations. As the theoretical model demonstrates, air quality affects the values of the choice variables only through its impact on health.

The remainder of this chapter is concerned with estimating equation (5) and the expression for $dB/d\alpha$ using the St. Louis health and the RAMS air quality data. Section 3 shows how the sample on which these estimates are based was constructed and how the variables in (5) were defined. Section 4 presents the empirical results.

3. Sample Construction and Variable Definitions

A. Sample Construction

For the purpose of this study only these individuals whose major activity was recorded as employed were included in this sample. (See Appendix 3 for a more complete description of these data.) The reason for excluding all others was that no data were available to assess their value of time; a necessary ingredient in computing the full prices. The value of time to an employed individual is imputed to be the wage rate and no comparable measure is available for others who may be retired, too young to work, etc. In order to calculate this measure, more detailed information is required. Of the 2197 employed persons in the survey, only 820 provided hourly wage data.

B. Variable Definitions

The variables used in the empirical analysis can be divided into five categories, those measuring: (1) consumption of medical services, (2) the price of medical services, (3) the wage rate, (4) socioeconomic-demographic characteristics, and (5) air quality. Each of these categories is considered sequentially.

1) Medical Care

Existing information in the St. Louis health data set measuring the consumption of medical services certainly are far from perfect. What is desired here is a variable that would measure whether an individual

received regular medical check ups; however, there was no question in the survey that explicitly answered this question. The number of times the individual visited a doctor during the year preceding the Individual Background Interview (1977 or 1978 depending on when the respondent was enrolled in the survey) and a yes/no question asking if a doctor was usually seen at least once a year were recorded for each individual. If an individual had a rash of illnesses or developed a chronic illness the year preceding survey, his answer to the first question would be biased upward for this would not have been his response in a "typical" year. His response would be measuring the amount of time spent ill rather than whether he sought preventive medical care. Or as another example, a pregnant woman may visit the doctor every two weeks toward the end of her term; therefore, her response will be much higher than it would have been the preceding year and it may in fact measure the health of her child.

The second question still does not measure preventive medical care; however it may come closer than the first question. That "usually" is included in the question at least partially controls for any bias that may come from an individual experiencing an atypical year for doctor visits. Those who have visited the doctor for annual check ups should respond affirmatively to this question in addition to those who visited a physician for the treatment of disease only. Even if an individual has seen the doctor for a specific treatment he will have received more preventive medical care than a person who has received no medical care at all for a physician may notice other physical problems in the process of treatment or he may give the patient advice on preventive measures necessary to stay healthy. Thus, the second question may measure preventive medical care more accurately than the first.

2) Price of Medical Care

The price of medical care was constructed to take into account both direct dollar outlays for the medical care, the time costs involved in commuting to and from the source of medical care, and the waiting time at the source of medical care. The only available information on the direct dollar outlays was a question regarding the usual charge by the doctor for an office visit. Many individuals have health insurance which pays for part or all of their office visits, thus their actual dollar outlay may be much lower than that which is recorded for the office charge. The only data available which may alleviate this problem are obtained from a question asking the portion of an office visit paid by any health insurance. Unfortunately, the answers to this question are categorized very generally (all or most, some or about half, little or none) and provide little useful information. With these difficulties in mind, the price of medical care is defined as:

$$q_M = \text{Office Visit Charge} + \text{hourly wage (commuting time} + \text{office waiting time)} \quad (7)$$

where commuting time + office waiting time are measured in hours. The data on usual office visit charges and hourly wage rates came from the

supplemental survey and therefore are in 1980 dollars. Note that q_M probably overestimates the true cost per visit which individuals face since no account is taken of any insurance reimbursement of medical fees. Moreover, the hourly wage rate also may overstate the value of time.

3) Wage Rate

As shown in equation (7), the hourly wage rate is used as the value of time in computing the price of medical care. That wage rate is defined as follows. For the entire eight week follow-up period (in 1978 or 1979), individuals kept track of where they were during each 24 hour day. The data thus gives a record of the number of hours worked during the entire eight weeks. From this record an average number of hours per week was computed. Also, in the supplemental survey (1980) was a question regarding take home pay from an individual's full time occupation and the pay period corresponding to this amount. The amount of take home pay per week are computed from these two items and by dividing this figure by the average number of hours worked per week a variable measuring the hourly wage was constructed. As indicated in the previous subsection, data on take home pay was missing for a substantial fraction of workers.

4) Socioeconomic-Demographic Variables

In this chapter, the α variable has frequently been discussed, for the sake of simplicity, as if it referred exclusively to changes in air quality. However, in the model presented in Chapter 3, a was defined to be a vector of exogenous variables that affect the efficiency with which medical services are used in producing the health stock. This subsection makes that specification more explicit by stating which additional variables also might be included. These are: (1) age, (2) education, (3) race, and (4) sex.

Age, simply given by years of age, is included to reflect the loss of efficiency in producing the health stock as the individual grows older. Education is the number of years of schooling completed by the individual. Education is used here to measure human capital, which may affect how efficiently that person will combine inputs to produce a given stock of health. There is some difficulty when education is used to measure human capital, however. The stock of human capital will not be constant over a person's lifetime as suggested by the use of education. Instead, it will tend to increase as an individual receives on-the-job-training and more informal education and it can also depreciate as a person grows older. The use of years of schooling also poses some difficulties as the number of years of schooling does not always correspond directly with years of education. Thus, there are at least two sources of measurement error when years of schooling is used to measure education but there is probably an equal chance that the error is negative as it is positive. Finally, race and sex are both used as exogenous variables to test the hypothesis that one race or one sex may be more efficient producers of health than the other therefore their demand for the health inputs may vary accordingly.

5) Air Quality Variables

The only remaining exogenous variables are those measuring air quality in the St. Louis area. As discussed earlier this posed some difficulty because of the inopportune timing of the survey and the collection of the RAMS air pollution data. Had the RAMS air pollution data and the survey been conducted at the same time, it would have been possible to measure the relationship between work loss days and the amount of air pollution in the area on those same work loss days or the days just prior to them. The hypothesis being tested in this case would be whether people adjust their behavior on a day to day basis. In order to test this, it would also be necessary to have daily information on an individual's personal habits besides the daily air pollution readings and the record of work loss days and when they occurred.

If it is assumed that the yearly averages of each pollutant for each station are "typical" for any given year these yearly averages can be used as a measure of air quality for the time that the survey took place. The use of yearly data for one particular area poses a difficulty for this analysis. This study proposes to test the hypothesis that individuals may compensate for a decrease in air quality by adjusting their health related behavior. By using only one SMSA area and using annual data the only variation in air quality will arise from certain areas of the SMSA having higher concentration of air pollution than others. This variation may still be quite large. For example, areas that are congested with rush hour traffic normally have a much higher concentration of photochemical oxidants than areas that experience little peak hour traffic. But if annual data are used, it would be better to have observations on more areas in order to obtain a broader range of pollution exposures.

The pollution variables include the mean levels of ozone, sulfur dioxide, oxides of nitrogen, small (inhalable) total suspended particulates, and small (inhalable) suspended lead particulates. Other pollution variables were available but these represent most aggregate measures of pollutants. In Appendix 1, synergisms between ozone and other pollutants were discussed. Researchers have found evidence that ozone in combination with sulfur dioxides and nitrogen oxides may be more harmful to human health than any of these pollutants taken alone. Therefore, it is necessary to include these air quality measures in the list of independent variables. These variables are constructed by simply multiplying the two relevant pollutants together. Also constructed was a variable combining-mean levels of ozone and TSP in order to account for any synergistic effects between these two pollutants.

4. Empirical Results

In this section, estimates of the willingness to pay for improved air quality are presented using the methodology underlying equation (5) for two separate samples of St. Louis workers. The first sample is composed of the 820 persons for whom wage information is available and the second sample is composed of 2197 workers. For both of these samples, the equation to be

estimated is shown in equation (8)

$$MED_i = MED(\alpha_i, AGE_i, EDUC_i, SEX_i, RACE_i, H_i) \quad (8)$$

where α_i now denotes a set of air quality measures and H_i denotes a set of measures of the health stock. Additionally, each of the variables used to estimate equation (8) is defined in Table 2. That table also gives the arithmetic mean for each variable in each of the two samples. Note that the categorical self-evaluation of health variable is not listed in Table 2. Preliminary estimates of equation (8) using that variable produced poor results and consequently it simply was dropped from further consideration.

Moreover, given the variable definitions in Table 2, the components of α_i are measured as levels of air pollution rather than as measures of air quality. Hence, the expected signs on the coefficients of α_i would be positive and those coefficients can be used directly, without multiplying by minus one, in computing $dB/d\alpha$. Given the manipulation used to derive equation (8), the expected signs on CHRO and Length should be negative since increases in these variables are associated with decreases in the health stock. Also, the expected signs on the four socioeconomic-demographic variables are as follows: (1) the coefficient of AGE would be positive if the aging process reduces the efficiency with which the health stock is produced, (2) the coefficient of EDUC would be negative if years of schooling increase the efficiency with which health is produced and (3) the coefficients of SEX and RACE should be positive if males and blacks tend to have lower health stocks. Several measures of air quality were tried as independent variables. Since ozone was of interest it was included in all equation specifications. Measures of lead content in the air were also included in some of the specifications to test the findings reported in Appendix 2. Sulfur dioxides and total suspended particulates have also been found to cause deleterious health effects; therefore they were included in some specifications. In Appendix 1, it was mentioned that synergisms among pollutants may be more harmful than each pollutant taken separately. Several researchers found evidence that ozone and sulfur dioxides and ozone and nitrogen oxides could combine synergistically to produce more harmful effects to human health than does ozone alone. Therefore these two synergisms were included in several of the specifications.

The decision of which pollutants to include depended partially on the correlations among pollutants. The only correlation found to be particularly high was SULDIOM with OXSULM. Care was taken not to include two highly correlated air quality measures in the same equation so as to make the results more reliable. Pearson correlations between the pollution and nonpollution variables as well as between the nonpollution variables themselves were on the order of .20 or smaller and certainly not high enough to suggest a serious multicollinearity problem.

Because of the discrete nature of the dependent variable, MED_i, and the inclusion of the health stock (a choice variable) as a covariate, a simultaneous equation logit model was developed for each equation estimated. Predicted values for CHRO and LENGTH were obtained from their

TABLE 2

VARIABLE DEFINITIONS*

VARIABLE	DEFINITION	SAMPLE SIZE 820 MEAN	SAMPLE SIZE 2197 MEAN
MED	1: Denotes respondent sees a doctor at least once annually	.747	.755
PMED	Price of medical care	40.744	45.38
HWAGE	Hourly wage	5.078	5.078
RACE	1: Indicates person is black	.277	.241
AGE	Age in years	39.28	40.33
EDUC	Years of schooling completed	12.65	12.79
SEX	1: Indicates respondents is male	.539	.583
OZONEM	Mean ozone level (ppm)	.019	.019
SULDIOM	Mean sulfur dioxide level (ppm)	.024	.024
TSPSMLM	Total suspended particulates, small size, mean levels, (micrograms/m^3)	22.81	23.77
LEADSMLM	Lead, small size, mean (micrograms/m^3)	706.084	705.42
OZSULM	OZONEM x SULDIOM	<.000	<.000
OZNITM	OZONEM x OXNITM	<.000	<.001
CHRO	1: Denotes respondent reports presence of a chronic illness	.105	.108
LENGTH	Number of years respondent has had chronic illness	1.374	1.327

* See Appendix 3 for more complete details.

respective reduced form equations and then substituted into equation (8). Since CHRO also is a discrete variable, its reduced form equation, which uses PMED$_i$, HWAGE$_i$, α_i, SEX$_i$, RACE$_i$, AGE$_i$, and EDUC$_i$ as covariates, is estimated in a logit framework. The reduced form equation for LENGTH, on the other hand, which uses the same covariates, is estimated as a Tobit model. The frequency distribution of the length variable is characterized by a large number of zeros (more than half of the values are zero) and then integer values ranging as high as 27.

From an econometric viewpoint, the simultaneous equation system described above certainly is not a typical one in view of the fact that the dependent variables either are discrete or truncated. Nelson and Olson (1978), however, have demonstrated that the procedure used here, which is analogous to two-stage least squares, produces consistent and asymptotically normal estimates. Furthermore, on the basis of a small sample simulation experiment, those same authors conclude that estimated standard errors for the reported coefficients tend to be biased upward. Therefore, tests for whether those coefficients are statistically significant would be conservative.

When estimating equation (8) in the simultaneous equation logit framework, the probability, P_j, that the jth individual falls into the category that is assigned a value of one is equal to:

$$P_j = \exp\{Z_j\beta\}/[1 + \exp\{Z_j\beta\}] \tag{9}$$

where Z_j is the jth individual's vector of covariates and β is a coefficient vector. In this model, the marginal effect of the kth independent variable on P_j is given by:

$$\frac{\partial P_j}{\partial Z_{jk}} \quad \beta_K P_j(1 - P_j) \tag{10}$$

Thus, by estimating equation (8) in the form shown in (19), the effect of a change in air pollution levels on the probability of usually seeing a doctor at least once per year can be calculated. That effect corresponds to the H_α/H_M term in the willingness to pay expression, $dB/d\alpha$. However, the correspondence is rather inexact since MED$_i$ is a dummy variable, whereas PI in the theoretical model measures the quantity of medical services consumed. Also, since MED$_i$ indicates whether a doctor is usually seen at least once per year, the resulting estimate of H_α/H_M probably understates the true value of that parameter; a factor that would offset the overestimate of q_M discussed previously.

The marginal effects evaluated at the mean of each independent variable are reported in Tables 3 and 4. In parentheses are the χ^2 statistics (distributed with one degree of freedom) for each variable. These statistics test the null hypothesis that the marginal effect of the variable in question is equal to zero. Recorded below each column of derivatives are: (1) the number of cases (NC), (2) the χ^2 statistic (with degrees of freedom) for the equation as a whole which is useful in testing

TABLE 3

ESTIMATES OF THE HEALTH PRODUCTION FUNCTION
(DERIVATIVES EVALUATED AT THE MEAN)
820 CASES

	1	2	3	4	5	6
ONE	-.064 (.046)	-.307 (.269)	-.040 (.018)	-.010 (.001)	-.290 (.231)	.226 (.392)
OZONEM	30.52** (5.04)	33.14** (5.15)	31.25** (5.40)	29.74** (4.86)	33.06** (4.86)	25.72* (3.31)
SULDIOM		-.416 (.055)	.390 (.042)			
TSPSMLM		-.009 (.322)			.008 (.266)	
LEADSMLM						-.0003 (2.07)
OXSULM					-8.82 (.009)	95.94 (.785)
OZNITM			-91.90 (.424)	-78.12 (.370)		
AGE	-.003 (1.34)	-.003 (1.66)	-.004 (1.96)	-.003 (1.56)	-.003 (1.61)	-.003 (1.16)
EDUC	-.012 (1.29)	-.011 (1.37)	-.009 (.949)	-.010 (.975)	-.012 (1.50)	-.013 (1.84)
SEX	-.198*** (29.62)	-.200*** (30.46)	-.202*** (30.73)	-.199*** (29.84)	-.1999*** (30.29)	-. 196*** (29.14)
RACE	.064 (2.12)	.065 (2.19)	.079* (3.01)	.077* (3.00)	.065 (2.60)	.069 (2.44)
$\bar{CHRO}$	-.100 (.008)	-.036 (.001)	.203 (.036)	.044 (.002)	-.078 (.006)	-.248 (.055)
$\hat{LENGTH}$	.049 (1.11)	-050 (1.20)	.041 (.823)	.042 (.808)	.052 (1.32)	.057 (1.68)
NC	820	820	820	820	820	820
χ^2 (df)	51.1(7)***	51.7(9)***	51.7(9)***	51.3(8)***	51.7(8)***	53.3(9)***
FCP	618	618	618	618	618	618

***denotes significance at 1% level.
**denotes significance at 5% level.
*denotes significance at 10% level.

TABLE 4

ESTIMATES OF THE HEALTH PRODUCTION FUNCTION
(DERIVATIVES EVALUATED AT THE MEAN)
2197 CASES

	1	2	3	4	5	6
ONE	-.793***	-1.38***	-.936***	-.949***	-1.32***	-1.08***
	(26.58)	(23.73)	(20.92)	(24.18)	(22.40)	(23.24)
OZONEM	53.85***	56.64***	54.28***	55.86***	53.64***	55.13***
	(52.53)	(52.86)	(44.66)	(49.02)	(52.46)	(50.58)
SULDIOM		2.97**	2.73**			
		(6.04)	(5.03)			
TSPSMLM		.020**			.020**	
		(6.60)			(6.59)	
LEADSMLM						.0003**
						(5.48)
OXSULM					151.71**	91.04
					(6.04)	(1.79)
OZNITM			121.65	197.94***		
			(2.10)	(7.14)		
AGE	-.012***	-.011***	-.011***	-.011***	-.011***	-.011***
	(41.51)	(39.85)	(35.41)	(39.76)	(40.4)	(39.84)
EDUC	.005	.005	.004	.004	.005	.004
	(1.26)	(1.44)	(.957)	(.975)	(1.44)	(.899)
SEX	-.274***	-.307***	-.300***	-.313***	-.311***	-.306***
	(119.94)	(119.40)	(113.29)	(117.2)	(119.6)	(118.9)
RACE	.101***	.105***	.094***	.088***	.106***	.104***
	(15.51)	(16.16)	(12.89)	(11.82)	(16.52)	(15.87)
$\hat{CHRO}$	5.71***	5.19***	4.85***	4.88***	5.20***	5.12***
	(41.08)	(42.46)	(36.84)	(40.69)	(42.96)	(41.98)
$\hat{LENGTH}$	-.035	-.033	-.027	-.030	-.033	-.028
	(1.52)	(1.49)	(.979)	(1.16)	(1.51)	(1.16)
NC	2197	2197	2197	2197	2197	2197
χ^2(df)	149.5(7)	152.8(9)	146.4(9)	149.59(8)	153.1(9)	151.1(9)
FCP	1669	1676	1668	1667	1678	1668

***denotes significance at 1% level.
**denotes significance at 5% level.
*denotes significance at 10% level.

the null hypothesis of no relationship between the dependent and independent variables, and (3) the fraction of cases correctly predicted by the equation (FCP).

Table 3, which provides the estimates of the health production function based on the 820 cases for which wage data are available, shows that $\hat{CHRO}$ has the correct sign but is not significantly different from zero at conventional levels. $\hat{LENGTH}$ is neither correctly signed nor significantly different from zero. One explanation for this outcome for these variables may lie in the poor fits achieved by the reduced form equations for CHRO and LENGTH. With respect to the performance of the air pollution variables, however, the derivative of OZONEM is positive and significant at either the 5 percent level or the 10 percent level in all six equations. Additionally, the value of the derivative of OZONEM range from 25.72 to 33.14 indicating that the effects are not particularly sensitive to the choices made concerning which other pollution variables to include. Qualitatively at least, the results tend to corroborate the evidence presented in Appendix 1. That material reported some of the physiological effects of ozone exposure to be impaired pulmonary function and sensory irritation such as burning eyes or scratchy throat; i.e., exactly the type of health effects for which medical attention might be sought. The other pollution variables, however, did not perform nearly as well. In all six equations, no air quality derivative except the one for OZONEM was significantly different from zero at even the 10 percent level. In fact, the derivative of LEADSMLM was the wrong sign and insignificant. In the present context, however, that result is plausible since the health effects of lead poisoning, as indicated in Appendix 2, are of a long term nature and do not necessarily result in symptoms requiring immediate medical attention. Low-level lead poisoning, for example, may lead to progressive disorders of the cardiovascular, renal, or neurological systems, or simply to a general dullness or irritability. Finally, none of the possible synergistic effects examined had significant derivatives.

Among the socieoeconomic-demographic variables, AGE entered consistently with the wrong sign and its derivative never was significant. That result may seem somewhat surprising, however, in a sample composed only of employed workers, there may not be sufficient variation in years of age to capture effect of this variable on the health stock. Moreover, the relation between age and the health stock may be strongly non-linear, with the largest effects occurring among those at the highest ages; i.e., those not well represented in the sample considered here. EDUC, on the other hand, was always the correct sign (negative) implying those with more schooling are more efficient producers of health. The derivative of EDUC, however, never was significantly different from zero. The derivatives of SEX and RACE indicate that in the sample considered females and blacks tend to have lower health stocks than males and whites respectively. Both derivatives are consistently and significantly different from zero, with the derivative of SEX always being significant at the 1 percent level.

Table 4 provides parallel results for the 2197 observation sample, which as has been indicated repeatedly, suffered from a severe incidence of missing data on hourly wages. The sign pattern of the derivatives of CHRO and LENGTH reversed in the larger sample-as compared with the situation found in Table 3. Thus, the sign for CHRO reported in Table 4 is wrong and the sign for LENGTH is correct. Distressingly, the derivative of CHRO is highly significant. The air pollution variables, on the other hand, performed better in the equations shown in Table 4 than in those shown in Table 3. The derivative of OZONEM is always significant at the 1 percent level. In addition, the derivatives of the remaining five pollution variables always are of the correct sign and generally are significant at the 5 percent level. That statement applies in particular to the coefficient of LEADSMLM shown in column 6. Thus, these results suggest that the health stock responds negatively to increases in concentrations of air pollutants in addition to ozone and that the synergism discussed in Appendix 1 may be important to consider (see columns 4 and 5). The socieoeconomic-demographic variables performed similarly with the results reported in Table 3, although the significance levels for the derivatives of AGE, SEX, and RACE were considerably higher.

With some trepidation, the results from Tables 3 and 4 can be used to make some illustrative willingness to pay estimates for a reduction in ozone levels. These benefit estimates are offered advisedly largely because of two problems, discussed previously, in calculating $dB/d\alpha$: (1) the dependent variable in the equations estimated is only an indicator of whether medical care was received and does not measure the level of consumption of medical services and (2) the price of medical care employed in the calculations may tend to overstate the price actually faced by the individuals in the sample. In any case, since ozone was the only air quality variable that performed consistently well, in the 820 observation data set and since it is of primary interest in this study, only reductions in that pollutant are used in making the benefit calculations. Because St. Louis experiences only a comparatively small number of days a year when the hourly average exceeds the national primary and secondary standards it will not take a large reduction in ozone levels to meet the standard. Therefore, reductions in the ozone level of 10 percent, 15 percent, 20 percent, and 30 percent of the mean ozone level (.019) have been used to calculate the benefits.

Willingness to pay estimates are presented based on the two equations presented in columns 3 in both Tables 3 and 4. Those equations had neither the highest and lowest OZONEM derivatives; therefore, the estimates presented would typify the results obtained from the remaining equations. Since the MED variable reflects whether a doctor usually is seen at least once per year, the willingness to pay estimates also would be annual figures. Moreover, all of the willingness to pay estimates reported are computed for individual "average" employed worker. In other words, the means of all independent variables, provided in Table 2, were used in the empirical approximation to H_α/H_M, as was the mean the price of medical care variable, q_M. Table 5 presents the willingness to pay calculations.

TABLE 5

WILLINGNESS TO PAY FOR REDUCTIONS IN OZONE LEVELS

	PERCENT REDUCTION IN MEAN OZONE LEVELS			
EQUATION	10%	15%	20%	30%
3 (Table 3)	$2.42	$3.63	$4.84	$7.27
3 (Table 4)	$2.69	$4.04	$5.39	$8.08

The estimates of willingness to pay for a 10 percent (.0019 ppm) reduction in mean ozone levels range from $2.42 to $2.69. For a 30 percent (.0057 ppm) reduction in mean ozone levels the willingness to pay estimates range from $7.27 to $8.08. It should again be stressed that these estimates are in annual terms and pertain to the "average" worker in St. Louis. Although these willingness to pay figures may appear to be small, they still are larger than those found by Seskin in his study of photochemical oxidant levels and acute illness in the Washington, D.C. area. Chapter 2 reported Seskin's finding that a 55.6 percent reduction in maximum one-hour average 1973 oxidant levels, necessary to meet the 1971-78 national standard, would result in a savings to Group Health Association plan members of $4490 in that year. Since during 1973, there were roughly 100,000 GHA members, annual benefits per member would have been, approximately $.04. Seskin's benefit estimate rises somewhat for 1974. During that year a 42.9 percent reduction in photochemical oxidant levels would have been necessary to meet the national standard; a reduction that would have resulted in benefit for GHA members of $12,140 or about $.12 per member. As noted in Chapter 2, however, the theoretical basis of Seskin's estimates appears to be weak. As a consequence, an explicit comparison can be made between his estimates and those presented here only with great difficulty if at all. Nevertheless, the results are consistent with the prediction made by Harrington and Portney (1983), that the compensating variation approach should give higher willingness to pay values than the cost of illness approach.

There are at least two reasons, however, why the willingness to pay estimates presented in Table 5 are comparatively low. First, the reductions in ozone levels contemplated are not large. As indicated, the mean of the variable OZONEM was .019 ppm, reflecting the fact that ozone levels in St. Louis are much lower than those elsewhere in the U.S. In the Los Angeles area, for example, average ozone concentrations would exceed that figure by a factor of more than five and peak ozone concentrations would be as high as .35 ppm. In any event, for the case of St. Louis, a 10 percent or even a 30 percent reduction in ozone levels is not particularly large in absolute terms. Second, the benefit estimates account only for the affects of improvements in air quality on health. A total benefit estimate for each individual might also account, for example, for reduced materials damage and improved visibility.

5. Conclusion

In this chapter, the empirical application of the theoretical model was presented and estimates of willingness to pay for the "average" employed person were derived for specific reductions in air pollution. A method was derived whereby the health stock of an individual did not need to be quantified. This allowed for possibly more reliable results as no valid measure of health stock could be found. It appears from the estimations of the implicit health production function that ozone levels do play an important role in the production of health. No other pollutants were shown to have as large of effect on the production of health as ozone. For this reason only willingness to pay estimates were made for a reduction in ozone. Those willingness to pay estimates reflected an annual per employed person benefit to St. Louis residents of between $2.42 and $2.69 for a 10 percent reduction in mean ozone levels and between $7.27 and $8.08 for a 30 percent reduction in mean ozone levels.

FOOTNOTES

1. Recall that in the 2197 observation sample, measures of the hourly wage were not available for most respondents. To circumvent this problem in the reduced form regressions, the arithmetic mean of HWAGE was substituted in place of the missing and unknown values. A detailed evaluation of this procedure is provided in a series of four papers by Afifi and Elashoff (1966, 1967, 1969a, 1969b).

REFERENCES

Afifi, A. and R. Elashoff. "Missing Observations in Multivariate Statistics, Part I," Journal of American Statistical Association, 61 (1966), 595-604.

_____. "Missing Observations in Multivariate Statistics, Part II," Journal of the American Statistical Association, 62 (1967), 10-29.

_____. "Missing Observations in Multivariate Statistics, Part III," Journal of the American Statistical Association, 64 (1969), 337-58.

_____. "Missing Observations in Multivariate Statistics, Part IV," Journal of the American Statistical Association, 64 (1969), 359-65.

Atkinson, S. "A Comparative Analysis of Macro Epidemiological Studies," mimeo, University of Wyoming (December 1981).

Bates, D.V., D.M. Bell, C.D. Burnham, M. Hazucha, J. Mantha, L.D. Pengelly, and S.F. Silverman. "Short-term Effects of Ozone on the Lung, Journal of Applied Physiology, 32 (1972), 176-181.

Bauchinger, M., E. Schmid, and H.J. Einbrodt. "Chromosome Aberrations in Lymphocytes after Occupational Exposure to Lead and Cadmium," Mutations Research, 40 (1976), 57-62 .

Bauchinger, M., E. Schmid, and D. Schmidt. "Chromosome Analysis of Traffic Policeman with Elevated Lead Levels," Mutations Research, 16 (1972), 407-412.

Beattie, A.D., M.R. Moore, A. Goldberg, M.J.W. Finlayson, J.F. Graham, E.M. Mackie, J.C. Main, D.A. McLaren, R.M. Murdock, and G.T. Stewart. "Role of Chronic Low Level Lead Exposure in the Aetiology of Mental Retardation," The Lancet, 7907 (1975), 589-598.

Beritic, T. "Lead Concentration Found in Human Blood in Association with Lead Colic," Archives of Enivomnental Health, 23 (1971), 289-291.

Bhagia, G.S. and H. Stoevener, Impact of Air Pollution on the Consumption of Medical Services, Publication No. 600/5-78-002, Environmental Protection Agency, Corvallis (January 1978).

Burness, H.S., A.M. Church, R.G. Cummings, and A.F. Mehr. "Respiratory Health Effects from Air Pollution; An Overview," mimeo, University of New Mexico (December 1981).

Chisholm, J.J. "The Use of Chelating Agents in the Treatment of Acute and Chronic Lead Intoxication in Childhood," Journal of Pediatrics, 73 (1968) 1-38.

Cooper, B.S. and D.P. Rice. "The Economic Cost of Illness Revisited," Social Security Bulletin, 39 (1976), 21-36.

Cooper, W.C. and W.R. Gaffey. "Mortality of Lead Workers," Journal of Occupational Medicine, 17 (1975), 100.

Crocker, T.D., W.D. Schulze, S. Ben-David, and A.V. Kneese, Methods Development for Assessing Air Pollution Control Benefits, Vol. I, Experiments in the Economics of Epidemiology, Publication No. 600/5-79-001a, Environmental Protection Agency (February 1979).

Cropper, M.L. "Measuring the Benefits from Reduced Morbidity," American Economic Review, 71 (May 1981), 235-240.

David, O.J., J. Clark, and K. Voeller. "Lead and Hyperactivity," The Lancet, 2 (1972), 900-903.

David, O.J., B. McGann, J. Clark, S. Hoffman, and S. Sverd. "Low Lead Levels and Mental Retardation," The Lancet, 4 (1976), 1376-1379.

De la Burde, B. and M.S. Choate. "Does Asymptomatic Lead Exposure in Children Have Latent Sequelae?," Journal of Pediatrics, 81 (1972), 1088-1091.

De la Burde, B. and M.S. Choate. "Early Asymptomatic Lead Exposure and Development at School Age,' Journal of Pediatrics, 87 (1975), 638-664.

DeLucia A.J. and W.C. Adams. "Effects of O_3 Inhalation During Exercise on Pulmonary Function and Blood Biochemistry," Journal of Applied Physiology, 43 (1977), 75-81.

Dingwall-Fordyce, J. and R.E. Lane. "A Follow-Up Study of Lead Workers," British Journal of Industrial Medicine, 20 (1963), 313-315.

Durham, W.R. "Air Pollution and Student Health," Archives of Environmental Health, 28 (1974) 241-254.

Equitable Environmental Health, Inc. "Health Effects Planning Study Concerning Ozone and Other Photochemical Oxidants," Report No. HAOX-02, Rockville, M.D. (1977).

Fahim, M.S., Z. Fahim, and D.G. Hall. "Effects of Subtoxic Lead Levels on Pregnant Women in the State of Missouri," International Conference on Heavy Metals in the Environment. Toronto, Ontario. October 27-31, (1975).

Fetner, R. "Ozone-induced Chromosome Breakage in Human Cell Cultures," Nature, 194 (1965), 793-794.

Folinsbee, L.J., F. Silverman, and R.J. Shephard. "Exercise Responses Following Ozone Exposure," Journal of Applied Physiology 38 (1975), 996-1001.

Folinsbee, L.J., F. Silverman, and R.J. Shephard. "Decrease of Maximum Work Performance Following Exposure," Journal of Applied Physiology, 42 (1977), 531-536.

Folinsbee, L.J., S.M. Horvath, P.B. Raven, J.F. Bedi, A.R. Morton, B.L. Drinkwater, N.W. Boldvan, and J.A. Gilner. "Influence of Exercise and Heat Stress on Pulmonary Function During Ozone Exposure," Journal of Applied Physiology 43 (1977), 409-413.

Folinsbee, L.J., J.F. Bedi, and S.M. Horvath. "Respiratory Responses in Humans Repeatedly Exposed to Low Concentrations of Ozone," EPA-600/J-80-329, 1980.

Forni, A., G. Cambiaghi and G.C. Secchi. "Initial Occupational Exposure to Lead," Archives of Environmental Health, 31 (1976), 73-78.

Freeman, A. Myrick. The Benefits of Environmental Improvement: Theory and Practice, Johns Hopkins University Press, Baltimore (1979).

Gerking, S. and W. Schulze. "What Do We Know About Benefits of Reduced Mortality from Air Pollution?," American Economic Review, 71 (May 1981), 228-334.

Gershanik, J.J., G.G. Brooks, and J.A. Little. "Blood Lead Values in Pregnant Women and Their Offspring," American Obstetrics and Gynecology, 119 (1974), 508-511.

Granic, J.L., S. Sussa, S. Granick, R.D. Levere, and A. Kappas. "Studies in Lead Poisoning II. Correlation Between the Ratio of Activated to Inactivated Aminolevulinic Acid Dehydratase of Whole Blood and the Blood Lead Level," Biochemical Medicine, 8, 149-159.

Grossman, M. "The Correlation Between Health and Schooling," in N.E. Terleckyj, ed., Household Production and Consumption, National Bureau of Economic Research, Studies in Income and Wealth, Vol. 40, New York (1976).

Grossman, M. (1972a). "On the Concept of Health Capital and the Demand for Health," Journal of Political Economy, 80 (March 1972), 223-255.

Grossman, M. (1972b). The Demand for Health: A Theoretical and Empirical Investigation, National Bureau of Economic Research, Occasional Paper 119, New York (1972).

Hackney, J.D., W.S. Linn, J.G. Mohler, E.E. Pedersen, P. Breisacher, and A. Russo. "Experimental Studies on Human Health Effects of Air Pollutants II. Four-hour Exposure to Ozone Alone and in Combination with Other Pollutant Gases," Archives of Environmental Health, 30 (1975a), 379-384.

Hackney, J.D ., W.S. Linn, S.K. Karuza, R.D. Buckley, D.C. Law, D.V. Bates, M. Hazucha, L.D. Pengelly, and F. Silverman. "Effects of Ozone Exposure in Canadians and Southern Californians. Evidence for Adaption?" Archives of Environmental Health, 32 (1977), 110-116.

Hackney, J.D., W.S. Linn, J.G. Mohler, and C.R. Collier. "Adaption to Short-term Respiratory Effects of Ozone in Men Exposed Repeatedly," Journal of Applied Physiology, 43 (1977), 82-85.

Hallenbeck, W., W. Kojala, and R. Allen et al. "Health Effects of Ozone and Other Photochemical Oxidants in the Chicago Area," II NR-79/27, Chicago, IL, 1979.

Hammer, D.I., V. Hasselbald, B. Portnoy, and P.F. Wehne. "The Los Angeles Student Nurse Study: Daily Symptom Reporting and Photochemical Oxidants," Archives of Environmental Health. 28 (1974), 255-260.

Harrington, W. and P.R. Portney. "Valuing the Benefits of Improved Human Health," mimeo, Resources for the Future, Washington, D.C., 1983.

Hazucha, M. and D.V. Bates. "Combined Effect of Ozone and Sulfur Dioxide on Human Pulmonary Function," Nature, 257 (1975), 50-51.

Hernberg, S., J. Nikkanen, G. Mellin, and H. Lilius. "δ-aminolevulinic Acid Dehydrase as a Measure of Lead Exposure." Archives of Environmental Health, 21 (1970), 140-145.

Jaksch, J. and H. Stoevener. Outpatient Medical Costs Related to Air Pollution in Portland, Oregon Area, Publication No. 600/5-84-017, U.S. Environmental Protection Agency (July 1974).

Kagawa, J. and T. Toyama. "Effects of Ozone and Brief Exercise on Specific Airway Conductance in Man," Archives of Environmental Health, 30 (1975), 36-39.

Kerr, H.D., T.J. Kulle, M.L. McIhany, and F. Swiodersky. "Effects of Ozone on Pulmonary Function in Normal Subjects. An Environmental-Chamber Study," American Review of Respiratory Disease, 111 (1975), 763-773,.

Lancranjan, I., J.I. Popescu, O. Garanesco, I. Klepsch, and M. Serbanesu. "Reproductive Ability of Workmen Occupationally Exposed to Lead," Archives of Environmental Health, 30 (1975), 396-401.

Lave, L.B. and E.P. Seskin. "An Analysis of the Association Between U.S. Mortality and Air Pollution," Journal of the American Statistical Association, 68 (June 1973), 284-290.

Lave, L.B. and E.P. Seskin. Air Pollution and Human Health, Johns Hopkins University Press, Baltimore (1977).

Legge, T.M. "Industrial Lead Poisoning," Journal of Hygiene, 1 (1901), 96.

Lilis, R., A. Fischbein, S. Diamond, H.A. Anderson, F.J. Selikoff, W.E. Blumberg, and J. Eisinger. "Lead Effects Among Secondary Lead Smelter Workers with Blood Lead Levels Below 80 mg/100 ml," Archives of Environmental Health, 32 (1977), 256-226.

Merz, T., M.A. Bender, H.D. Kerr, and T.J. Kulle. "Observations of Aberrations in Chromosomes of Lymphocytes From Human Subjects Exposed to Ozone at a Concentration of 0.5 ppm for 6 and 10 Hr," Mutations Research, 31 (1975), 299-302.

Mohler, J.G., N.R. Johnson, and B.W. Armstrong. "Airway Resistance in Human Subjects Acutely Exposed to .6 to .8 ppm Ozone," Second International Clean Air Congress, International Union of Air Pollution Prevention Association, Washington, D.C., December 6-11, 1970, pp. 195-199.

Motley, H.L., R.H. Smart, and C.I. Leftwich. "Effect of Polluted Los Angeles Air (SMOG) on Lung Volume Measurements," Journal of the American Medical Association, 171 (1959), 1469-1477.

National Academy of Sciences, Ozone and Other Photochemical Oxidants. Committee on Medical and Biologic Effects of Environmental Pollution, Division of Medical Sciences, National Research Council, Washington, D.C. (1977).

National Research Council. Lead in the Environment. Committee on Lead in the Human Environment, Environmental Studies Board, Commission on Natural Resources. Washington, D.C. (1980).

Needleman, H.L., C. Gunnoe, A. Leviton, R. Reed, H. Peresie, C. Maher, and P. Barnett. "Deficits in Psychologic and Classroom Performance of Children with Elevated Dentine Lead Levels," New England Journal of Medicine. 300 (1979), 689-695.

Nelson, F. and L. Olson. "Specification and Estimation of a Simultaneous Equation Model with Limited Dependent Variables," International Eocnomic Review, 19 (1978), 695-710.

Nelson, W.C., M.H. Lykins, J. Mackey, V.A. Newill, J.F. Finklea, and D.J. Hammer. "Mortality Among Orchard Workers Exposed to Lead Arsenate Spray: A Cohort Study," Journal of Chronic Diseases, 26 (1973), 105-118.

Nordman, C.H., and S. Hernberg. "Blood Lead Levels and Erythrocyte δ-Aminolevulinic Acid Dehydrotase Activity of Selected Population Groups in Helinski," Scandanavian Journal of Work-Environment-Health, 1 (1975), 219-232.

O'Riordan, M.L., and H.J. Evans. "Absence of Significant Chromosome Damage in Males Occupationally Exposed to Lead," Nature, 247 (1974), 50-53.

Ostro, B. and R. Anderson. "The Effects of Air Pollution on Work Loss and Morbidity," Journal of Environmental Economics and Management, forthcoming (1983).

Pearlman, M.E., J.F. Finklea, C.M. Shy, J. VanBruggen, and V.A. Newill. "Chronic Oxidant Exposure and Epidemic Influenza," Environmental Research, 4 (1971), 129-140.

Perion, J. and C.B. Ernhart. "The Relation of Subclinical Lead Level to Cognitive and Sensorimotor Impairment in Black Preschoolers," Journal of Learning Disorders; 7 (1974), 26-30.

Perlroth, M.G., D.P. Tschudy, H.S. Marver, et al. "Acute Intermittent Porphyria: New Morphologic and Biochemical Findings," American Journal of Medicine, 41 (1966), 149-162.

Piomelli, S. and B. Davidow. "The FEP Concentration: A Promising Screening Test for Lead Poisoning," Pediatric Research, 6 (1972), 366.

Pueschel, S.M., L. Kopito and H. Schwachman. "A Screening and Follow Up Study of Children with an Increased Lead Burden," Journal of American Medical Association, 333 (1972), 462-466.

Roels, H., J.P. Buchel, R. Lauwerys, G. Hubermont, P. Bruaux, F. Clasys-Thoreau, A.L. Fontaine, and J. Van Overscheide. "Impact of Air Pollution by Lead on the Heme Biosynthetic Pathway in School-Age Children," Archives of Environmental Health, 31 (1976), 310-316.

Rosenzweig, M. and T.P. Schultz, (1982a). "The Behavior of Mothers as Inputs to Child Health: The Determinants of Birthweight, Gestation, and Rate of Fetal Growth," in V. Fuchs, ed., Economic Aspects of Health, National Bureau of Economic Research, Chicago (1982).

Rosenzweig, M. and T.P. Schultz, (1982b). "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs and Their Effects on Birthweight," mimeo (August 1982).

Sassa, S., J.L. Granick, S. Granick, A. Kappas and R.D. Levere. "Studies in Lead Poisoning. I. Microanalysis of Erythrocyte Protoporphyrin Levels by Spectro Fluorometry in the Detection of Chronic Lead Intoxication in the Subclinical Range," Biochemical Medicine, 8 (1973), 135-148.

Schoettlin, C.E., and E. Landau. "Air Pollution and Asthmatic Attacks in the Los Angeles Area," Public Health Reports, 76 (1961), 545-548.

Seskin, E. Air Pollution and Health in Washington, D.C.: Some Acute Health Effects of Air Pollution, National Bureau of Economic Research, Inc., Publication No. EPA/600/5-77/010, Washington, D.C. (July 1977).

Seskin, E., (1979a). "Pollution and Health in Washington, D.C.," Journal of Urban Economics, 6 (July 1979), 275-291.

Seskin, E., (1979b). An Analysis of Air Pollution and Its Health Effects: Washington, D.C., Metropolitan Area, Resources for the Future, Publication No. EPA/600/5-79/002, Washington, D.C. (February 1979).

Sterling, T.D., J.J. Phair, S.V. Pollack, D.A. Schurnsky, and I. DeGroot. "Urban Morbidity and Air Pollution: A First Report," Archives of Environmental Health, 13 (1966), 158-170.

Silver, W. and R. Rodriguez-Torres. "Electrocardiographic Studies in Children with Lead Poisoning," Pediatrics 41 (1968), 1124-1127.

Strothmann, J.A., et al. "Documentation of the Regional Air Pollution Study (RAPS) and Related Investigations in the St. Louis Air Quality Control Region," Research Triangel Park: Environmental Sciences Research Laboratory (December 1979).

U.S. Department of Health, Education, and Welfare, National Center for Health Statistics, "Prevalence of Selected Chronic Respiratory Conditions, United States - 1970," USGPO: Washington, D.C. (1973).

U.S. Environmental Protection Agency. Air Pollution: Air Quality Criteria for Photochemical Oxidants. Publication No. NATO/CCMS-29 Washington, D.C. (1974).

U.S. Environmental Protection Agency. Air Quality Criteria for Lead. Office of Research and Development. EPA-600/8-77-017. Washington, D.C.: U.S. Environmental Protection Agency (1977).

U.S. Environmental Protection Agency. Lead: Ambient Water Quality Criteria. Criteria and Standards Division, Office of Water Planning and Standards. Washington, D.C.: U.S. Environmental Protection Agency (1979).

Wada, O., K. Takea, Y. Varo, T. Ono, M. Nagahashi, and H. Seki. "δ-Aminolevulinic Acid Dehydratase in Low Level Lead Exposure," Archives of Environmental Health, 31 (1976), 211-215.

Wayne, W.S., P.F. Wehrle, and R.E. Carroll. "Oxidant Air Pollution and Athletic Performance," Journal of American Medical Association, 199 (1967), 901-904.

Williams, H.W., W.T. Caraway, and W.A. DeYoung. "Inactuation of Antibodies: A Causative Factor of Brain Pathology in Acute Lead Intoxication," Archives of Neurological Psychiatry, 72 (1954), 579-582,.

World Health Organization. Enviornmental Health Criteria, 3. Lead. Geneva: World Health Organization (1977).

Young, W.A., D.B. Shaw, and D.V. Bates. "Effect of Low Concentrations of Ozone on Pulmonary Function in Man," Journal of Appoied Physiology, 19 (1964), 765-768.

APPENDIX 1

HEALTH EFFECTS OF OZONE ON HUMANS

1. Introduction

Ozone is one of several photochemical oxidants that can be observed in the atmosphere. It is a secondary pollutant; that is, it is formed as a result of chemical reactions involving other pollutants in the atmosphere and sunlight. (Primary pollutants are those emitted directly by pollution sources.) The pollutants most responsible for the formation of ozone are the nitrogen oxides and hydrocarbons. The internal combustion engine is a major source of these pollutants, although there are other stationary sources such as electric power generating plants which contribute heavily to emission of nitrogen oxides.

The amount of ozone found in the atmosphere will depend upon the time of day, meteorologic conditions, and the amount of nitrogen oxides present. Early in the morning the concentration is low. As the day progresses, the amount of ozone in the air increases. The primary influence is that of sunlight intensity. Also, rush hour traffic will cause ozone levels to rise dramatically. The highest concentrations of ozone are usually found between the hours of 10 a.m. and 6 p.m.

The concentration of ozone in the atmosphere also will vary widely with location. Table 1 illustrates this diversity. During the years 1964-1967, St. Louis had a maximal hourly average of .35 ppm with a peak concentration of .85 ppm while Chicago had a maximal hourly average of only .13 ppm and a peak concentration only slightly higher at .19 ppm. Los Angeles had a higher maximal hourly average than St. Louis during this period of time, .58 ppm, but the peak concentration only reached .65 ppm (NAS 1977). Ozone concentrations will even vary widely within an SMSA. Table 2 presents a summary of ozone levels as measured by the RAMS network in the St. Louis SMSA during the years 1974 to 1977. RAMS station 15 exceeded a maximal hourly average of .01 ppm on 64.68 percent of the total days with available data and a maximal average of .08 ppm on 9.79 percent of the total days. RAMS station 17 exceeded a maximal hourly average of .01 ppm on 75.36 percent of the total days but on only 2.87 percent of the total days was a maximal hourly average of .08 ppm exceeded. This variation is due for the most part to the amount of traffic and the presence of stationary pollution sources in the area.

It has been estimated that many major U.S. cities will not attain the ambient air standard for ozone set by the EPA in 1979. The national primary and secondary standard for ozone is an hourly average of .12 ppm not to be exceeded more than once a year. Referring to Table 2 and

TABLE 1

SUMMARY OF MAXIMAL OXIDANT CONCENTRATIONS IN SELECTED CITIES, 1964-1967[a]

Station	Total Days With Available Data	Total Days with Maximal Hourly Average ≥ Concentration Specified						Maximal Hourly Avg. (ppm)	Peak Concentration (ppm)	Yearly Average (ppm)
		.05 ppm		.10 ppm		.15 ppm				
		No. of Days	% of Days	No. of Days	% of Days	No. of Days	% of Days			
Los Angeles	730	540	74.0	354	48.5	220	30.1	.58	.65	.030
Denver	285	226	79.3	51	17.9	14	4.9	.25	.31	.036
St. Louis, MO	582	362	62.2	59	10.1	14	2.4	.35	.85	,031
Philadelphia	556	233	41.9	60	10.9	13	2.3	.21	.25	.026
Cincinnati	613	319	52.0	55	9.0	IO	1.6	.26	.32	,030
Washington, D.C.	577	313	54.2	65	11.3	7	1.2	.21	.24	.029
Chicago	530	269	50.8	24	4.5	0	0	.13	.19	.028

[a] From NAS 1977 (derived from U.S. DHEW)

TABLE 2

SUMMARY OF OXIDANT CONCENTRATION IN ST. LOUIS, 1974-1977 (RAMS AIR POLLUTION DATA)

Station	Total Days With Available Data	Total Days With Maximal Hourly Average > Concentration Specified									
		.01 ppm		.05 ppm		.08 ppm		.12 ppm		.16 ppm	
		No. of Days	% of Days	No. of Days	% of Days	No. of Days	% of Days	No. of Days	% of Days	No. of Days	% of Days
1	420	303	72.14	132	31.43	46	10.95	9	2.14	2	.5
2	438	338	77.17	129	29.45	40	9.13	10	2.28	1	.2
3	428	283	66.12	104	24.30	26	6.07	5	1.17	2	.47
4	424	279	65.80	78	18.40	22	5.19	4	.94	0	0
5	431	295	68.45	102	23.67	28	6.50	4	.93	1	.23
6	421	303	71.97	129	30.64	39	9.26	10	2.38	5	1.19
7	434	288	66.36	87	20.05	27	6.22	3	.69	2	.46
8	433	323	74.60	124	28.64	39	6.24	10	.69	2	.116
9	420	300	71.43	178	42.38	65	15.48	12	2.86	1	.48
10	427	328	78.10	150	35.13	42	9.84	9	2.10	0	0
11	428	323	75.47	123	28.74	39	9.11	10	2.34	2	.47
12	428	316	73.83	112	26.17	32	7.48	10	2.34	3	.70
13	432	324	75.00	121	28.01	37	8.56	7	1.62	3	.69
14	421	316	75.06	158	37.53	52	12.35	17	4.04	5	1.18
15	419	271	64.68	113	26.97	41	9.79	10	2.39	2	.46
17	418	315	75.36	108	25.84	12	2.87	0	0	0	0
19	422	301	71.33	147	34.83	53	12.56	13	3.08	4	.95
20	402	270	67.16	108	26.87	34	8.46	10	2.49	0	0
21	422	319	75.59	154	36.49	54	12.80	16	3.79	6	1.42
Average % of Days St, Louis Exceeded Stated Concentration			71.87		29.24		8.89		2.01		.52

Average Ozone Level (1974-1977) .019

averaging across stations, St. Louis exceeded an hourly average of .12 ppm on 2.01 percent of the days with available data. This translates into nearly eight days a year where an hourly average of .12 ppm was exceeded; thus, during the years 1974-1977, St. Louis was in violation of the present air quality standards for ozone.

It is interesting to note, however, the decline in ozone levels that St. Louis experienced from 1967 to 1975. From 1964 to 1967, the maximal hourly average exceeded .05 ppm on 62.2 percent of the total days while from 1974-1977 this happened on only 29.24 percent of the total days. The average ozone level declined from .031 ppm during 1964 to 1967 to .019 ppm during 1974 to 1977.

2. Health Effects on Humans

Two types of studies have been conducted to determine the health effects of ozone on humans. These are controlled experimental studies under laboratory conditions and epidemiological investigations of ozone-exposed groups. From these studies, ozone has been found to produce adverse effects on pulmonary function, numerous discomforts (including headaches, irritation and soreness of the throat, nose, mouth, and trachea, substernal discomfort, coughing, wheezing, and malaise), changes in blood chemistry, and possibly chromosomal aberrations.

A. Pulmonary Function and Sensory Irritation

1) Controlled experimental studies

Controlled experimental studies of the health effects of ozone on humans offer several advantages (Hallenbeck et al. 1979). These experiments can be carried out under specific conditions. That is, many factors such as temperature, humidity, health status of the participants, duration of exposure, and level of exercise can be controlled for and thus, the effects that are produced can be directly related to the pollutant, ozone. Also, there are several pulmonary tests that can be carried out in the laboratory so that specific effects of ozone on the lungs can be identified.

Ozone concentrations of 4-30 ppm have been found to cause death in laboratory animals after exposure at these levels for 3-4 hours. Industrial exposure at 9, ppm has produced severe pneumonia in workers (Young et al. 1964). Ozone levels below 1 ppm are the most interesting to study though because they are known to occur in industries, smog, and in the cabins of modern aircraft. Pulmonary function and sensory irritation have been studied most frequently for levels of ozone exceeding .30 ppm. Very few studies have been conducted for ozone levels of .20 ppm, .10 ppm, or below. Those studies that have been carried out at these levels have found no conclusive evidence of adverse effects (Equitable Environmental Health, Inc. 1977). Referring back to Table 2, St. Louis never experienced an hourly average ozone level of .30 ppm during the years 1974 to 1977 and

an hourly average of .16 ppm was exceeded on only .52 percent of the total days monitored.

Decrements in pulmonary function have been reported for ozone levels greater than .30 ppm. Along with these changes in pulmonary function many symptoms of sensory irritation have been reported. Below is a review of several of these studies.

Bates (1972) exposed 10 healthy subjects to .75 ppm of ozone for two hours while the subjects were resting. Symptoms reported by the subjects included substernal soreness and cough. A few complained of symptoms of dyspnea (difficulty in breathing) and pharyngitis (inflammation of the pharynx). The pulmonary tests showed decrements in vital capacity (maximal expiration following maximal inspiration), increased airway resistance, and decreased dynamic compliance of the lungs (greater stiffness of the lungs). The researchers concluded that .75 ppm of ozone, in absence of other pollutants, is too high a level for a two hour exposure of the general population. Folinsbee et al. (1975) reported a decrease in maximum oxygen uptake (absorption) following exposure to ozone at .75 ppm with intermittent exercise. Because of this decrease in maximum oxygen uptake, maximum work performance was reduced 11 percent. Kagawa and Toyama (1975) found that peak concentration of .90 ppm of ozone for 5 minutes produced adverse effects when the subject undertakes exercise in it. They found a highly significant decrease in airway conductance in the four subjects who were exposed though none of the subjects complained of any symptoms.

In an earlier study, before researchers used environmental chambers to simulate ambient air with ozone pollution, eleven subjects breathed .60-.80 ppm of ozone through a mouthpiece for two hours (Mohler et al. 1970). This level of exposure produced decrements in vital capacity, forced expiratory volume, and other pulmonary functions. Subjects complained of substernal soreness and tracheal irritation for 6 to 12 hours after the exposure. These symptoms were accompanied by a dry cough. The effects on pulmonary function and the symptoms experienced could be partially due to the way the subjects were exposed, i.e. through a mouthpiece.

Other studies have produced similar results to those studies mentioned (see EPA 1974). It appears that there are definite physiological changes that take place in the lungs and these are accompanied by various symptoms when subjects are exposed to .75 ppm for two or more hours.

Experimental exposures at .50 ppm have generally had similar effects on pulmonary function as exposure at .75 ppm. Hackney et al. (1975a) exposed 2 groups of 4 subjects each to ozone levels of .50 ppm for 4 hours. Subjects in Group A were healthy with no history of pulmonary problems while subjects in Group B had a history of hyperreactive airways (they had frequently suffered from allergies, smog sensitivity, or asthma). Group A showed few or no effects; they complained only of mild pharyngitis and substernal discomfort. Group B, however, developed marked respiratory symptoms and physiological changes. Pulmonary tests revealed decreases in lung capacity, forced expiratory flow rates and volume and forced vital capacity, reductions in inspiratory capacity due to substernal pain, and

increased airway resistance. Symptoms in Group B included substernal pain, wheezing, malaise, and cough. The authors concluded that subjects with a history of pulmonary hyperactivity were significantly affected while healthy subjects were minimally affected by exposure at .50 ppm. However, subsequent exposure of a second group of 7 healthy subjects at .50 ppm ozone 2 hours a day for 2 consecutive days with intermittent exercise resulted in definite decreases in pulmonary function (1975b). Subjects showed symptoms of cough. substernal discomfort, and malaise. The effects were greatest on the second consecutive day of exposure. This finding is consistent with the results of other studies that have found effects of ozone on pulmonary function to be more severe during exercise. In these same experiments, some effects were found in the same group at exposure levels of .37 ppm but not after exposure at .25 ppm. There were definite decreases in pulmonary function and various symptoms were observed at .37 ppm.

In a study by Folinsbee et al. (1977), 14 nonsmoking males were exposed to .50 ppm ozone for 2 hours under different environmental conditions. The subjects were divided into 2 groups; one that performed from 60 to 90 minutes of exercise and one that performed 30 to 60 minutes. The temperature of the laboratory chamber was also varied. The authors found that pulmonary function changed the greatest amount immediately after exercise instead of at the end of the exposure. The decrease in forced volume capacity and forced expiratory volume following exercise was twice as large as the decrements at the end of exposure. The greatest decrease occurred when heat and ozone were combined. The symptoms observed in the subjects were chest discomfort and difficulty in taking a breath. The authors concluded that effects of ozone are most severe immediately after exercise, and heat stress may modify the overall effect of ozone on pulmonary function.

Several other studies have also found pulmonary function is affected at .50 ppm ozone. Kerr et al. (1975) found significant changes from control values of several pulmonary tests (airway conductance, pulmonary resistance, and forced expiratory volume) when 20 subjects were exposed to .50 ppm ozone for 6 hours with two intermittent exercise sessions of 15 minutes each. Folinsbee et al. (1975) studied the response of 28 subjects after ozone exposure of .50 ppm for 2 hours while exercising intermittently at 45, 60, and 75 percent of maximum aerobic power. The major response noted was an increase in respiratory frequency during exercise following ozone exposure. It was concluded that through its irritant properties ozone modifies the normal ventilatory response to exercise.

There is also evidence that ozone can produce adverse effects on pulmonary function and discomfort at concentrations below .50 ppm. One study by Hackney et al. (1975b) that examined effects of ozone exposure of .37 ppm has already been noted. The authors found definite decreases in pulmonary function, and the subjects experienced various discomforts. Hazucha and Bates (1975) reported a marked decrease in pulmonary function among healthy subjects performing light exercise while exposed to .37 ppm of ozone and .37 ppm sulfur dioxide for 2 hours. Throat irritation, coughing, and chest pain were also observed. These effects persisted

several hours after exposure. Folinsbee et al. (1975) noted an increase in respiratory frequency when subjects were exposed to .37 ppm ozone for 2 hours while exercising intermittently.

Effects were studied by DeLucia and Adams (1977) on 6 males after exposure for 1 hour to .15 ppm or .30 ppm ozone at rest or performing exercise at 25, 45, or 65 percent of maximum oxygen uptake. Ventilation volume and maximum oxygen uptake were unaffected by the most severe exposures and exercise. Host subjects did however demonstrate signs of toxicity (headaches, congestion, and wheezing) during the most stressful protocols. Two of the most sensitive individuals were unable to complete 1 hour of 65% maximum oxygen uptake exercise while breathing .30 ppm ozone.

In most of these studies, an association between symptoms and changes in lung function was usually found. Hypersusceptibles do not develop symptoms upon 2 hour exposure to ozone at .20-.25 ppm. However, symptoms develop in normal as well as hypersusceptibles after exposure for 2 or more hours at levels greater than .37 ppm with or without exercise. Exposure to .15 or .30 ppm ozone for one hour, accompanied by stressful exercise was sufficient to produce discomfort in even normal subjects. The symptoms that subjects experienced most frequently were substernal pain, coughing and wheezing, chest tightness, sore throat, and tracheal irritation. At lower levels of exposure, headaches were frequently complained of. Some hypersusceptible subjects were unable to perform normal tasks or complete the study when exposed to levels from .37 ppm ozone while exercising.

The experimental studies described above indicate that significant adverse effects on pulmonary function occur in humans at ozone concentrations as low as .37 ppm. There is evidence that hypersusceptibles (those with a history of respiratory problems) may develop symptoms at levels lower than this. In many of these studies, subjects engaged in some form of exercise to test the response of ozone on exercising individuals. These studies concluded that the effects of ozone concentrations at any level are much more pronounced when an individual is engaging in "heavy work." The effects of ozone are thus related to the volume of ozone breathed per unit. The more air an individual is breathing in the higher the actual exposure to ozone will be.

Several researchers have reported an apparent adaptation to pulmonary effects of repeated or chronic exposure to ozone. In a recent study by Hackney et al. (1977) 6 subjects were exposed to .50 ppm ozone 2 hours a day for 4 successive days under conditions simulating ambient pollution exposures. Five of these subjects showed pulmonary function decrement on exposure day 1 to day 3. This decrement was largely reversed on day 4. These results suggest some people do not continue to experience the same decrements in lung function after repeated exposures to ozone.

Another study by Hackney et al. (1977) compared the responses of 4 Canadians (whose previous total ambient ozone exposure was low) and those of 4 southern Californians to ozone exposures of .37 ppm for two hours with light exercise. The Canadians, on average, showed greater clinical and physiological reactivity to exposure. The explanation that seemed most

plausible to the investigators was that southern Californians have adapted to chronic ambient ozone exposure.

The purpose of a study by Folinsbee et al. (1980) was to determine whether there was any cumulative effects of repeated exposure to ozone and if adaptation to repeated exposures occurred. Three groups of subjects were used, each exposed to a different concentration of ozone: group 1 (n=10) was exposed to .20 ppm ozone, group 2 (n = 10) was exposed to .35 ppm ozone, and group 3 (n=8) was exposed to .50 ppm ozone. Subjects were exposed for 2 hours to filtered air on day 1 and day 5 of the experiment and to the ozone air for 2 hours on day 2 through day 4. No acute or cumulative effects were seen for those repeatedly exposed to .20 ppm ozone. There were decrements in pulmonary function for those exposed to .35 ppm on day 2 and 3, the greatest effect being seen day 3, but the effects were absent on days 4 and 5. The greatest change in pulmonary function came from the group repeatedly exposed to .50 ppm ozone. The most significant effects occurred day 3. There were significant effects day 4 but they were of a lesser magnitude. The authors concluded that there are some short term cumulative effects. These are followed by a period of resistance or adaption in which there is a marked lessening of the effects on pulmonary function and of subjective symptoms.

Synergistic effects of ozone and other air pollutants have been reported. In a study by Hazucha and Bates (1975), healthy young subjects were continuously exposed over a 2 hour period to a mixture of .37 ppm ozone and .37 ppm sulfur dioxide. They found that the combination of these two pollutants had a much greater effect on pulmonary function than either had individually. Greater decrements in pulmonary tests were observed following the combined exposure than was observed following exposure to each gas alone. However, a study that was designed to confirm this conclusion, using the same levels of ozone and sulfur dioxide found a substantially smaller decrement in pulmonary function than what had been previously reported.

Hackney et al. (1975a) exposed subjects to a mixture of ozone and nitrogen dioxide and ozone, nitrogen dioxide, and carbon monoxide. Ozone concentrations measured .25 ppm and .50 ppm. Added to these concentrations was .30 ppm nitrogen dioxide or .30 ppm nitrogen dioxide and 30 ppm carbon monoxide. The authors concluded that no additional effects were produced nor were effects more severe due to the increase of nitrogen dioxide and carbon monoxide. However, no conclusions about synergistic effects can be drawn from this study because of the possibility of these effects being present at different pollutant levels or environmental conditions.

2) Epidemiological Studies

A variety of epidemiological studies attempting to associate various characteristics of human health with daily concentrations of ozone have been carried out in the last two decades, mainly in the Los Angeles area. Many of the epidemiological studies of the health effects of ozone on human beings suffer from serious defects. Among these are lack of information about subjects studied, subjectivity of response due to knowledge of the

level of exposure, lack of analysis of the threshhold of the response, incomplete separation of the effects on health of competing meteorological and pollutant variables, and lack of appropriate weighting of time spent indoors and outdoors and the associated differences in pollutant exposure Hallenbeck 1979). Within these limits, some definite conclusions can be drawn as to the association of some health indicators with levels of ozone,

A relationship between mortality and concentrations of ozone has not been clearly demonstrated. Two studies relating the effects of ozone on the mortality of the elderly in Los Angeles County have been reported by the California State Health Department (EPA 1974). A marked increase in mortality among the elderly was found during a period of high photcchemical-oxidant concentrations during a 2 week period. However, there was also a sharp increase in temperature during this period that contributed to the increase in mortality. When mortality rates during high oxidant concentrations accompanied by low temperatures were examined, no relationship between oxidant levels and mortality could be discerned. The other study compared daily mortality and transfer to a hospital among nursing home residents to daily temperatures and the occurrence of smog-alert days with ozone levels of .30 ppm or higher. No correlation between mortality and smog-alert days in the absence of high temperatures could be found.

Sterling et al. (1966) studied the influence of ozone on the number of people admitted to hospitals for a 7 month period in 1961. They grouped diseases into "highly relevant," "relevant", and "irrelevant" categories. Highly relevant diseases were allergic disorders, inflammatory diseases of the eye, acute upper respiratory infections, influenza, and bronchitis. "Relevant" disorders included disease of the heart, rheumatic fever, vascular diseases, and all other respiratory diseases. All other disorders were considered irrelevant. Correlation coefficients between admissions for "highly relevant" diseases and nine pollution indices measured were all statistically significant at the .05 level. Correlations between "relevant" diseases and pollution were significant for oxidant, ozone, and sulfur dioxide. For all pollutants except sulfur dioxide, correlations between irrelevant diseases and pollution were negative cr not significant. This study must be interpreted carefully; the relationships, though significant, were minimal. Confirmation of this study over a longer period is needed.

Several studies have also found some evidence that ozone may aggravate existing respiratory conditions. Schoettlin and Landau (1956) studied 137 patients being treated for asthma in the Pasadena area. No correlation between air pollution of any kind and asthmatic attacks was significant. However, a significantly greater number of asthma attacks occurred on days when ozone levels were greater than .25 ppm than on days when they were less than .25 ppm. This result may suggest that there is a threshold level of ozone above which an increase in asthmatic attacks may occur.

Studies were done by Motley et al. (1959) on 66 volunteers, 46 of whom had pulmonary emphysema. Oxidants were removed from the air of the laboratory room by activated charcoal filters. Pulmonary functions

improved in the volunteers with emphysema if they entered the room on a "smoggy" day and remained there for 40 hours or more. No change in pulmonary function was seen for normal subjects or patients with emphysema who entered the room on a "nonsmoggy" day. The smoking habits of subjects were not accounted for nor were the effects of individual pollutants separated from one another so no real conclusions can be drawn.

Other studies have suggested an association between exposure to ozone and both increased frequency of asthma attacks and decrements in pulmonary function (EPA 1974). However all of these studies have serious limitations. Variations in smoking habits haven't been taken into account nor have the effects of other pollutants either singly or in combination with ozone been controlled for.

Studies have also been done to examine the effects ozone may have on the healthy population. Hammer et al. (1974) studied the effects of ozone levels on student nurses in Los Angeles for the period October 1961 to June 1964. Diaries were kept by student nurses on any discomfort or illness they had during each day of the study period. Headache frequency rose slightly at and above ozone levels of .25-.29 ppm. Rates for eye discomfort began to increase at .15-.19 ppm; rates for cough and chest discomfort showed an increase at .30-.39 ppm. Symptom frequencies in this study were more closely related to photochemical oxidants than carbon monoxide, nitrogen dioxide, or daily minimum temperatures.

Effects of ozone on the health of college students in the Los Angeles and San Francisco Bay area were examined by Durham (1974). Health data were collected from 1970-71. Each time a student visited the health service for a new illness, a sheet requesting demographic information, smoking habits, and physicians' diagnosis was completed. Coefficients of correlation between levels of pollution and weather variables on day 1 and proportions of new illness on the same day and on days up through day 7 were made. The illnesses most strongly associated with pollution were, in descending order, pharyngitis, bronchitis, tonsillitis, colds and sore throat. The pollutants most closely correlated with illness were in descending order, peak oxidant, mean sulfur dioxide, mean nitrogen dioxide, and mean nitric oxide. Stronger associations between pollution and illness were found in the Los Angeles area. Correlations of pollution with bronchitis were greatest when lagged 5 to 6 days; correlations of pollution with combined respiratory disease were greatest when lagged zero to 3 days; and correlations of pollution with asthma, eye irritation, headache, and hay fever were greatest when lagged zero to 1 day. This study strongly suggests a relationship between ozone levels and specific types of illnesses.

Pearlman et al. (1971) studied the effects of oxidants upon epidemic influenza incidence and duration of the illness. During 1968-69, morbidity rates of 3500 children from 5 southern Californian communities representing a definite gradient in chronic ozone exposure were studied to see if there was a relationship between chronic ozone exposure and epidemics among school children. Information from several sources (absenteeism from school

and questionnaries filled out by parents) revealed no statistically significant morbidity differences corresponding to ozone exposure.

Effects of Los Angeles oxidizing type air pollution on athletic performance was examined by Wayne et al. (1967) in 21 competitive meets of high school cross country track runners from 1959-64. It was hypothesized that running times should improve as the season progresses. Team performance was evaluated by the percentage of boys who didn't improve over their last meet. The highest correlation to team performance was that of oxidant levels in the hour before the race. Carbon monoxide levels, temperature, or humidity showed no relationship with team performance. The authors asserted that the decrease in performance may be directly physiological or may be due to decreased motivation caused by discomfort It is noted that athletes often complain of chest discomfort after competing in Los Angeles. Whether or not the decrement in team performance is due to physiological changes or decreased motivation, the study still indicates that ozone had an effect on team performance.

Though no conclusive evidence can be drawn from these studies, there is still an indication that low levels of ozone result in adverse human responses. Many participants suffered from discomfort when the concentration of ozone in the atmosphere was as low as .15 ppm. Symptoms reported at this exposure included eye irritation, sore throat, shortness of breath, cough, headache, watering of eyes, and hoarseness. Asthmatic attacks increased when ozone concentrations were above .25 ppm. At .3-.39 ppm, chest discomfort was reported by participants in these studies. Therefore, it seems safe to say that members of the general population suffer discomfort at ozone levels that are found in urban areas.

B. Effects on Blood Chemistry

At almost every level of ozone above .2-.25 ppm, small but significant changes in human blood biochemistry are seen. The changes include increases in red blood cell membrane fragility and serum E levels, decreases in hemoglobin concentrations, and alterations in the activities of several enzymes in the blood. Although there is a wide spectrum of ozone interference on blood biochemistry, the physiological Significance is unknown.

C. Mutagenesis

Ozone may potentiate chromosomal aberrations. Several studies have been done to evaluate these possible effects. Merz et al. (1975) studied the effects of exposure to .50 ppm. ozone for 6 to 10 hours on chromosomal aberrations. Ho true chromosome type aberrations were found. However, chromatid deletions and single strand breaks increased after exposure. The authors reported an increased frequency in these abnormalities 2 weeks after exposure with a return to normal frequency 6 weeks after exposure.

Another study (Fetner 1965) found chromatid deletions were produced as an exponential function, of the exposure to ozone. The authors concluded

that exposure to ozone is capable of producing chromatid breakages in human cell cultures.

3. Conclusion

Conclusive evidence exists that high ozone levels are detrimental to human health. Very pronounced physiological effects of ozone on pulmonary function have been found for ozone levels of .50 ppm or higher. Less pronounced effects have occurred in studies where the level of ozone administered to subjects was .37 ppm with these effects becoming more pronounced when subjects were engaged in heavy work. Sensory irritation and discomfort have been reported at ozone levels as low as .15 ppm, a level not uncommon in many urban areas.

The next chapter summarizes the relationships found by researchers between levels of lead in the body and the resulting physiological effects. WHile it is a fairly direct procedure to find a relationship between the level of ozone in the air and its physiological effects, it is much more difficult to compare actual lead concentrations in the ambient air to the effects this concentration may have on health. Lead occurs in media other than the ambient air, and the contributions each media has on an individual's total exposure to and absorption of lead is not clearly known. Therefore, it will be an impossible task to relate concentration so flead only in the ambient air to specific physiological effects; only the total lead concentration found in the subjects' bodies can be used to find a relationship between lead levels and health.

APPENDIX 2

HUMAN HEALTH EFFECTS OF LEAD

1. Introduction

Every member of the general population is exposed to elevated levels of lead from numerous sources. It is found in the air, water, foods, tobacco, soil, dust, and other items. The exposure to lead will depend upon both how much is found in these different medium and the amount actually inhaled or ingested. Some groups of people may be subject to much higher exposures than others. For example, children may eat soil or dust that contain lead deposited from the air or paint on walls. Also, some workers are exposed to large amounts of lead due to the type of occupation they are in.

Exposure to lead for the general adult population may be summarized as follows. Average inhalation from the air is about 15 µg/day in urban areas. Inhalation may be as low as 1 µg/day in rural areas to more than 100 µg/day at some urban sites. Average ingestion from diet is about 200 to 300 µg/day, with a range anywhere from 100 µg to more than 2000 µg/day. Ingestion from water ranges from 1 to more than 500 µg/day. About 10 µg of lead per pack of cigarettes smoke is inhaled. Ingestion of lead from dust, soil, or paint is negligible in adults; however, ingestion may be much higher in children.

It is clear from this data that diet is the major pathway of exposure. However, no conclusions about actual body burden can be made until it is known how much of this inhaled or ingested lead is absorbed. Also, it is clear that typical exposure can include a wide range of values. Urban residents or smokers will be exposed to more lead than rural residents or nonsmokers. Other characteristics of individuals such as age, sex, and socioeconomic status will affect the degree of lead exposure and the amount absorbed.

The relationship between exposure to lead and the level of lead in the body has not been clearly defined. The EPA (1977) estimated that of the lead in the air inhaled, between 20 and 40 percent of it would be deposited in the lungs. It is not known how much of the deposited lead is absorbed into the body. From controlled studies, it has been estimated by the EPA (1977) that for every 1 µg/m in the air, the blood lead level in an individual increases by approximately 1 to 2 µg/100 ml. The amount absorbed from the gastrointestinal tract is about 8 to 10% of the amount ingested. The EPA estimated that for every 100 µg/day of oral intake of lead an increase in blood lead levels of between 4.4 to 18.3 µg/100 ml takes place. The amount of lead absorbed into the body from inhaled

cigarette smoke is not known though it is assumed to be about the same amount as that inhaled from the air.

The relative contributions of each medium to the total lead body burden is also important in considering lead exposure. The contributions will vary widely in accordance to the residence of an individual and the characteristics of that individual. For example, an urban smoker will get a much higher percentage of lead from the air and tobacco smoke while a rural nonsmoker will get very little from these sources. In both cases, diet will probably contribute over 50 percent of the total lead absorbed, but for the rural nonsmoker the contribution will probably be about 90 percent while the contribution to the urban smoker will be about 50 percent. Air and tobacco smoke may account for approximately 15 percent each of the lead absorbed into the body of the urban smoker depending of course on the concentration of lead in the air and the number of cigarettes smoked. For the rural nonsmoker, the contribution of these two medium to lead in the body is negligible.

It is difficult to estimate a typical blood lead level for the general population. The differences in estimated blood lead levels are influenced by the degree and type of exposure, the characteristics of the individuals themselves, and the method of sampling and analytical techniques used in the estimation. The average blood lead level of urban adults has been estimated to be between 20 and 25 µg/100 ml by many researchers. Other studies have found both lower and higher blood lead levels in urban adults. The range of values of blood lead levels for the general population is probably from 10 µg/100 ml to 30 µg/100 ml.

A. Ambient Air Exposures of Lead in the Recent Past

Since 1957, samples of particulate matter collected at many urban and nonurban National Air Surveillance Network (NASN) sites have been analyzed for lead. However, only since 1966 have reliable procedures for measuring lead levels been used; therefore, only this data has been thoroughly analyzed. This data from NASN was studied for trends over the ten-year preiod from 1965 through 1974. It was found that urban lead concentrations increased from 1965 to 1971 and then declined. The principle reason for the decline was the introduction of automobile engines around 1970 that used gasoline with lower lead content. Practically all cars built after 1970 were able to use regular gasoline instead of the more leaded premium fuels. Even lower levels of lead in the air are probably now present because of the large use of unleaded gasolines. This decline will continue as more people switch from cars that use regular gasoline to cars that use unleaded gasoline.

In Table 1, lead concentrations and particle size are presented for the four quarters of 1970 for six urban areas. The average annual total lead concentration ranged from a high quarter of 3.2 µg/m^3 in Chicago to a low quarter of 1.3 µg/m^3 in Washington, D.C. St. Louis experienced a high quarter of 1.9 µg/m^3 and a low of 1.6 µg/m^3. This table indicates tghat many U.S. urban areas in 1970 were exceeding the EPA standard of 1.5 µg/m^3 maximum arithmetic mean average over a calendar quarter. A

TABLE 1

QUARTERLY AND ANNUAL SIZE DISTRIBUTIONS OF LEAD-BEARING PARTICLES FOR SIX CITIES, 1970[a]

City and Quarter of Year	Average Concentration $\mu g/m^3$	Average Mass Media Diameter μ_m	Percentage of Particles $\leq \mu_m$
Chicago, IL			
1	2.4	1.43	41
2	3.5	.51	65
3	3.5	.56	65
4	2.9	.54	64
Total Year	3.2	.68	59
Cincinnati, OH			
1	1.0	.25	79
2	2.2	.41	74
3	1.9	.54	69
4	2.1	.65	67
Total Year	1.8	.48	72
Denver, CO			
1	2.0	.43	76
2	1.1	.58	68
3	1.4	.52	69
4	3.0	.56	66
Total Year	1.8	.50	70
Philadelphia, PA			
1	1.5	.36	74
2	1.2	.38	74
3	1.8	.70	62
4	1.9	.45	70
Total Year	1.6	.47	70
St. Louis, MO			
1	1.9	.46	68
2	1.6	.63	53
3	1.8	.78	59
4	1.8	.95	53
Total Year	1.8	.69	62
Washington, D.C.			
1	1.3	.36	76
2	1.0	.39	73
3	1.3	.41	74
4	1.8	.54	71
Total Year	1.3	.42	74

Source: EPA (1977)

large number of southern California cities had annual average concentrations of 3.0 $\mu g/m^3$ or greater during this period. This was probably due to the heavy automobile traffic in these areas and also the topography and meteorological condition sthat favor retention of pollutants.

The average max median diameter for these six cities ranged from .69 mm in St. Louis to .42 mm in Washington, D.C. Forty-nine to 74 percent of the lead was associated with particles less than 1 mm in diameter. The smaller particles can become embedded in the lungs and therefore cause more health problems than the larger particles.

Table 2 presents a summary of lead concentrations in the St. Louis metropolitan area for the five quarters in the period January , 1976 to March, 1977. When these figures are compared with those in Table 1, the amount of lead present in the ambient air appears to have decline significantly since 1970. In fact, the maximum arithmetic mean over one quarter in 1976 is 1.040 $\mu g/m^3$ compared to 1.9 $\mu g/m^3$ in 1970. A large part of this reduction is probably due to the increased use of unleaded gasoline in automobiles. From this data, it can be concluded that St. Louis was meeting the EPA standards for lead in the ambient air in the period when the RAMS data was gathered. The level of lead in the ambient air declined approximately 100 percent over the period 1970 to 1976 therefore bringing St. Louis in compliance with the primary and secondary EPA standards.

It is difficult to know the contribution this level of lead in the ambient air has on the total lead burden in an individual. As mentioned in the previous section, the EPA (1977) has estimated that for every 1 $\mu g/m^3$ in the air, the blood level increases by approximately 1 to 2 $\mu g/100$ ml. If this is true, then the amount of lead in the air in St. Louis is probably contributing little to the total level of lead in the bodies of those living in the St. Louis SMSA. The health effects derived from airborne lead as compared to the lead in other media may actually be quite small.

2. Health Effects of Lead

The toxicity of lead has been recognized since ancient times when workers exposed to lead were observed to suffer more frequently from symptoms of lead encephalopathy. Lead affects many organs and organ systems. The effects of lead on the hematopoietic system (the blood forming system) and the central nervous system and the amount of lead exposure at which these effects occur have been most extensively studied and documented. In addition to the effects lead has on these two systems, lead also has chronic toxic effects on the kidneys, the liver, the skeleton and the gastrointestinal, cardiovascular, endocrine, immune, reproductive, and peripheral (neuromuscular) systems. It is now thought that there may be relationships between lead and chromosomal abnormalities, mutations, and cancer.

TABLE 2

QUARTERLY LEAD CONCENTRATIONS IN THE ST. LOUIS AREA, 1976-1977[a] ARITHMETIC MEAN OVER A QUARTER
(VALID CASES IN PARENTHESES)

Station	Quarter 1			Quarter 2			Quarter 3			Quarter 4			Quarter 5		
	Small size[b] µg/m^3	Large size[c] µg/m^3	Total µg/m^3	Small size µg/m^3	Large size µg/m^3	Total µg/m^3	Small size µg/m^3	Large size µg/m^3	Total µg/m^3	Small size µg/m^3	Large size µg/m^3	Total µg/m^3	Small size µg/m^3	Large size µg/m^3	Total µg/m^3
3	.461 (306)	.144 (311)	.605	.462 (336)	.134 (339)	.596	.573 (285)	.142 (276)	.715	.683 (305)	.158 (299)	.841	.421 (240)	.108 (240)	.529
5	.600 (357)	.187 (355)	.787	.523 (346)	.167 (345)	.690	.629 (329)	.180 (326)	.809	.813 (259)	.227 (255)	1.040	.610 (247)	.168 (246)	.778
6	.668 (173)	.207 (173)	.875	.633 (174)	.205 (176)	.839	.696 (105)	.193 (103)	.889	*	*	*	*	*	*
8	.623 (171)	.247 (172)	.870	.512 (166)	.144 (166)	.656	.571 (166)	.143 (166)	.714	.626 (117)	.128 (118)	.754	.496 (111)	.119 (111)	.615
12	.613 (179)	.199 (179)	.812	.617 (170)	.185 (172)	.802	.804 (151)	.199 (153)	1.033	.850 (155)	.146 (152)	.996	.594 (127)	.152 (131)	.746
15	.307 (171)	.071 (169)	.378	.292 (154)	.081 (152)	.373	.329 (216)	.049 (216)	.378	.426 (156)	.064 (154)	.490	.456 (137)	.062 (138)	.518
18	.242 (159)	.086 (159)	.328	.255 (165)	.078 (164)	.333	.284 (134)	.086 (150)	.370	.360 (159)	.102 (158)	.462	.251 (139)	.097 (137)	.348
20	.445 (127)	.108 (132)	.553	.499 (174)	.121 (177)	.620	.596 (147)	.145 (148)	.741	.602 (142)	.129 (142)	.731	.480 (135)	.107 (132)	.587
22	.168 (154)	.033 (152)	.201	.156 (167)	.028 (163)	.184	.193 (132)	.026 (133)	.219	.197 (152)	.033 (143)	.230	.160 (106)	.027 (104)	.187
24	.183 (164)	*	*	.117 (109)	*	*	.174 (134)	*	*	*	*	*	*	*	*

[a] RAMS Air Pollution Data
[b] < 2 micrometers
[c] > 2 micrometers
* No cases

Dates for Quarters:
1: 1/1/76 - 3/31/76
2: 4/1/76 - 6/30/76
3: 7/1/76 - 9/30/76
4: 10/1/76 - 12/31/76
5: 1/1/77 - 3/14/77

A. Effects on the Nervous System and Behavior

That workers exposed to large amounts of lead suffer more often from encephalopathy has been recognized since the time of Hippocrates. The symptoms observed were dullness, irritability, headaches, loss of memory, and restlessness which often progressed to delirium, coma, convulsions, and even death. These symptoms were also described in infants and young children. The minimum level of lead exposure that results in encephalopathy is still not clearly known. For most adults, lead encephalopathy does not occur until blood lead levels well in excess of 120 µg/100 ml are reached, though there is evidence of acute encephalopathy and death occurring at blood lead levels below 120 µg/100 ml. For children, the minimum blood lead level required to bring on encephalopathy is probably between 80 and 200 µg/100 ml (EPA 1977). After encephalopathy has occurred and a person has survived, there may be permanent sequelae. It has been found that learning ability is impaired in children who have suffered from lead encephalopathy and in some cases, mental retardation has resulted from the attack (NAS 1980).

Due to the dramatic decrease in the incidence of encephalopathy, a well defined dose-response relationship will probably never be known. Instead the effects of lead exposure at much lower levels than those resulting in encephalopathy have been extensively studied. Those effects that occur at lower exposure levels and the dose required to bring about these effects have been studied mainly for children. Because children are still developing, they are more vulnerable to the effects of lead on the nervous system than adults are.

The concern today is whether children with elevated lead exposure may be experiencing subtle neurological damage without ever exhibiting any of the symptoms of lead encephalopathy. There have been several studies showing higher blood lead levels in mentally retarded children than in control groups. David et al. (1976) compared two groups of mild and borderline mentally retarded children: those whose cause of retardation was known and those whose cause was unknown. The group of children "etiology unknown" had statistically significantly raised blood lead concentrations, and the mentally retarded group with "probable etiology" showed no significant difference in blood lead concentrations from those of the normal control group. They concluded that the association between lead and mental retardation occurs over a much wider range than previously thought, and this association is independent of encephalopathic lead poisoning. In fact, they believe that any rise in lead exposure above 24.5 µg/100 ml must be regarded as potentially noxious.

De la Burde and Choate (1972, 1975) compared the performance on a series of psychologic tests of 4 year old children who had asymptomatic lead exposure between 1 and 3 years of age (blood lead levels ranged from 40 µg/ml to 100 µg/100 ml) to a group of children of the same age and socioeconomic background who presumably did not have significant exposure to lead. They found significant differences on psychologic tests between the 2 groups. Exposed children had significant deficits in global IQ and associative abilities, in visual and fine motor

coordination and in behavior. They then did a follow up study of the same children at 7 years of age in order to determine whether the differences detected on the previous tests still existed. They found that the deficits of the lead exposed children on the psychologic tests were still present indicating permanent damage done due to lead exposure. They also found that the behavior disturbances seen in the lead exposed children at 4 years of age became obvious and more of a problem in the school setting. They concluded that there is a significant relationship between asymptomatic lead exposure and deficits in cognitive, perceptual, and behavioral functioning and thus, there is a need for early detection of subclinical lead exposure and for adequate preventive measures.

David et al. (1972) compared the incidence of elevated blood lead levels in five groups of children: (1) a pure hyperactive group with no cause for hyperactivity (2) a hyperactive group with probable cause (prematurity) (3) a hyperactive group with possible cause (4) a group who had recovered from lead poisoning and (5) a nonhyperactive group. The pure hyperactive group had statistically significant higher blood lead levels (mean = 26.2 ± 8 µg/100 ml) than the control group (mean = 22.2 ± 9.6 µg/100 ml). The hyperactive children with probable cause did not show statistically significant differences in blood lead levels from the control group. While a causal relationship between hyperactivity and increased exposure to lead can't be proved, the findings of this study do seem to support the hypothesis that a relationship between moderate lead exposure and altered motor activity exists.

Another study by Perino and Ernhart (1974) of the same general design as the study done by de la Burde and Choate concluded that neurobehavioral deficits do occur in children only moderately exposed to lead. They found significant deficits on motor functions and behavioral tests given to preschoolers who had blood lead levels of 40 to 70 µg/100 ml.

In all of these studies, a causal relationship cannot be shown. It has been asserted that intellectual deficits and hyperactivity could contribute to excessive lead exposure because of the pica habits of these children. However, Beattie et al. (1975) found a strong association between household water lead levels in the homes of pregnant women and mental retardation of the children born to these women. Blood lead levels were significantly higher in the retarded children than the control group a few days after birth. Lead levels of 25.4 ± 12.1 µg/100 ml were measured in the retarded group and 17.8 ± 4.9 µg/100 ml were measured in the control group. In this case, the children's high levels of blood lead cannot be attributed to their behavior. Instead, it can be inferred that lead was a causal factor in the children's mental retardation.

Most of the epidemiological studies done in this area have serious limitations and are open to criticism (NAS 1980). Many of the studies deal with unusual populations of children, such as those in schools for the retarded so it is uncertain whether their findings can be applied to the "average" child. Also many studies have not controlled for other relevant factors involved in child development such as parents IQ and socioeconomic data. Almost all of the studies have relied upon blood lead levels as a

marker of the degree of lead exposure because blood can be obtained easily and harmlessly. However, the level of lead in the blood is not an exact measure of exposure because of storage and movement of lead in other parts of the body, and it is not a direct indicator of lead levels in tissues and organs. Blood is thought to measure primarily recent exposure to lead. Lead in the blood and soft tissues is turned over rapidly; it stays in these areas about four to six weeks. Therefore, when past exposure to lead is important to the study, as it is in the neurobehavioral studies, a different measure for lead exposure must be used. The lead content of deciduous teeth is believed to be a good indicator of total lead exposure over the life of the tooth. The amount of lead in shed teeth has been found to be highly correlated with blood lead levels.

Recently, Needleman et al. (1979) addressed these methodological problems in a study that attempted to measure the neuropsychologic effects of subclinical exposure of lead by comparing the performance of a high lead group to a low lead group on a battery of tests. The lead levels in deciduous teeth were used as a measure of exposure, and only those children whose tooth lead levels fell below the 10th percentile and above the 90th percentile of the entire sample were used as the control group and study group, respectively. They found deficits in several of the neuropsychological tests given to the high lead group, and these remained even after 39 nonlead variables that might affect learning ability and behavior were taken into account. Also, when the two groups were evaluated by teachers the high lead group scored significantly lower on 9 of the 11 indices of classroom performance.

Adults have passed the developmental stages that make children so vulnerable to lead exposure, but the central nervous system is still sensitive to lead toxicity. Various effects on sensory, psychomotor, and psychological functions have been reported in adults with blood lead levels between 40 and 80 µg/100 ml. In one recent study (Lillis et al. 1977), 55 percent of a group of workers with mean blood lead levels of 56 µg/100 ml reported central nervous system symptoms such as tiredness, sleeplessness, irritability, and headache and 39 percent reported muscle and joint pain.

Peripheral neuropathy (as indicated by slowed reflexes and muscular weakness) and altered sensitivity to pain have been observed in workers whose blood lead levels exceed 50 µg/100 ml. Case histories confirm the occurrence of lead induced neuropathies (as indicated by nerve conduction velocity) and other neurological signs such as tremors and wrist and foot drop at blood lead levels of 60 to 80 µg/100 ml and in some cases as low as 30 µg/100 ml (past exposure may have been higher) (EPA 1977).

In summary, there is sufficient evidence to indicate that neurological and behavioral effects do occur at exposures to lead less than those which cause lead encephalopathy. The minimum exposure level or the length of the exposure required is not clearly defined. There is no conclusive evidence that behavior or intellectual development is impaired at blood lead levels below 40 to 50 µg/100 ml, but as mentioned before blood lead levels may be

poor markers of exposure. Recent studies relying on lead in the teeth indicate that neurobehavioral impairment may occur at exposures encountered by some urban children.

B. Effects on Heme Synthesis

The effects of lead on the formation of hemoglobin (the protein that transports oxygen from the respiratory system to each cell) have been widely studied. These effects are detectable at lower levels of lead exposure than is the case with any other organ or organ system. Thus the hematopoietic system is termed the "critical effects organ system."

The effects of lead on heme synthesis are understood much better than the effects of lead on globin synthesis. The process of heme synthesis starts from two small building blocks, glycine and succinate. From these two building blocks a complex molecule, protoporphyrin IX, is formed. The synthesis culminates with the insertion of iron at the center of the porphyrin ring. Heme then combines with specific proteins, one of which is globin, forming hemoglobin (see Figure 1).

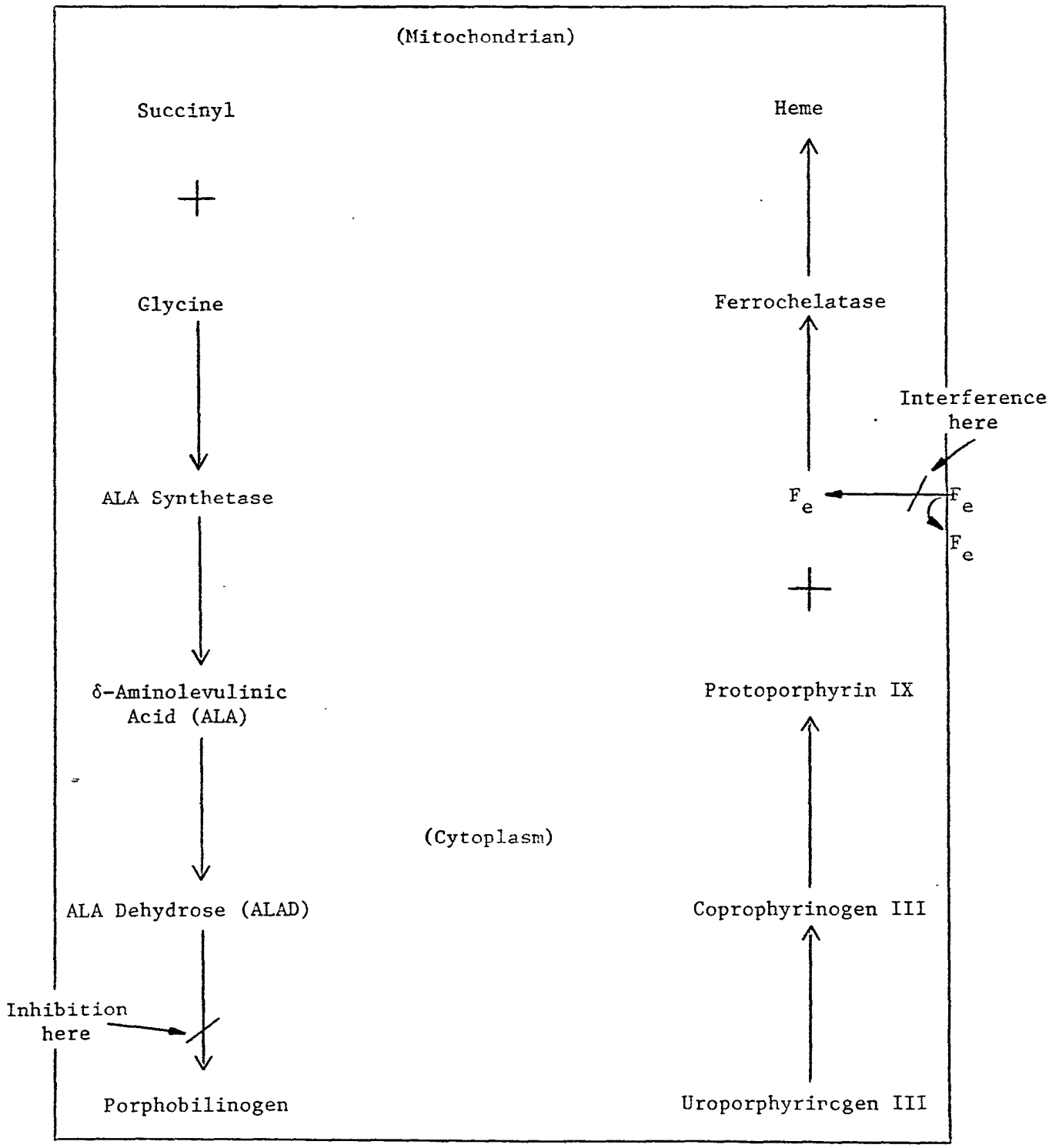
Lead interferes at several points in the synthesis of heme. The two most important points are the inhibition of the enzyme, δ-aminolevulinie acid dehydratase (ALAD) and the interference in the incorporation of iron into the heme. ALAD is the enzyme that acts to connect two molecules of delta aminolevulinic acid (ALA) to form the porphobilinogen ring. As a . result of this interference in the synthesis, precursors to heme, delta aminolevulinic acid, coproporphyrin (CP) and protoporphyrin IX (PROTO) or free erythrocyte protoporphyrin (FEP) can accumulate to elevated levels in the blood.

Inhibition of ALAD Is reflected in increased levels of its substrate, ALA, in the urine. Inhibition starts at a very low blood lead level. Granick et al. (1973) found that inhibition of ALAD in children started as low as 15 µg/100 ml. Nordman and Hernberg (1975) found a statistically significant correlation between ALAD and blood lead levels not exceeding 10 µg/100 ml. Because the inhibition of the enzyme ALAD is one of the earliest effects of lead on heme synthesis, the amount of ALA in the urine is often used as a detector of lead poisoning. Studies have shown that an increase in blood lead levels bring an exponential increase in ALA urinary excretion (Hernberg et al. 1970; Wado 1976).

When lead interferes with the insertion of iron into protoporphyrin IX, it causes the unutilized protoporphyrin IX to accumulate. The exact method of interference is not clearly known; it may be caused by an inhibition of the enzyme, heme synthetase, or the inhibited entry of iron into the mitochondrian. Large amounts of PROTO can accumulate and occupy the available heme pockets in hemoglobin. It remains incorporated into the hemoglobin molecule for the life of the cell, about 120 days. There is also an accumulation of free erythrocyte protoporphyrin (FEP) in the blood and an accumulation of coproporphyrin, a precursor to PROTO, in the urine. The amount of FEP in the erythrocytes has been used as a sensitive indicator of the amount of

Figure 1

Effects of Lead on Heme Metabolism



lead in the body since the development of a simple, accurate instrument for its measurement in 1972. Elevation of FEP in the erythrocytes first appears at 15-30 µg/100 ml in women and children and at 25 µg/100 ml in men (Roels 1976; Promelli 1972; Sassa 1973).

The health significance of the effects of low levels of lead in the blood on heme synthesis are not clearly known. A decrement in hemoglobin resulting in anemia first appears at blood lead levels of 40-50 µg/100 ml. (EPA 1977). This is slightly above the levels where an increase in ALA or FEP in the blood is seen. There is also some evidence that anemia may be caused in part by a shortened lifespan of erythrocytes caused by damage to the erythrocyte membrane. However, the inhibitory effect lead has on the synthesis of heme accounts for most of the decreased concentration of hemoglobin.

Because man appears to possess a large excess of ALAD, the health significance of ALAD inhibition remains doubtful (Roels 1976). The elevation of FEP probably has greater physiological significance; it is believed that a significant modification of FEP is not tolerable (Perlroth et al. 1966). It is possible that heme precursors are toxic in themselves. Elevated levels on ALA may have toxic effects on neuromuscular functions though this evidence was derived from animal studies using high doses of ALA (NAS 1980). The EPA (1977) concluded that elevation of FEP in the blood may also indicate impairment of mitochondrial function and cell respiration. Therefore, the elevation of FEP in the blood may be a good indicator of critical toxic effects of lead on the body.

C. Effects on the Renal System

There is considerable information on the effects of lead to the kidneys. There are two different types of effects that can occur due to lead exposure. One is an acute form of renal tubular damage and the other is a chronic reduction in the ability of the kidneys to remove substances from the bloodstream. The former occurs with short term exposure and is reversible while the latter is considered to be of a slow, progressive nature.

The acute condition which results from lead poisoning is characterized by aminoaciduria (an excess of amino acids in the urine), glucosuria (glucose in the urine), and hyperphosphatura (an excess of phosphates in the urine). These symptoms reflect proximal tubular damages. (The tubules function is to send back into the blood chemicals that the body needs and leave waste products trapped outside.) Therefore, the lead interferes with the filtering process of the kidney which can deprive the body of nutrition. In a group of children with slight lead-related neurological signs, aminoaciduria was found in 8 of 43 children. Their blood lead levels ranged from 40 to 120 µg/100 ml (Pueschel et al. 1972). Chisholm (1968) found 20 of 23 children with chronic lead poisoning to have signs of aminoaciduria. There is evidence that acute lead induced renal effects can be successfully treated. When Chisholm treated the 20 children who showed signs of aminoaciduria, renal function in all returned to normal.

When an individual has been subject to prolonged exposure to lead, a different effect on the renal system can result. This disease is referred to as chronic lead nephropathy. It is characterized by slow development of the lead-exposed kidneys with arteriosclerotic changes, interstitial fibrosis, glomerular atrophy and degeneration of the vessels. All of these changes act to reduce the ability of the kidneys to remove substances from the blood stream. This disease can end in renal failure, even long after the exposure to lead is ended. Evidence of a causal relationship between chronic nephropathy and childhood lead poisoning was found in two studies that compared the deaths from chronic nephritis among people under 30 years of age in Queensland, Australia (WHO 1977). These cases involved childhood exposure with a latency of 10-30 years for the development of renal impairment. Chronic lead nephropathy is more common among people who have had a fairly high exposure to lead for more than 10 years than those exposed less than 10 years. This disease is difficult to detect in its early stages; in fact, individuals may lose up to two-thirds of the functional capacity of the kidneys and still be asymptomatic (EPA 1977).

In summary, acute renal effects can occur in children and adults with subtle signs of lead poisoning. Damage occurs in the proximal tubular cells resulting in decreased efficiency in filtering out proteins and other substances. This condition can be successfully treated. Prolonged exposure can lead to more serious effects on the renal system that are irreversible and may result in renal failure. Little is known about dose-response relationships for these effects.

D. Reproductive Effects

There is extensive evidence from occupationally exposed populations that lead,has serious effects on the human reproductive system. The effects include an increase in the number of miscarriages and stillborns, disruption of the ovarian cycle, loss of fertility in men, loss of libido and reduced potency in men and the increased likelihood of abnormal pregnancy.

Data on the adverse effects of lead on the reproductive system have existed since the turn of the century. A study done by Legge (1901) on 77 female lead workers found that out of 212 pregnancies, there were 90 miscarriages, 21 stillborns, and of the 101 live births, 40 infants died in the first year. The exposure to lead in this study was very high. However, a more recent study suggests that miscarriages occur in women with.only modest exposure (EPA 1979). The women is this study were wives of lead workers and not actually occupationally exposed to lead themselves. No actual levels of lead exposure were reported so no conclusions can be drawn about the minimally toxic level of exposure that would result in miscarriage.

It has also been found that subtoxic lead absorption during pregnancy is associated with preterm deliveries and premature fetal membrane rupture in term infants. Fahim et al. (1976) found the incidence of premature membrane rupture and preterm delivery to be 17 percent for an area 30 to 50 miles west of a lead mining area in Missouri and only .41 percent in a

Missouri urban area where no lead mining activities existed. The maternal and fetal blood lead levels at birth for the deliveries with preterm membrane rupture were 26 and 13 µg/100 ml, whereas for normal deliveries they were about 14 and 4 µg/100 ml respectively.

Lead can affect fertility as well as conception. Lancranjan et al. (1975) reported that moderately increased lead absorption (blood lead means = 52.8 µg/100 ml) resulted in decreased fertile ability in men. They found an increased number of malformed sperm, decreased number of total sperm, and sperm with decreased motility. They concluded that hypofertility induced by lead is probably due to lead's direct toxic effect on the gonads. Lead has also been associated with loss of libido and impotency but no dose-response relationship has been defined for these effects (EPA 1977). Lead also affects fertility in women. It has been estimated that short term exposure at ambient air levels of less than 7 $\mu g/m^3$ may cause an increase in anovular cycles (menstrual cycles in which ovulation does not occur) and disturbance in the lutein phase (ovulation occurs and the uterine walls in prepared for the egg to be implanted in this phase) (EPA, 1977).

Lead can pass through the placental barrier and accumulate in fetal tissues after about the 12th week of pregnancy. There is a high correlation between blood lead levels of women immediately after birth and the blood lead levels of their infants. Gershanik et al. (1974) found a correlation coefficient of .6377 for lead levels in infants and their corresponding mothers. There is little evidence to suggest that lead causes congenital malformations. However, since the neurobehavioral system of children is so susceptible to low exposures to lead, it is generally assumed that the fetus and newborn would be at least as susceptible. The study by Beattie (1975) (mentioned previously in the section on neurobehavioral effects) found an association between a high concentration of lead in the drinking water of pregnant women and mental retardation in their corresponding infants. Because there is so little evidence of the effects of prenatal lead exposure on mental functions in infants, it is impossible to define a dose-response relationship.

E. Carcinogenic Effects

The epidemiological evidence of the relationship between lead exposure and cancer is questionable. Several studies have been done to find the cause of death among people overly exposed to lead. Most of these studies found no significant correlation between lead exposure and deaths caused by cancer. However, there is evidence that lead causes cancer in laboratory animals.

A study done by Dingwall-Fordyce and Lane (1963) examined the causes of death among 267 men who died between 1926 and 1961. The men were classified as to lead exposure according to the nature of their work and if, they worked in an occupation with high lead exposure, according to their urinary lead excretion. They found in comparing the high lead exposure group to the low or negligible lead exposure groups that there was not a significant excess number of deaths caused by malignant tumors in the

high lead exposure group. There was an excess number of deaths but these were due mainly to vascular lesions of the central nervous system.

In another relevant study, beginning in 1938, orchardists who had sprayed fruit trees with lead arsenate at one time and the rest of the population living in the area of the spraying were classified into groups according to the degree of exposure to lead. In 1968, a follow up study was begun (Nelson et al. 1973). The status (living or dead) of the persons in the study was determined and if they had died the cause of death was determined. There was no suggestion of a relationship between lead exposure and death from three major causes of death: heart disease, cancer, and stroke.

In a more recent study (Cooper 1975) of mortality among lead smelter workers and lead battery workers, the causes of death that showed a statistically significant elevation included "all malignant neoplasms" and cancer of "other sites" in battery workers. However, the author concluded that the excess deaths due to neoplasms cannot be attributed to lead "because there was no consistent association between the incidence of cancer deaths and either length of employment or estimated exposures to lead".

There have been many studies done concerning the relationship of lead exposure and cancer mortality in laboratory animals. Several studies have shown that lead causes renal tumors in rats and hamsters, and one study each has associated lead with renal cancer in mice, lung tumors in hamsters and brain tumors in rats (EPA 1979). In all of these studies, the exposure to lead was very high, many times more than what a typical individual would be exposed to. Therefore, a dose-response relationship for much lower exposures to lead cannot be extrapolated from the data.

There is too little data to make a conclusion about the carcinogenicity of lead. If it is a carcinogen, then it is probably a weak one since the epidemiological studies done shown no significant association between lead exposure and cancer mortality. Further research is needed before any conclusions can be made.

F. Effects on the Cardiovascular System

When exposed to a high level of lead over a prolonged period of time, arteriosclerotic changes (hardening of the arteries) can take place in the kidney and lead to chronic renal disease (see section on renal effects). Dingwell-Fordyce and Lane (1963) reported a marked increase in mortality among lead workers due to cerebrovascular disease. This observation applied only to workers exposed to lead during the first quarter of the century, when lead exposure on the job was very bad. The same increase was not seen in workers employed more recently. In a more recent epidemiological study (Cooper 1975), there was no excess mortality due to diseases associated with hypertension or vasculopathy. It would appear from the studies done that lead affects the vascular system only with very high industrial exposure-- like that seen around the turn of the century. These effects could either be direct effects on the blood vessels

themselves or a consequence of the toxic effects lead has on the renal system.

There is evidence that lead has toxic effects on the heart. Cases of structural and functional changes of the myocardiun (the muscular substance of the heart) have been described in adults and children with clinical lead poisoning (EPA 1977). In many cases when encephalopathy has been treated, the electrocardiographic abnormalities disappeared. Silver and Rodriguez-Torres found abnormal cardiograms in 21 of 30 children who had symptoms of lead toxicity. After therapy, the abnormalities were found in only 4 children.

Conclusive evidence of the effects of lead exposure on the cardiovascular system is not available so no dose-response relationships can be defined. The effects that have been noted may only affect those who have been exposed to a vary high level of lead.

G. Chromosomal Effects

The study of lead effects on chromosomes is technically difficult and therefore, it is hard to make any conclusions about these effects. The significance of the implication of injury to chromosomes is great. Each chromosome must separate correctly into two chromatids during cell division, and these chromatids must be equally redistributed in order to reproduce stable new cells for the maintenance of healthy tissue. Any chromosomal aberrations can be responsible for consequences as serious as defects in offspring of the affected individual (EPA 1977).

There are both negative and positive reports in this area. O'Riordan and Evans (1974) examined the lymphocytes of 70 male workers occupationally exposed to lead and found that chromosomal aberration frequencies were low and not significantly different from the control group. Bauchinger and Schmid (1976) also reported no increase in chromosomal aberration yield levels over the control group when they studied lymphocyte chromosomes of 20 industrial workers who showed a 20 to 30 percent increase in blood lead levels over the general population.

Studies done on 11 subjects (Forni et al. 1976) before and during initial occupation exposure to moderate quantities of lead fumes in a storage battery plant showed a doubling of the rate of abnormal cell divisions after one month. This rate increased after two months, remained in this range for seven months and then decreased somewhat. It must be noted here the small number of subjects studied making the results less reliable.

Chromosomal aberrations in lymphocytes of 24 workers in a zinc smelting plant who had increased blood levels of lead and cadmium were significantly higher than those of 15 members of the control group. Though it can't be determined which of the metals was to blame the authors suggest that the aberrations could be caused by the synergistic effects of several metal compounds (Bauchinger and Schmid 1972).

The evidence of chromosomal aberrations is conflicting and inconclusive. No clear dose-response relationship can be defined for the effects of lead on chromosomes. Because there have been studies showing an increase in chromosomal aberrations in lymphocytes, it may be that chromosomal damage occurs in other cells, possibly eggs and sperm (EPA 1977). The effects of lead on reproduction (notably miscarriages and stillborns) discussed previously may in part be due to the effects of lead on genetic material.

H. Gastrointestinal Effects

It has long been known that colic is associated with lead poisoning. Colic is usually an early symptom of poisoning.. It is seen most commonly in industrial exposure ccases but it is also a symptom of lead poisoning in infants and children (EPA 1977).

There is insufficient data by which to define a dose-response relationship for the effects of lead on the gastrointestinal system. One study of 64 men with blood lead levels between 40 and 80 µg/100 ml found 13 of those men to be suffering from colic and constipation (Beritic 1971). However, no conclusions as to the exposure required for colic to occur can be made from this study.

I. Effects on the Endocrine System

Effects of lead on the endocrine system are not well defined. It is known to decrease the thyroid function and possibly interfere with pituitary function (EPA 1977). There is not enough evidence in this area to make any conclusion about the effects or to define any dose-response relationship.

J. Effects on the Immunological System

Animal studies have indicated that exposure to lead may interfere with normal susceptibility to infection. Decrease production of antibodies may be one of the factors involved in this increased susceptibility. Williams et al. (1954) reported that lead binds antibodies in vitro and could do so in vivo. Again there is not enough evidence to make any conclusions about the effects of lead on the immunological system.

3. Conclusion

Though there is much known about the toxic effects of lead on the body, there is still controversy over what level of lead in the body should be considered safe. There is general consensus that toxic effects do occur in the body at minimum blood lead levels of 40 to 50 µg/100 ml. However, many researchers feel that a safe maximum blood lead level is closer to 25 µg/100 ml. Adverse neurobehavioral effects have been found to occur at blood lead levels of 20 to 25 µg/100 ml in children, and FEP elevation, which the EPA has concluded may indicate serious toxic effects, occurs at blood lead levels of 15 to 30 µg/100 ml. Therefore, the general population

may be subject to exposures to lead that produce deleterious effects. Whether this is the case may never be known since it would be virtually impossible to demonstrate these effects if the majority of the population were suffering from them.

The next chapter presents an empirical formulation of the theoretical model discussed in Chapter 3 using the St. Louis morbidity data set and the RAMS air pollution data. Specifically, the relationship between the presence of ozone and lead in the ambient air and the production of health is estimated. Other pollutants will also be examined for their effect on the production of health The results of this estimation allow benefits derived from a reduction in air pollution to be calculated.

APPENDIX 3

DATA-BASE DOCUMENTATION

The purpose of this appendix is to provide the reader with detailed information regarding the St. Louis health survey. Before doing so, however, note that these data together with the air quality data obtained from the RAPS, the city of St. Louis, the county of St. Louis, and the Illinois Environmental Protection Agency can be obtained from:

Professor Shelby Gerking
Department of Economics
University Station, P.O. Box 3985
University of Wyoming
Laramie, WY 82071

Telephone (307) 766-4890

Data requests will be subject to a charge equal to the price of the computer tapes required.

The sections below reproduce word-for-word key material concerning the data base that was presented in Koontz (1981). A copy of the complete version of this volume may be obtained from Gerking at the above address at a price set equal to photoreproduction charges. Copies may also be available from Geomet Technologies, Inc., 1801 Research Boulevard, Rockville, Maryland 20850; (301) 424-9133.

The following section is reproduced verbatim from Koontz (1981).

Section 1.0

INTRODUCTION

This section describes the history of the study under which data described herein were collected, outlines the contents of this report, and notes a companion report which is forthcoming. In April of 1974, GEOMET, Incorporated contracted with the Environmental Protection Agency (EPA) under EPA Contract Number 68-01-2280 to perform a household interview survey intended to provide quantifications of air pollution-induced morbidity costs for individuals residing in the St. Louis Standard Metropolitan Statistical Area (SMSA). The study was intended to take advantage of comprehensive air quality data deriving from a network of 25 monitoring stations located throughout the SMSA and known as the Regional Air Monitoring System (RAMS).

During Phase I, Stage 1 of the study, GEOMET staff reviewed related literature and ongoing research and specified statistical models and an analytical plan for estimating relationships between human morbidity (and associated costs) and air pollution. In a Phase I, Stage 1 draft report dated September 1974, GEOMET staff concluded (and EPA staff subsequently concurred) that a formal procedural pretest of the instruments and study design was needed before undertaking the full-scale data collection effort in Phase II. GEOMET then began Phase I, Stage 2 of the contract by developing survey instruments and a sampling plan for collecting the requisite data in a representative fashion. Following approval of the pretest instruments and methodology by the Project Officer and clearance within EPA, the pretest family of forms and a supporting statement were submitted to the Office of Management and Budget (OMB) in March 1975; simultaneous clearance of the pretest and full-scale survey was sought. OMB reviewers initially disapproved the entire package but subsequently gave clearance for the pretest only in August 1975, with approval of the full-scale effort contingent upon pretest outcomes. During this period, a final version of the Phase I, Stage 1 report was prepared, which was dated July 1975 and which was included in a revised OMB submission package.

Phase I, Stage 2 continued with the administration of the pretest survey to 72 St. Louis households between November 1975 and January 1976. Air quality data for this time period were available by March 1976. Following the development, programming and application of comprehensive procedures for pollution exposure estimation, careful analysis of the pretest data, and questionnaire modifications based on the pretest experience, a new OMB package for the full-scale survey was sent to the Project Officer in draft form in May 1977 and in final form in July 1977. EPA approved and forwarded the package to OMB in November 1977. OMB clearance was ultimately obtained in January 1978.

Owing to the delays in receiving OMB clearance and the careful analysis and documentation required to secure such clearance, the RAMS network was no longer in operation by the time clearance was ultimately obtained. However, a network of monitoring stations operated by local agencies provided sufficient coverage of the St. Louis urbanized area to permit implementation of the survey within that subset of the SMSA. GEOMET staff submitted a Phase I, Stage 2 draft report to EPA in March 1978 which described the pretest materials, methods, and results and which outlined the plans for full-scale data collection during Phase II. Following selection of the household sample for Phase II, receipt of a letter of endorsement from the Missouri Governor's office (Appendix A), and conduct of an interviewer training session in St. Louis, data collection efforts commenced. Households were interviewed between June 1978 and July 1979.

Under EPA contract 68-01-5773, Task Order 8, GEOMET staff edited, coded, keyed, verified, and processed the data from the Phase II survey; this activity took place between November 1979 and July 1980. The EPA Project Officer also decided, on the basis of research conducted by a group of economists at the University of Wyoming, that additional information should be obtained from households completing all aspects of the Phase II (or basic) survey in order to enhance the analytic potential of the basic data base. GEOMET staff consulted in the design of a supplemental data form for this purpose and were responsible for coordinating field efforts, processing the resultant data, and merging it with the basic data. Supplemental data were collected between April and August, 1980. Editing, coding, and keying efforts were completed by the end of 1980. Final error resolution, merging, and documentation took place during the first half of 1981.

In January 1980, the research and development component of GEOMET became a subsidiary of the parent company under the name of GEOMET Technologies, Inc. (GTI). In subsequent references to the company authoring this report, the subsidiary name is utilized.

The purpose of this report is to provide potential users of the data base described herein with a sufficient understanding of the design aspects and information sources underlying the data to permit proper utility of all data elements during subsequent analyses. Section 2.0 of the report provides an overview of the survey design and a summary of the survey response rates. Section 3.0 outlines the general contents of all survey forms and the manner in which they were administered. Section 4.0 provides a detailed description of the data base organization and contents. Section 5.0 highlights additional factors to be considered in using the data base. Appendices to the report illustrate survey instruments and interviewing guidelines and list selected coding schemes which are too lengthy for inclusion in the main body of the report. A forthcoming report will describe the manner in which respondent activity patterns coded in the data base should be utilized in order to estimate exposure to ambient pollutant concentrations both prior to and contemporaneous with the time during which each respondent participated in the basic survey.

The following section is reproduced verbatim from Koontz (1981).

Section 2.0

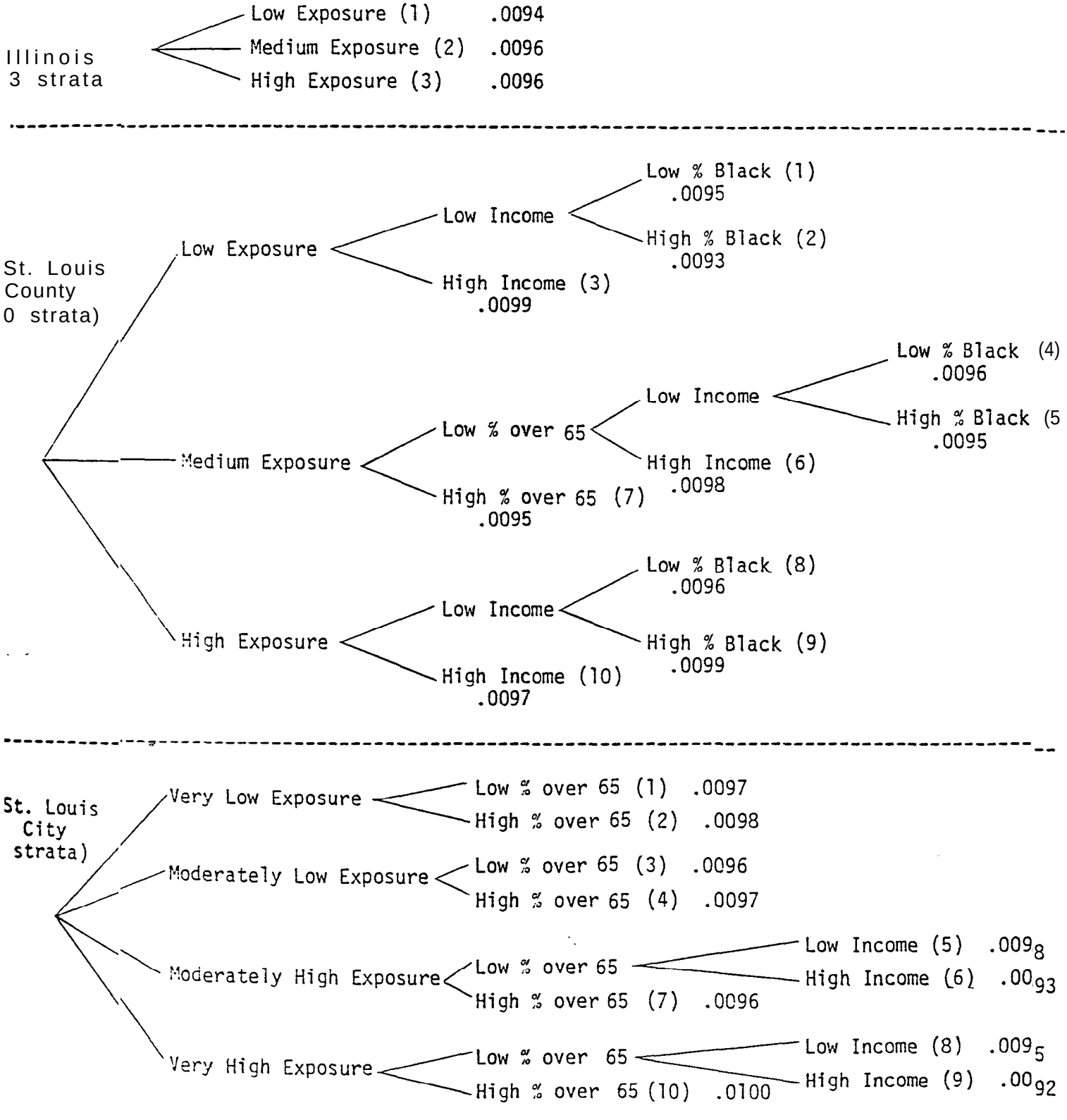
SURVEY DESIGN AND RESPONSE RATES

A stratified random sample of 4,160 households was solicited for the basic survey. The sampling frame from which these households were drawn is represented by 238 census tracts within the urbanized area of the St. Louis SMSA which were bounded by the network of continuous monitoring stations operating at the time during which the sample was drawn (Spring, 1978). According to 1970 census data, the 238 tracts collectively contain 432,162 households, or nearly 60 percent of the households within the entire SMSA. The overall sampling fraction for the study, then, was 4,160/432,162 or approximately one percent.

The multistage sampling strategy used for the study was designed to yield an ultimate sample of households with equal selection probabilities, so that subsequent analyses would not require any weighting. In general terms, the sampling strategy required: (1) stratification of census tracts and subsequent selection using the probability-proportional-to-size (PPS) technique; (2) selection of blocks within selected tracts, again using the technique; and (3) selection of 4 households per selected block to compensate for prior unequal selection probabilities and to facilitate survey logistics. During stages (1) and (2), it was necessary to combine selected tracts/blocks to ensure that a sufficient number of blocks/households would be available at subsequent sampling stages. A primary stratification of census tracts was made by geopolitical boundary (i.e., St. Louis City, St. Louis County, Illinois); a secondary stratification was made within the three primary strata according to historical pollution levels; where heterogeneity remained in any substrata with respect to income, age, or race, further stratification was made based on one or more of these variables. The ultimate number of strata which resulted from this process was 23; these strata and their respective sampling fractions (ranging from .0092 to .0100) are displayed in Figure 1.

In selecting census tracts for the survey, 1970 census counts of dwelling units were used as the measure of size for Illinois in applying the PPS technique. For St. Louis City and County, updated counts based on 1976 compilations by the Union Electric Company were utilized as the measure of size. In selecting blocks for St. Louis City and County, updated counts were made by visiting all blocks within tracts whose updated counts differed from 1970 census counts by more than 10 percent. In St. Louis County, where some significant areas of new home construction had emerged since 1970, it was necessary to add some newly created blocks to those listed in the U.S. Census Bureau Block Statistics publication for the St. Louis SMSA. Since updated counts were not obtained for all tracts and blocks, the measures of size doubtlessly differed from the actual sizes in many instances. Thus it is more proper to speak of individuals

Figure 1. Sampling Strata and Sampling Fractions for Basic Survey.



participating in the survey as having approximately equal selection probabilities.

The 4,160 households solicited for the study were enrolled on a weekly basis, at the rate of 80 per week over 52 weeks. The census blocks selected for the survey were systematically allocated across the 52 enrollment weeks so that the study sample would represent the universe both on an overall basis and at any particular point in time. Each household participating in the survey was administered a background interview and four biweekly followup interviews; thus each household participating in the survey to completion contributed eight weeks of followup information. Since the last group of households was enrolled in the survey during week 52 and was followed for eight weeks (including week 52), the actual duration of the survey was 59 weeks. The nature of the enrollment process was such that a maximum of 80 households would participate during the first and last weeks of the survey and a maximum of 640 households would participate during the central portion of the survey (see Figure 2). Another interesting feature of the sampling design is that enrollment weeks separated by one calendar year and having fewer than 640 solicited households each do represent 640 solicited households when paired (i.e., weeks 1 and 53, weeks 2 and 54,..., weeks 7 and 59). Households in pairs of enrollment weeks related in this manner could conceivably be treated as seasonally equivalent and grouped together at some analytical stage.

Only households which participated in the basic survey to completion were eligible for the supplemental data collection effort. Prior respondents were located on the basis of home/work addresses and home telephone numbers reported during the basic survey. If a respondent could not be located through the household's telephone number or through the employer of any household member, a letter was sent to the last reported address which was labeled "Do Not Forward -- Address Correction Requested". This approach identified many previously unlocated respondents as still at the same household address but with a new, unpublished phone number or as having moved, with the new address supplied by the U.S. Postal Service. For households which had moved but for which no forwarding address was available, neighbors and local social/religious organizations were contacted in a final effort to locate the respondent.

In Table 1, the response rates for the basic and supplemental surveys are listed both by enrollment week and in total. Basic survey respondents are segregated as "Background Interview Only", "Partial Followup", or "Complete Followup", with the latter group eligible for supplemental data collection. The summary figures at the bottom of the table indicate that 3,063 or 73.63 percent of the 4,160 solicited households participated in the basic survey to completion, with an additional 93 households (2.24 percent) providing partial followup information. A total of 3,256 (78.27 percent) of the households participated in the background interview portion of the survey. The 2,594 households providing supplemental information represent 62.36 percent of the 4,160 households solicited for the basic survey or 84.69 percent of those eligible for the supplemental data collection effort. The major reasons for nonresponse during the background interview were outright refusal, vacant dwelling unit, or not-at-home

Figure 2. Household Enrollment Process for Basic Survey.

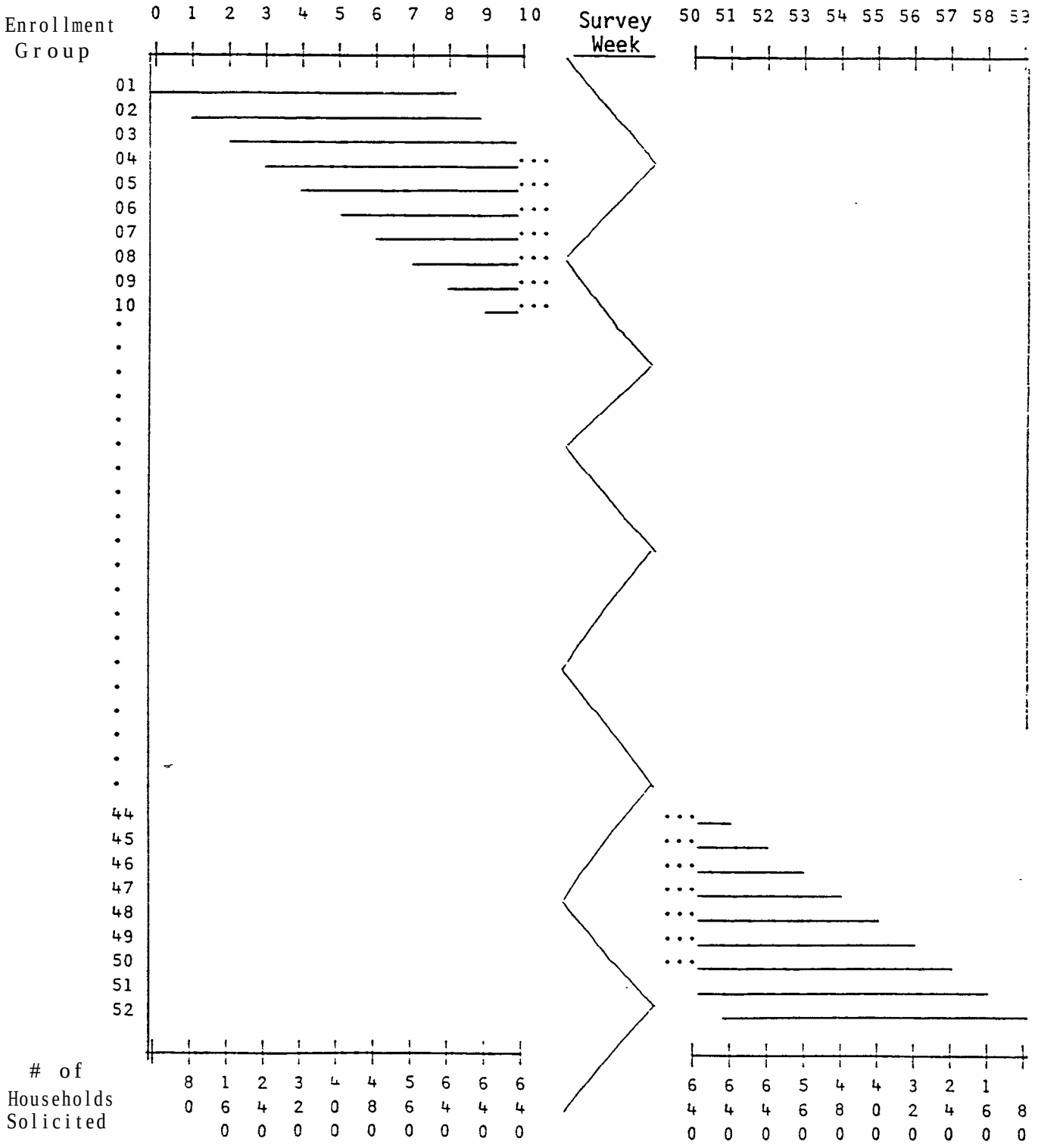


Table 1. Response Rates for Basic and Supplemental Surveys.

Week Number	No. Interview	Background Interview Only	Partial Follow-up	Complete Follow-up*	Supplemental Interview
1	14	3	-	63	51
2	23	10	2	45	37
3	13	-	6	61	57
4	23	2	1	54	49
5	25	5	3	47	43
6	15	1	2	62	50
7	16	1	1	62	49
8	22	3	4	51	45
9	19	3	3	55	46
10	18	3	6	53	48
11	23	4	2	51	47
12	16	3	1	60	52
13	18	2	-	60	56
14	15	3	3	59	51
15	25	1	-	53	49
16	15	1	3	61	56
17	21	-	1	58	48
18	19	4	-	57	36
19	18	1	1	60	49
20	14	2	1	63	57
21	17	3	2	58	51
22	21	-	-	59	52
23	12	3	1	64	55
23	12	1	-	67	47
25	22	-	1	57	45
26	19	1	-	60	58
27	16	-	-	64	55
28	16	3	1	60	47
29	13	1	-	66	54
30	23	-	3	54	46
31	16	-	4	60	52
32	15	2	1	62	50
33	19	2	-	59	51
34	20	-	1	59	52
35	18	2	2	58	44
36	11	1	-	68	59
37	19	2	1	58	52
38	14	1	2	63	51
39	20	-	3	57	51
40	13	2	3	62	50
41	11	-	1	68	57
42	18	3	3	56	48
43	17	3	2	58	48
44	10	1	3	66	53
45	17	4	1	58	44
46	14	-	6	60	52
47	21	1	-	58	56
48	14	3	2	61	47
49	15	3	-	62	51
50	19	-	2	69	45
51	19	4	7	50	45
52	20	2	1	57	50
TOTAL ALL WEEKS	904	100	93	3.063	2,594
% OF SOLICITED HOUSEHOLDS(4,160)	21.73	2.40	2.24	73.63	62.36 (84.69**)

* Eligible for Supplemental Interview

** % of Households Eligible for Supplemental Interview

during any of three attempted contacts. The major reasons for nonresponse during followup were outright refusal or relocation outside the St. Louis area. The major reasons for nonresponse during supplemental data collection were outright refusal or failure to locate/contact.

The following section is reproduced verbatim from Koontz (1981).

SECTION 3.0

ADMINISTRATION AND GENERAL CONTENTS OF SURVEY FORMS

Three forms - a "Household Background Questionnaire," an "Individual Background Interview," and a "Telephone Interview Form" - were designed to meet the basic survey requirements. The first two instruments were to be utilized during the initial interview with each household agreeing to participate in the survey. The third instrument was to be administered during each of four biweekly follow-up periods for each family agreeing to continue participation in the survey. A two-day training session was held prior to conduct of the survey during which the Interviewers' Manual (Appendix B) was reviewed and mock interviews were conducted. The basic role and contents of each form are discussed below.

The Household Background Questionnaire (Appendix C) was designed to identify all household members, their basic demographic characteristics, and the socioeconomic status (SES) of the household. The demographic characteristics included age, sex, race, and education. The presence or absence of 3 types of health insurance (medicare, medicaid, private) was asked for each individual. A categorical question was phrased concerning overall income. The final questions elicited responses which enable the calculation of the number of persons per room for the household.

The Individual Background Interview (Appendix D) was designed to identify baseline health levels, chief activities and commuting routines, and smoking habits of each individual in participating households. The respondent was asked to reply to questions on behalf of each household member. To measure baseline health levels, questions were formulated concerning (1) the respondent's subjective appraisal of each individual's health status, (2) chronic conditions suffered by individuals, and (3) the approximate number of contacts with the health delivery system by each individual during the past year. A subsequent series of questions concerned typical commuting and waiting times associated with a doctor visit.

To ascertain individuals' major activities and usual commuting routines, a series of questions were posed which (1) established each individual's major activity and (2) provided skip patterns (based on the major activity and age of each individual) aimed at determining the typical locations, hours, and days for activities away from the home. For individuals whose hours at a given activity might not be known in advance, provision was made to record those hours during the followup telephone interviews. For individuals who worked full-time or part-time, questions were formed to identify (1) the nature of the job, (2) the extent of work outdoors, and (3) vacation/sick leave policies. The final questions comprising the Individual Background Interview concerned the extent and duration of cigarette smoking, as well as the extent of pipe and cigar

smoking, for individuals of age 14 years or older who were smoking at the time the survey was administered. At the conclusion of the background interview, the respondent was asked to participate in the followup portion of the basic survey and was offered a "Major Activity and Illness Calendar" (Appendix E) as a memory aid for subsequent callbacks.

The Telephone Interview Form (Appendix F) was designed to elicit followup morbidity data and changes in normal routines for members of participating households. The following health consequences were examined during the followup interview:

- Disability
  - Days confined to bed
  - Days confined to the house
  - Days of restricted activity.
- Absenteeism
  - Days of absence from work (including work around the house for housewives)
  - Days of absence from school.
- Contacts with the Health Delivery System
  - Days in a hospital
  - Days in a nursing home
  - Days on which a physician was seen or phoned for advice.
- Receipt of Ancillary Services
  - Shots, x-rays, or lab tests associated with physician visits
  - Prescriptions.

The chief health problem associated with each type of health consequence was also sought during each telephone interview. The basic activity changes examined were:

- Changes from Normal Routine
  - Days out of town
  - Days at home for vacation or holiday from work/school
- Changes in Job Status (finding a job, changing jobs, or losing a job)
- Hour Updates for Selected Activities (for individuals whose usual hours at such activities could not be ascertained during the background interview).

A form (Appendix G) was subsequently designed to supplement the information obtained during the background portion of the survey. This form was administered to households who participated in the background portion of the survey and who remained in the survey throughout their four biweekly followup periods. Prior to administration of the supplemental data form, a one-day training session was held during which written instructions (Appendix H) were reviewed and mock interviews were conducted. The basic contents of the form are as follows:

- Part A - Residence Questions

  - Presence of air conditioning and type of heating/cooking fuel for residence at time of background survey
  - Residential history.

- Part B - Work Questions

  - Length of employment, exposure to irritants, membership in labor union, vacation/sick leave availability, and salary for jobs held at time of background survey
  - Description and length of major occupation, including indoor vs. outdoor work and exposure to irritants, for retired individuals.

- Part C - Income and Health Insurance Questions

  - Proportion of income not related to jobs and sources of that income
  - Usual charge for doctor visit and portion covered by insurance.

- Part D - General Health Questions

  - Height, current vs. maximum weight, exercise patterns, and time spent outdoors/watching TV
  - Length of long-term health problems identified in background survey
  - Smoking history for individuals 14 years or older who were nonsmokers at time of background survey.

- Part E - Diet Questions

  - Average number of hot meals per day and frequency of dieting
  - Consumption of cured meats, sweets, salty snacks, caffeinated drinks, and drinks with sugar
  - Alcohol consumption for individuals 14 years or older.

The following section is reproduced verbatim from Koontz (1981).

Section 4.0

DATA BASE ORGANIZATION AND CONTENTS

The magnetic tape supplied by GTI to EPA together with this report has the following general specifications:

- IBM standard label
- EBCDIC character code
- 9 - track
- 6250 bpi
- odd parity.

The tape contains 52 files, one for each enrollment week. The files are labeled as follows:

- STLOUIS.WEEK01 (households enrolled during week 01 of the basic survey)
- STLOUIS.WEEK02 (households enrolled during week 02)

- STLOUIS.WEEK52 (households enrolled during week 52).

Each file contains one record for each individual participating in the background portion of the survey, whether or not he/she participated in subsequent portions of the study. Each record consists of 2007 numeric characters (i.e., LRECL = 2007); each block contains six records (i.e., BLKSIZE = 12,042). Files are in numerical order, with a primary sort by household number and a secondary sort by person number. No information is contained in the data base which permits identification of particular households or individuals. Only the individual given custody of the survey data forms and control cards/sheets is capable of linking numerical identifiers to their respective households/individuals. The forms and control cards/sheets should be kept in a locked room/file, with access given only to EPA researchers/contractors using the data base who wish to verify information as needed or resolve questions arising from the data.

Each record can be segregated according to the following general components:

- Identifiers - columns 1 - 7
- Background interview information - columns 8 - 101
- Location information for use in exposure estimation - columns 102 - 1490
- Followup interview information - columns 1491 - 1886
- Supplemental information - columns 1887 - 2007

An important link between the basic survey and the supplemental survey is column 1887, which identifies whether or not households were eligible for the supplemental survey and whether or not they participated in it. Individuals in households either not eligible for or not participating in the supplemental effort have blanks throughout columns 1888 - 2007 of their respective records.

A detailed description of the data base contents constitutes the remainder of this section. For each item in the data base, its location in the record (column number or numbers) is provided together with the source of the information (form and question number or numbers), codes applied to the item and comments (where applicable).

Column Number(s)	Item Number and Description
1-2	1. Week Number Source: Face sheet (p.1) of Household Background Questionnaire (HBQ) Codes: 01 to 52 Comments: Refers to week during which household was enrolled in the basic survey; '01' refers to households enrolled during week beginning June 4, 1978; '02' refers to households enrolled during week beginning June 11, 1973; . . . '52' refers to households enrolled during week beginning May 27, 1979.
3-5	2. Household Number (Family Number) Source: Face sheet of HBQ Codes: 001 to 080 Comments: Maximum of 80 households enrolled per week; column 3 always zero; only columns 4-5 need be accessed.
6-7	3. Person Number Source: Table below question A5 on p.3 of HBQ Codes: 01 to 14 Comments: 14 believed to be upper bound, but may be higher: can be ascertained by generating frequency distribution.
8-9	4. Age (at time of HBQ) Source: Table below question A5 on p.3 of HBQ Codes: 00 - less than 1 year of age 01 to 96 - age in years 97 - 97 years of age or older 98 - refused 99 - don't know
10-11	5. Education (last year completed at time of HBQ) Source: Table below question A5 on p.3 of HBQ Codes: 01 to 20 - years of education completed 97 - special education 98 - refused 99 - don't know Comments: Question asked only for individuals of age 14 years or older; 20 believed to be upper bound, but may be higher; can be ascertained by generating frequency distribution.

Column Number(s)	Item Number and Description
12	6. Medicare Insurance Coverage Source: Question C1 on p.6 of HBQ; table below question A5 on p.3 of HBQ Codes: 0 - no 1 - yes 8 - refused 9 - don't know
13	7. Medicaid Insurance Coverage source: Question C2 on p.6 of HBQ; table below question A5 on p.3 of HBQ Codes: 0 - no 1 - yes 8 - refused 9 - don't know
14	8. Private Insurance Coverage Source: Question C3 on p.6 of HBQ; table below question A5 on p.3 of HBQ Codes: 0 - no 1 - yes 8 - refused 9 - don't know
15	9. Sex Source: Table below question A5 on p.3 of HBQ Codes: 1 - male 2 - female
16-17	10. Total Number of Persons in the Household Source: Question A6 on p.4 of HBQ Codes: 01 to 14 Comments: Should generally be equal to highest person number in household; 14 believed to be upper bound, but may be higher.
18	11. Sate of Respondent Source: Question A7 on p.4 of HBQ Codes: 1 - white 2 - black 3 - Asian or Pacific-islander 4 - Alaskan native or American Indian 9 - don't know Comments: Response categories per OMB requirements.

Column Number (s)	Item Number and Description
19	12. Hispanic (or Spanish) Origin of Respondent Source: Question A3 on p.4 of HBQ Codes: 1 - yes 2 - no 9 - don't know Comments: Question required by OMB.
20-22	13. Number of Rooms (Including Bathrooms) in the Household's Dwelling Unit Source: Question C4 on p.7 of HBQ Codes: 010 to 200 Comments: 3-digit code with implicit decimal point between 2nd and 3rd digits; e.g., 8 rooms coded as 080, 10-1/2 rooms coded as 105; 200 (20 rooms) believed to be upper bound.
23-24	14. Number of Bathrooms in the Household's Dwelling Unit Source: Question C5 on p.7 of HBQ Codes: 00 to 50 Comments: 2-digit code with implicit decimal point between 1st and 2nd digits; e.g., 2 bathrooms coded as 20, 1-1/2 bathrooms coded as 15; 50 (5 bathrooms) believed to be upper bound.
25	15. Total Combined Family Income Source: Question C5 on p.7 of HBQ Codes: 0 - refused 1 - under $2,000 per year 2 - $2,000 to $2,999 per year 3 - $3,000 to $5,999 per year 4 - $6,000 to $9,999 per year 5 - $10,000 to $14,999 per year 6 - $15,000 to $19,999 per year 7 - $20,OOO to $29,999 per year 3 - $30,OOO and over per year 9 - don't know Comments: Includes income from wages, salaries, social security or retirement benefits, help from relatives, rent from property, etc.
26	16. Estimated Family Income Source: Entered in red by coder opposite question C6 on p.7 of HBQ Codes: 1 to 8 - same as for item 15 9 - no estimate needed (code 1-8 applied for item 15) or no estimate possible Comments: 'Guesstimate' by coder on basis of information supplied by other households in same census block; no estimate needed if response to item 15 received: no estimate possible if no other households in census block reported income.

Column Number(s)	Item Number and Description
27-28	17. Perceived Health Status Source: Question B1 on p.1 of Individual Background Interview (IBI) Codes: 01 - very good 02 - good 03 - fair 04 - poor Comments: Column 27 always zero; only column 28 need be accessed.
29-33	18. Long-term Health Problem or Handicap #1 Source: Questions B2 and B3 on p.1 of IBI Codes: 00000 - no problem *or* handicap 10050 to 99990 - see Comments Comments: Coded according to "A Reason for Visit Classification for Ambulatory Care" developed by National Center for Health Statistics (see Appendix I); 5-digit code with implicit decimal point between 4th and 5th digits; 1st digit, originally alpha, recoded to numeric as follows: 1=S, 2=D, 3=X, 4=T, 5=J, 6=R, 7=A, 9=U; thus a problem reported as "shortness of breath" would have been initially coded as "S415.0" and would appear in the data base as "14150"; if reported as "asthma", the problem would have been initially coded as "D625.0" and would appear in the data base as "26250".
34-33	19. Long-term Health Problem or Handicap #2 source: Questions B2 and B3 on p.1 of IBI Codes: 00000 - no problem or handicap 10050 to 99990 - see Comments Comments: Coded in same manner as item 18; if more than 2 problems or handicaps, priority was given to problems associated with the respiratory *or* circulatory systems.
39-41	20. Number of Visits to Physician or Health Care Facility in Past Year Source: Question B4 on p.2 of IBI Codes: 000 to 097 - number of visits 098 - 98 or more visits 099 - don't know Comments: Includes maternity or preventive visits, visits to eye doctors, psychiatrists, chiropractors; excludes visits to dentists; column 38 always zero; only columns 40-41 need be accessed.

Column Number(s)	Item Number and Description
42-44	21. Usual Commuting Time to Usual Source of Medical Care Source: Question B5 on p..? of IBI Codes: 000 to 998 - commuting time in minutes 999 - don't know
45-47	22. Usual Waiting Time at Usual Source of Medical Care Source: Question B6 on p.2 of IBI Codes: 000 to 998 - waiting time in minutes 999 - don't know
48	23. Major Activity SOURCE: Question B7 on p.3 of IBI Codes: 1 - employed 2 - keeps house 3 - goes to school 4 - pre-school play 5 - retired 6 - unemployed 7 - other Comments: "Other" category often used as activity category for children on vacation from school, e.g., when background interviews were administered during summer months.
49-50	24. Current Tobacco Smoking Status Source: Question B37 on p.15 of IBI Codes: 00 - not applicable 01 - yes 02 - no 09 - don't know Comments: Smoking questions asked only for individuals of age 14 years or older; column 49 always zero; only column 50 need be accessed.
51	25. Current Cigarette Smoking Status Source: Question B38 on p.15 of IBI Codes: 0 - not applicable 1 - yes 2 - no 9 - don't know Comments: Question asked Only for current tobacco smokers.

Column Number(s)	Item Number and Description
52	26. Number of Cigarettes Currently Smoked Per Day Source: Question B39 on p.15 of 131 Codes: 0 - not applicable 1 - under 10 (less than 1/2 pack) 2 - 10 to 15 (between 1/2 and one pack) 3 - 16 to 25 (one pack) 4 - 26 to 35 (1-1/2 packs>) 5 - 36 to 45 (two packs) 6 - 46 or more (more than two packs) 9 - don't know Comments: Question asked only for current cigarette smokers
53	27. Length of Time Smoking Source: Question B40 on p.15 of IBI Codes: 0 - not applicable 1 - less than one year 2 - 1 to 3 years 3 - 4 to 5 years 4 - 6 to 10 years 5 - 11 to 20 years 6 - over 20 years 9 - don't know Comments: Question asked only for current cigarette smokers.
54	28. Current Pipe Smoking Status Source: Question B41 on p.16 of IBI Codes: 0 - not applicable 1 - yes 2 - no 9 - don't know Comments: Question asked only for current tobacco smokers.
55-56	29. Number of Pipefuls Currently Smoked Per Day Source: Question B42 on p.16 of IBI Codes: 00 - not applicable (less than one per day if response to question B41 is "yes.") 01 to 98 - number per day 99 - don't know Comments: Question asked only for current pipe smokers.
57	30. Current Cigar/Cigarillo Smoking Status Source: Question B43 on p.16 of IBI Codes: 0 - not applicable 1 - yes 2 - no 9 - don't know Comments: Question asked only for current tobacco smokers.

Column Number(s)	Item: Number and Description
58-59	31. Number of Cigars/Cigarillos Currently Smoked Per Day Source: Question B44 on p.16 of IBI Codes: 00 - not applicable (less than one per day if response to question 343 is "yes") 01 to 98 - number per day 99 - don't know Comments: Question asked only for current cigar/cigarillo smokers.
60-62	32. Full-time (or Primary) Occupation Source: Question B15 on p.7 of IBI Codes: Per U.S. Census Bureau Index of Occupations (Appendix J) 000 - not applicable 995 or higher - don't know
63-66	33. Full-time (or Primary) Industry Source: Question B17 on p.7 of IBI Codes: Per Office of Management and Budget Standard Industrial Classification Manual (Appendix K) 0000 - not applicable 9999 - don't know
67	34. Full-time Occupation - Mostly Indoors, Outdoors, or in Vehicle? Source: Question B18 on p.7 of IBI Codes: 0 - not applicable 1 - inside 2 - outdoors 3 - combination inside/outdoors 4 - riding in vehicle 9 - don't know
68	35. Full-time Occupation - Paid Sick Leave Days Allowed Per Year Source: Question B21 on p.8 of IBI Codes: 0 - not applicable 1 - none 2 - 5 days or less (one week) 3 - 5 to 10 days (two weeks) 4 - 11 to 15 days (three weeks) 5 - 6 to 22 days (one month) 6 - over 23 days or more than one month (no limit) 9 - don't know

Column Number (s)	Item Number and Description
69-71	36. Part-time (or Second) Occupation Source: Question B27 on p.11 of IBI Codes: Same as item 32
72-75	37. Part-time (or Second) Industry Source: Question B29 on p.11 of IBI Codes: Same as item 33
76	38. Part-time Occupation - Mostly Indoors, Outdoors, or in Vehicle? Source: Question B30 on p.11 of IBI Codes: Same as item 34
77	39. Part-time Occupation - Paid Sick Leave Days Allowed Per Year Source: Question B33 on p.12 of IBI Codes: Same as item 35
78-80	40. Last Occupation for Unemployed Individuals Source: Question B22 on p.9 of IBI Codes: Same as item 32
81-84	41. Last Industry for Unemployed Individuals Source: Question B23 on p.9 of IBI Codes: Same as item 33
85	42. Last Occupation - Mostly Indoors, Outdoors, or in Vehicle? Source: Question B24 on p.9 of IBI Codes: Same as item 34
86	43. Last Occupation - Paid Sick Leave Days Allowed Per Year Source: Question B25 on p.9 of IBI Codes: Same as item 35
87-89	44. Usual Travel Time for Other Activity (Usually Babysitter) Source: Question B11 on p.4 of IBI Codes: 000 - not applicable 001 to 998 - travel time in minutes 999 - don't know Comments: Has already been utilized in making hour-by-hour location assignments (see item 59); no further use anticipated.

Column Number(s)	Item Number and Description
90-92	45. Usual Travel Time for Full-time (of Primary) School Source: Question B14 on p.5 of IBI Codes and Comments: Same as item 44.
93-95	46. Usual Travel Time for Full-time (or Primary1 Occupation Source: Question B20 on p.8 of IBI Codes and Comments: Same as item 42.
96-98	47. Usual Travel Time for Part-time (or Second) Occupation Source: Question B32 on p.12 of IBI Codes and Comments: Same as item 44.

Column Number(s)	Item Number and Description
99-101	48. Usual Travel Time for Part-time (or Second) School Source: Question B35 on p.13 of IBI Codes and Comments: Same as item 44.
102-104	49. Home Location Code Source: Face sheet (p.1) of HBQ Codes: 001 to 336 - legitimate codes (see Appendix L for coordinates) 340 - Background interview only (no followup information) Comments: Should always be equal to 340 or a number between 001 and 336; if not, treat location as unknown; if equal to 340, all hour-by-hour location assignments (item 59) have been set to zero.
105-107	50. Location Code for Other Activities Source: Question B9 on p.4 of IBI Codes: 000 - not applicable 001 to 336 - legitimate codes (see Appendix L for coordinates) 337 - outside the St. Louis urbanized area 333 - location varies 339 - location unknown or refused 340 - background interview only (even though location may be known) Comments: Codes 337, 338, 339 should be treated as unknown for purposes of exposure estimation.
108-110	51. Location Code for Full-time (or Primary) School Source: Question B12 on p.5 of IBI Codes and Comments: Same as item 50.
111-113	52. Location Code for Full-time (or Primary) Occupation Source: Question B16 on p.7 of IBI Codes and Comments: Same as item 50.
114-116	53. Location Code for Part-time (or Second) Occupation Source: Question B23 on p.11 of IBI Codes and Comments: Same as it item 50.
117-119	54. Location Code for Part-time (or Second) School Source: Question B34 on p.13 of IBI Codes and Comments: Same as item 50.

Column Number(s)	Item Number and Description
120-122	55. Dummy Location Code Source: None Codes: Always equal to 000
123-125	56. Location Code for New Work Address (if Changed Jobs) Source: Question I on pp.2-3 of TIF Codes and Comments: Same as item 50.
126-128	57. Location Code for New School Address (if Changed Schools) Source: Question I on pp.2-3 of TIF Codes and Comments: Same as item 50; most frequently used for children going to camp during summer.
129-131	58. Location Code for New Home Address (if Moved) Source: Face sheet of HBQ, Telephone Interview Form (TIF), or Family Control Sheet Codes and Comments: Same as item 50.
132-1475	59. Hour-by-Hour Location Assignments for Followup Period Source: Question B10 on p.4 of IBI, B13 on p.5 of IBI, B19 on p.8 of IBI, B31 on p.12 of IBI, B35 on p.13 of IBI, Questions E and I on pp.2-3 of TIF, Irregular Hours Update on p.4 of TIF Codes: 0 - Location unknown 1 - Use home location code (item 49) 2 - Use location code for other activities (item 50) 3 - Use location code for full-time school (item 51) 4 - Use location code for full-time occupation (item 52) 5 - Use location code for part-time occupation (item 53) 6 - Use location code for part-time school (item 54) 7 - Use location code for new home address (item 58) 8 - use location code for new occupation address (item 56) 9 - Use location code for new school address (item 57)

Column Number(s)	Item Number and Description
132-1475	59. Hour-by-Hour Location Assignments (Cont'd) Comments: This field contains 1344 characters, one for each hour during the followup period (8 weeks x 7 days/week x 24 hours/day); for households coded as week 01 (item 1), followup period began on June 4, 1978; for households coded as week 02, followup period began on June 11, !978; ... for households coded as week 52, followup period began on May 27, 1979; this item, in combination with items 49-58, is all that is required for exposure estimation; individuals with background interviews only (home location code of 340) have already been set to code 0 for all 1344 hours; individuals out of town, in the hospital or nursing home, or not providing irregular hour updates have been set to code 0 for all 24 hours on days when such events occurred.
1476	60. Warning Flag for Irregular Hours Source: Question B10 on p.4 of IBI, B13 on p.5 of IBI, B19 on p.8 of IBI, B31 on p.12 of IBI, B35 on p.13 of IBI, question I on pp.2-3 of TIF Codes: 0 - regular hours for all activities 5 - irregular (variable) hours for one or more activity Comments: Can be used to identify individuals with irregular hours.
1477-1481	61. Warning Flags Indicating Location Assignment Overlaps Source: Same as item 60 Codes: Column (1477-1481) indicates the location overwritten; value of flag (2-6, 8, 9) indicates the location responsible for the overwrite and actually stored in the hour-by-hour location assignments (item 59); Column 1477 or flag value 2 - other activity location Column 1478 or flag value 3 - full-time school location Column 1479 or flag value 4 - full-time occupation location

Column Number(s)	Item Number and Description
1477-1481	61. Warning Flags Indicating Location Assignment Overlaps (Cont'd) Column 1480 or flag value 5 - part-time occupation location Column 1481 or flag value 6 - part-time school location Flag value 8 - new occupation location Flag value 9 - new school location Comments: Can be used to access frequency of location assignment overlaps.
1482	62. Dummy Flag Source: None Codes: Always zero Comments: Of no use.
1483	63. Warning Flag for Hours with Location Assignment Unknown Source: Question E and I on pp.2-3 of TIF Codes: 0 - location assignment known for all hours or background interview only 1 - one or more hour with location assignment unknown Comments: Can be used to identify individuals (other than those with background interview only) having one or more followup hours with location assignment unknown (code 0 for item 59).
1484	64. Warning Flag for Hours Approximated on Followup Source: Question I on pp.2-3 of TIF Codes: 0 - no hours approximated 1 - one or more hour approximated Comments: Can be used to identify individuals with irregular hours who did not report hours for selected weeks but whose hours were guessed by coder on the basis of information reported for other weeks; see item 76 for further details.
1485	65. Warning Flag for Lost to Followup Source: Question I on pp.2-3 of TIF Codes: 0 - followup completed or background interview only 1 - lost to followup Comments: Can be used to identify individuals (other than those with background interview only) lost to followup; see item 76 for further details.

Column Number(s)	Item Number and Description
1486	66. Warning Flag for Unknown Travel Time Source: See items 44-43 Codes: 0 - all travel times known 2 - unknown travel time for other activity 3 - unknown travel time for full-time school 4 - unknown travel time for full-time occupation 5 - unknown travel time for part-time occupation 6 - unknown travel time for part-time school Comments: Unknown travel times set to zero in calculation of hour-by-hour location assignments (item 59).
1487	67. Warning Flag for Saturday p.m. to Sunday a.m. Overlap Source: Same as item 59 Codes: 0 - no overlap 1 - overlap Comments: Individuals with activities on Saturday evenings and into Sunday mornings required special treatment since their such hours overlapped survey weeks; has already been utilized in making hour-by-hour location assignments (see item 59); no further use anticipated.
1488	68. Warning Flag for Approximated Location Source: See items 49-58 Codes: 0 - no locations approximated 1 - one or more location approximated Comments: Can be used to identify individuals with one or more activity outside the area covered by legitimate location codes; for each activity outside the coded area but within 5 miles of it, the activity was given the nearest code and the flag value was set to one.
1489	69. Warning Flag for Invalid Location Source: See items 50-58 Codes: 0 - all locations valid 1 - one or more location invalid Comments: Can be used to identify individuals with location code of 337, 338, or 339 (see item 50) for one or more activity; such location codes will result in unknown exposure estimates for associated activities; individuals with background interview only and having activities away from home also coded as “1”.

Column Number(s)	Item Number and Description
1490	70. Edit Field Comments: Of no use.
1491-1504	71. Activity Restrictions for Followup Period 1 (Days 1-14) Source: Questions E.1.a-d on pp.2-3 of TIF Codes: 0 - no activity restriction 1 - in hospital 2 - in nursing home 3 - in bed at home 4 - confined to the house 5 - cut down on usual activities Comments: This field contains 14 characters, one for each day of Followup Period 1; for households coded as week 01 (item 1), Followup Period 1 began on June 4, 1978; for households coded as week 02, Followup Period 1 began on June 11, 1978; . . . for households coded as week 52, Followup Period 1 began on May 27, 1979.

Column Number(s)	Item Number and Description

1505-1518 72. Work/School Absenteeism for Followup Period 1

Source: Question E2 on pp.2-3 of TIF

Codes:
0 - no absenteeism
1 - missed work
2 - missed school
3 - missed work around the house
4 - missed work and school
5 - missed work and work around the house
6 - missed school and work around the house
7 - missed work, school, and work around the house

Comments: Same as item 71; reporting of missed work around the house not considered reliable.

1519-1532 73. Physician Contacts for Followup Period 1

Source: Questions F.1.a-b on pp.2-3 of TIF

Codes:
0 - no physician contact
1 - physician seen in office
2 - physician seen in hospital outpatient department
3 - physician seen in hospital emergency room
4 - physician seen in private clinic
5 - physician seen in neighborhood health center
6 - physician seen in public health clinic
7 - physician seen in own home
8 - physician seen elsewhere
9 - physician phoned for medical advice

Comments: Same as item 71.

1532-1546 74. Ancillary Services during Physician Visits for Followup Period 1

Sources: Question F2 on pp.2-3 of TIF

Codes:
0 - no ancillary services
1 - shot
2 - x-ray
3 - test
4 - shot and x-ray
5 - shot and test
6 - x-ray and test
7 - shot, x-ray, and test

Comments: Same as item 71.

1547-1560 75. Prescriptions Filled or Renewed for Followup Period 1

Source: Question G.1.a on pp.2-3 of TIF

Codes:
0 - no prescription filled or renewed
1 - prescription filled or renewed

Comments: Same as item 71.

Column Number(s)	Item Number and Description
1561-1574	76. Activity Changes for Followup Period 1 Source: Question H.1.a-b on pp.2-3 of TIF Codes: 0 - no activity changes 1 - out of town 2 - home from work for vacation or holiday 3 - home from school for vacation or holiday 4 - home from work and school for vacation or holiday 5 - individual lost to followup 6 - individual and rest of household lost to followup 7 - irregular hours not reported but guessed from other reported hours 8 - irregular hours not reported but guessed on basis of limited information 9 - irregular hours not reported and not guessed Comments: Same as item 71; has already been utilized in making hour-by-hour location assignments (see item 59); no further use anticipated.
1575-1579	77. Reason Associated with Items 71 and 72 Source: Question E.1.e on pp.2-3 of TIF Codes and Comments: See item 18; only one reason coded for each followup period; priority given to problems associated with respiratory or circulatory systems if more than one reported.
580-1584	78. Reason Associated with Items 73 and 74 Source: Question F.1.c on pp.2-3 of TIF Codes and Comments: Same as item 77.
1585-1589	79. Reason Associated with Item 75 Source: Question G.1.b on pp.2-3 of TIF Codes and Comments: Same as item 77.
1590-1603	80. Activity Restrictions for Followup Period 2 (Days 15-28) Source and Codes: Same as item 71 Comments: This field contains 14 characters, one for each day of Followup Period 2; for households coded as week 01 (item 1), Followup Period 2 began on June 18, 1978; for households coded as week 02, Followup Period 2 began on June 25, 1978; ... for households coded as week 52, Followup Period 2 began on June 10, 1979.

Column Number(s)	Item Number and Description
1604-1617	81. Work/School Absenteeism for Followup Period 2 Source and Codes: Same as item 72 Comments: Same as item 80.
1618-1631	82. Physician Contacts for Followup Period 2 Source and Codes: Same as item 73 Comments: Same as item 80.
1632-1645	83. Ancillary Services during Physician Visits for Followup Period 2 Source and Codes: Same as item 74 Comments: Same as item 80.
1646-1659	84. Prescriptions Filled or Renewed for Followup Period 2 Source and Codes: Same as item 75 Comments: Same as item 80.
1660-1673	85. Activity Changes for Followup Period 2 Source and Codes: Same as item 76 Comments: Same as item 30.
1674-1678	86. Reason Associated with Items 80 and 81 Source: Same as item 77 Codes and Comments: Same as item 77.
1679-1683	87. Reason Associated with Items 82 and 83 Source: Same as item 73 Codes and Comments: Same as item 77.
1684-1688	88. Reason Associated with Item 84 Source: Same as item 79 Codes and Comments: Same as item 77.
1689-1702	89. Activity Restrictions for Followup Period 3 (Days 29-42) Source and Codes: Same as item 71 Comments: This field contains 14 characters, one for each day of Followup Period 3; for households coded as week 01 (item 1), Followup Period 3 began on July 2, 1978; for households coded as week 02, Followup Period 3 began on July 9, 1978; ... for households coded as week 52, Followup Period 3 began on June 24, 1979.

Column Number(s)	Item Number and Description
1703-1716	90. Work/School Absenteeism for Followup Period 3 Source and Codes: Same as item 72 Comments: Same as item 89.
1717-1730	91. Physician Contacts for Followup Period 3 Source and Codes: Same as item 73 Comments: Same as item 89.
1731-1744	92. Ancillary Services during Physician Visits for Followup Period 3 Source and Codes: Same as item 74 Comments: Same as item 89.
1745-1758	93. Prescriptions Filled or Renewed for Followup Period 3 Source and Codes: Same as item 75 Comments: Same as item 89.
1759-1772	94. Activity Changes for Followup Period 3 Source and Codes: Same as item 76 Comments: Same as item 89.
1773-1777	95. Reason Associated with Items 89 and 90 Source: Same as item 77 Codes and Comments: Same as item 77.
1778-1782	96. Reason Associated with Items 91 and 92 Source: Same as item 78 Codes and Comments: Same as item 77.
1783-1787	97. Reason Associated with Item 93 Source: Same as item 79 Codes and Comments: Same as item 77.
1788-1801	98. Activity Restrictions for Followup Period 4 (Days 43-56) Source and Codes: Same as item 71 Comments: This field contains 14 characters, one for each day of followup Period 4; for households coded as week 01 (item 1), Followup Period 4 began on July 16, 1978; for households coded as week 02, Followup Period 4 began on July 23, 1978; ... for households coded as week 52, Followup Period 4 began on July 8, 1979.

Column Number(s)	Item Number and Description
1802-1815	99. Work/School Absenteeism for Followup Period 4 Source and Codes: Same as item 72 Comments: Same as item 98.
1816-1829	100. Physician Contacts for Followup Period 4 Source and Codes: Same as item 73 Comments: Same as item 98.
1830-1843	101. Ancillary Services during Physician Visits for Followup Period 4 Source and Codes: Same as item 74 Comments: Same as item 98.
1844-1857	102. Prescriptions Filled or Renewed for Followup Period 4 Source and Codes: Same as item 75 Comments: Same as item 98.
1858-1871	103. Activity Changes for Followup Period 4 Source and Codes: Same as item 76 Comments: Same as item 98.
1872-1876	104. Reason Associated with Items 98 and 99 Source: Same as item 77 Codes and Comments: Same as item 77.
1877-1881	105. Reason Associated with Items 100 and 101 Source: Same as item 78 Codes and Comments: Same as item 77.
1882-1886	106. Reason Associated with Item 102 Source: Same as item 79 Codes and Comments: Same as item 77.
1887	107. Status Regarding Supplemental Data Collection Source: Various control cards/sheets Codes: 1 - household eligible for supplemental data collection (completed basic survey through all 4 followup periods), supplemental data collected 2 - household eligible for supplemental data collection, supplemental data not collected 3 - household not eligible for supplemental data collection Comments: All individuals with codes "2" or "3" will have blanks in subsequent columns.

Column Number(s)	Item Number and Description
1888	103. Type of Air Conditioning at Time of Basic Survey Source: Question A1 on Supplemental Data Form (SDF) Codes: 1 - central air conditioning 2 - window air conditioning 3 - both types 4 - neither type 9 - don't know/refused
1889	109. Fuel Used Most to Heat Home at Time of Basic Survey Source: Question A2 on SDF Codes: 1 - electricity 2 - oil 3 - gas 4 - other 9 - don't know/refused
1890	110. Fuel Used Most for Heating Water at Time of Basic Survey Source: Question A3 on SDF Codes: Same as item 109.
1891	111. Fuel Used Most for Cooking at Time of Basic Survey Source: Question A4 on SDF Codes: Same as item 109.
1892	112. Buying or Renting Home at Time of Basic Survey Source: Question AS on SDF Codes: 1 - buying 2 - renting 3 - other 9 - don't know/refused Comments: Some respondents whose homes were completely paid for were coded as "buying"; others were coded as "other"; codes 1 and 3 should be combined and treated as buying/bought home.
1893-1894	113. Years Living at Home at Time of Basic Survey Source: Question A6 on SDF Codes: 00 - less than one year 01 to 98 - number of years 99 - don't know/refused Comments: Can be used to determine individuals with stable residential histories for purposes of historical exposure estimation.
1895-1896	114. Years Living in St. Louis Metro Area at Time of Basic Survey Source: Question A7 on SDF Codes: Same as item 113.

Column Number(s)	Item Number and Description
1897	115. Type of Place Lived Most Often Up to Age 14 Source: Question A8 on SDF Codes: 1 - metropolitan area (over 1 million population) 2 - large city (over 100,000) 3 - smaller city (over 10,000) 4 - small town/rural area 9 - don't know/refused Comments: Responses not considered reliable.
1898	116. Type of Place Lived Most Often Since Age 14 Source: Question A9 on SDF Codes and Comments: Same as item 115; individuals of age 13 or under coded as blank.
1899	117. Questions B1-7 Asked Regarding Full-time (or Primary) Occupation Source: Line labeled J1 Under Part B on SDF Answer Sheet Codes: 0 - questions not asked (no full-time occupation) 1 - questions asked Comments: Individuals with Code "0" Have Blank Responses for items 119-128.
1900	118. Questions B1-7 Asked Regarding Part-time (or Second) Occupation Source: Line labeled J2 under Part B on SDF Answer Sheet Codes: 0 - questions not asked (no part-time occupation) 1 - questions asked Comments: Individuals with code "0" have blank responses for items 129-138.
1901-1902	119. Years Worked for Employer at Time of Basic Survey for Full-time Occupation Source: Question B1 for J1 on SDF Codes: 00 - less than 1 year 01 to 98 - number of years 99 - don't know/refused Comments: Can be used to determine individuals with stable job histories for purposes of historical exposure estimation.
1903	120. Exposure to Irritants in Full-time Occupation Source: Question 32 for J1 on SDF Codes: 1 - all/most of the time 2 - sometimes/occasionally 3 - little/none of the time
1904	121. Membership in Labor Union at Time of Basic Survey for Full-time Occupation Source: Question B3 for J1 on SDF Codes: 1 - yes 2 - no 9 - don't know/refused

Column Number(s)	Item Number and Description
1905	122. Paid Vacation Days Available at Time of Basic Survey for Full-time Occupation Source: Question B4 for J1 on SDF Codes: 1 - Yes 2 - no 9 - don't know/refused
1906	123. Paid Sick Leave Days Available at Time of Basic Survey for Full-time Occupation Source: Question B5 for J1 on SDF Codes: 1 - yes 2 - no 9 - don't know/refused
1907	124. Would Money Be lost from Missing Full-time Occupation at Time of Basic Survey? Source: Question B6 for J1 on SDF Codes: 1 - yes 2 - no 9 - don't know
1908-1910	125. Amount of Money Lost from Missing Full-time Occupation at Time of Basic Survey Source: Question B6 for J1 on SDF Codes: Blank - not applicable (item 124 coded 2 or 9) 001 to 998 - amount lost in dollars 999 - don't know/refused
1911-1914	126. Take-home Pay from Full-time Occupation at Time of Basic Survey Source: Question B7 for J1 on SDF Codes: 0000 - not paid (e.g., voluntary work) 0001 to 9998 - amount paid in dollars 9999 - don't know/refused
1915	127. Pay period corresponding to Item 126 Source: Question B7 for J1 on SDF Codes: 1 - daily 2 - weekly 3 - every 2 weeks 4 - twice a month 5 - monthly 6 - three times a month 7 - hourly 8 - other (usually commission or self-employed) 9 - don't know/refused

Column Number(s)	Item Number and Description
1916	128. Weekly Take-home Pay Categories For full-time Occupation at Time of Basic Survey Source: Question B7 for J1 on SDF Codes: 0 - not applicable (response given for Items 126 and 127) 1 - under $50 2 - $50 to $99 3 - $100 to $199 4 - $200 to $299 5 - $300 to $399 6 - $400 and over 9 - don't know/refused
1917-1918	129. Years Worked for Employer at Time of Basic Survey for Part-time Occupation Source: Question B1 for J2 on SDF Codes and Comments: Same as item 119.
1919	130. Exposure to Irritants in Part-time Occupation Source: Question B2 for J2 on SDF Codes: Same as item 120
1920	131. Membership in Labor Union at Time of Basic Survey for Part-time Occupation Source: Question B3 for J2 on SDF Codes: Same as item 121
1921	132. Paid Vacation Days Available at Time of Basic Survey for Part-time Occupation Source: Question B4 for J2 on SDF Codes: Same as item 122
1922	133. Paid Sick Leave Days Available at Time of Basic Survey for Part-time Occupation Source: Question B5 for J2 on SDF Codes: Same as item 123
1923	134. Would Money Be Lost from Missing Part-time Occupation at Time of Basic Survey? Source: Question B6 for J2 on SDF Codes: Same as item 124
1924-1926	135. Amount of Money Lost from Missing Part-time Occupation at Time of Basic Survey Source: Question B6 for J2 on SDF Codes: Same as item 125; blank if item 134 coded 2 or 9

Column Number(s)	Item Number and Description
1927-1930	136. Take-home Pay from Part-time Occupation at Time of Basic Survey Source: Question B7 for J2 on SDF Codes: Same as item 126
1931	137. Pay Period Corresponding to Item 134 Source: Question B7 for J2 on SDF Codes: Same as item 127
1932	138. Weekly Take-home Pay Categories for Part-time Occupation at Time of Basic Survey Source: Question B7 for J2 on SDF Codes: Same as item 128
1933	139. Questions B8-11 Asked Regarding Principal Occupation Prior to Retirement Source: Item B8 on SDF answer sheet--whether circled or not Codes: 0 - item not circled, questions not asked (not retired) 1 - item circled, questions asked Comments: Questions asked for individuals indicating "Retired" as major activity at time of basic survey (item 23) or age 65 years and over and not working at time of basic survey; individuals with code "0" have blank responses to items 140-144.
1934-1936	140. Principal Occupation Prior to Retirement Source: Question B8 on SDF Codes: Per U.S. Census Bureau Index of Occupations (Appendix J1 blank - not applicable 995 or higher - don't know/refused
1937-1940	141. Principal Industry Prior to Retirement Source: Question B8 on SDF Codes: Per Office of Management and Budget Standard Industrial Classification Manual (Appendix K) blank - not applicable 9999 - don't know/refused
1941-1942	142. Years in Principal Occupation Prior to Retirement Source: Question B9 on SDF Codes: blank - not applicable 00 - less than one year 01 to 98 - number of years 99 - don't know/refused

Column Number(s)	Item Number and Description

1943 143. Principal Occupation Prior to Retirement - Mostly Indoors, Outdoors, or in Vehicle?

Source: Question B10 on SDF

Codes: blank - not applicable
1 - inside
2 - outdoors
3 - combination inside/outdoors
4 - riding in vehicle
9 - don't know/refused

1944 144. Exposure to Irritants in Principal Occupation Prior to Retirement

Source: Question B11 on SDF

Codes: blank - not applicable
1 - all/most of the time
2 - sometimes/occasionally
3 - little/none of the time
9 - don't know/refused

1945 145. Portion of Family Income During Year of Basic Survey from Sources Other than Jobs

Source: Question C1 on SDF

Codes: 1 - less than 2%
2 - 2 to 10%
3 - 10 to 25%
4 - 25 to 50%
5 - over 50%
9 - don't know/refused

1946 146. Portion of Family Income from Retirement Benefits

Source: Question C2a on SDF

Codes: 0 - None
1 - less than 2%
2 - 2 to 10%
3 - 10 to 25%
4 - 25 to 50%
5 - over 50%
6 - portion greater than zero but not known/refused
9 - don't know/refused

1947 147. Portion of Family Income from Interest, Dividends, Rent, Estates, Trusts, or Capital Gains

Source: Question C2b on SDF

Codes: Same as item 146

1943 148. Portion of Family Income from Unemployment or Strike Benefits

Source: Question C2c on SDF

Codes: Same as item 146

Column Number(s)	Item Number and Description
1949	149. Portion of Family Income from Disability Benefits Source: Question C2d on SDF Codes: Same as item 146
1950	150. Portion of Family Income from Welfare Payments or other Public Assistance Source: Question C2e on SDF Codes: Same as item 146
1951	151. Portion of Family Income from Alimony, Child Support, or Regular Contributions from Persons Not Part of Household Source: Question C2f on SDF Codes: Same as i tern 146
1952-1953	152. Usual Charge by Doctor for Office Visit Source: Question C3 on SDF Codes: 00 to 98 - amount charged in dollars 99 - don't know/refused
1954	153. Portion of Office Visit Charge Usually Paid by Any Health Insurance Source: Question C4 on SDF Codes: 1 - all or most 2 - some or about half 3 - little or none 4 - never been to doctor/not applicable 9 - don't know/refused
1955-1957	154. Current Height Without Shoes Source: Question D1 on SDF Codes: 000 to 998 - height in feet (column 1955) and inches (columns 1956-1957) 999 - don't know/refused
1959-1960	155. Current Weight Without Shoes and Clothes (Normal Weight if Pregnant) Source: Question D2 on SDF Codes: 000 to 998 - weight in pounds 999 - don't know/refused
1961-1963	156. Most Ever Weighed (Not Counting Pregnancies) Source: Question D3 on SDF Codes: Same as item 155
1964-1965	157. Years between Current and Maximum Weight (if Different) Source: Question D3 on SDF Codes: blank - not applicable 00 - less than one year 01 to 98 - number of years 99 - don't know/refused

Column Number(s)	Item Number and Description
1966	158. Doctor Usually Seen at Least Once a Year Source: Question D4 on SDF Codes: 1 - yes 2 - no 9 - don't know/refused
1967	159. Regular Program of Physical Exercise Followed Source: Question D5 on SDF Codes: 1 - yes 2 - no 9 - don't know/refused Comments: Regular program of exercise is considered to be one which an individual follows at least 3 days per week throughout the year.
1968-1969	160. Years Following Regular Program of Physical Exercise Source: Question D5 on SDF Codes: blank - not applicable 00 - less than one year 01 to 98 - number of years 99 - don't know/refused
1970	161. Frequency of Strenous Activities (Outside of Any Job) Source: Question D6 on SDF Codes: 1 - frequently 2 - occasionally 3 - rarely 9 - don't know/refused
1971-1972	162. Hours Spent Outdoors in a Typical Summer Week (Outside of Any Job) Source: Quest ion D7 on SDF Codes: 00 to 98 - number of hours 99 - don't know/refused
1973-1974	163. Hours Spent Watching or Sitting in Front of the TV in a-Typical Summer Week Source: Question D3 on SDF Codes: Same as item 162
1375	164. Question Asked Regarding Length of Long-term Health Problem(s) Source: Lines labeled P1/P2 Opposite Question D9 on SDF Answer Sheet Codes: 0 - no questions asked (no long-term health problems) 1 - questions asked Comments: Individuals with code "0" have blank responses for item 165; long-term health problems are those coded as items 8-19.

Column Number(s)	Item Number and Description
1976-1977	165. Years with Long-term Health Problem at Time of Basic Survey Source: Question D9 on SDF Codes: blank - not applicable 00 - less than one year 01 to 98 - number of years 99 - don't know/refused
1978	166. Questions Asked Regarding Past Cigarette Smoking Source: Item D10 on SDF answer sheet--whether circled or not Codes: 0 - item not circled, questions not asked (current cigarette smoker) 1 - item circled, questions asked Comments: Questions asked for individuals aged 14 years or over who were not smoking cigarettes at time of the basic survey; individuals with code "0" have blank responses to items 167-169.
1979-1980	167. Number of Cigarettes Usually Smoked Per Day Source: Question D13 on SDF Codes: blank - none 00 - less than one cigarette 01 to 98 - number of cigarettes 99 - don't know/refused
1981-1982	168. Age Started Smoking Cigarettes Source: Question D10 on SDF Codes: blank - never smoked cigarettes 01 to 98 - age in years 99 - don't know/refused
1983-1984	169. Age Stopped Smoking Cigarettes Source: Question D10 on SDF Codes: Same as item 168
1985	170. Questions Asked Regarding Past Cigar/Pipe Smoking Source: Question D11 on SDF answer sheet--whether circled or not Codes: 0 - item not circled, questions not asked (current cigar/pipe smoker) 1 - item circled, questions asked Comments: Questions asked for individuals aged 14 years or over who were not smoking cigars or pipes at the time of the basic survey; individuals with code "0" have blank responses to items 171-174.

Column Number(s)	Item Number and Description
1986-1987	171. Number of Cigars Usually Smoked Per Day Source: Question D11 on SDF Codes: blank - none 00 - less than one cigar 01 to 98 - number of cigars 99 - don't know/refused
1988-1989	172. Number of Pipefuls Usually Smoked Per Day Source: Question D11 on SDF Codes: blank - none 00 - less than one pipeful 01 to 98 - number of pipefuls 99 - don't know/refused
1990-1991	173. Age Started Smoking Cigars/Pipes Source: Question D11 on SDF Codes: blank - never smoked cigars or pipes 01 to 98 - age in years 99 - don't know/refused
1992-1993	174. Age Stopped Smoking Cigars/Pipes Source: Question D11 on SDF Codes: Same as item 173
1994	175. Average Number of Hot Meals Eaten Per Day Source: Question E1 on SDF Codes: 0 to 8 - number of meals 9 - don't know/refused
1995	176. Frequency of Dieting to Keep Weight Down Source: Question E2 on SDF Codes: 1 - regularly/often 2 - sometimes/occasionally 3 - rarely/never 9 - don't know/refused
1996	177. Frequency of Eating Cured Heats Source: Question E3a on SDF Codes: 1 - almost daily 2 - one to three times a week 3 - one to three times a month 4 - less often 9 - don't know/refused
1997	178. Frequency of Eating Candies, Pastries, or Other Sweets Source: Question E3b on SDF Codes: Same as item 177
1998	179. Frequency of Eating Salty Snacks Source: Question E3c on SDF Codes: Same as item 177

Column Number(s)	Item Number and Description
1999-2000	180. Frequency of Drinking Caffeinated Drinks Source: Question E4a on SDF Codes: 00 to 98 - number of times per day 99 - don't know/refused
2001-2002	181. Frequency of Drinking Drinks with Sugar Source: Question E4b on SDF Codes: Same as item 180
2003	182. Frequency of Drinking Alcoholic Beverages Source: Question E5 on SDF Codes: blank - not applicable (less than 14 years of age) 1 - almost daily 2 - several times a week 3 - several times a month 4 - several times a year 5 - less often 9 - don't know/refused Comments: Items 182-185 asked only for individuals aged 14 years or over; individuals reporting that they never drink coded as "5" for item 182 and blanks for items 183-185.
2004	183. Type of Alcoholic Beverage Drank Most Frequently Source: Question E6 on SDF Codes: blank - not applicable (less than 14 years of age or never drinks) 1 - beer 2 - wine 3 - liquor 9 - don't know/refused Comments: Response used in forming questions for items 184-185.
2005-2006	184. Number of Alcoholic Beverages Drank Per 24 Hours Source: Question E7 on SDF Codes: blank - not applicable (less than 14 years of age or never drinks) 01 to 98 - number of beverages per 24 hours 99 - don't know/refused
2007	185. Frequency of Drinking Enough to Become Intoxicated Source: Question E8 on SDF Codes: blank - not applicable (less than 14 years of age or never drinks) 1 - almost daily 2 - weekly 3 - monthly 4 - less often 5 - never 9 - don't know/refused