Table 4.12

Reduced Form Equation

MDOC 55 -- 10 11 13 14 21 24 54 25 27 41

VAR	B	T
X10-NONW	50.447	1.1020
X11-MAGE	1.3513	0.50439
X13-IN69	0.61649D-02	0.86708
X14-EDUC	1.9399	1.3418
X21-DENS	161.53	0.26107
X24-COLD	-0.12806	-0.53850
X54-CI68	0.45771	1.4919
X25-XPRO	0.22302D-01	1.4136
X27-XCAR	0.22804D-02	0.72804
X41-6SFT	0.24002D-01	1.9402
CONSTANT	-1691.3	-2.7124

R-SQUARE= 0.3877
SSR= 0.1062D+06 DF= 49

Table 4.13

Total Mortality

MO70 1 -- 61 10 11 21 24 54 25 27 41

VAR	B	T
X61-DOCH	-0.53031D-01	-4.7501
X10-NONW	5.6092	5.0448
X11-MAGE	0.66172	12.360
X21-DENS	31.910	2.4494
X24-COLD	0.14370D-01	3.1093
X54-CI68	0.21896D-01	3.0758
X25-XPRO	0.19325D-02	3.7525
X27-XCAR	-0.82907D-04	-1.5347
X41-6SFT	0.36251D-03	1.5938
CONSTANT	-78.675	-3.7169

R-SQUARE= 0.8195
SSR= 49.74 DF= 50

MO70 1 -- 61 10 11 21 24 54 25 27 41 31 57 60

VAR	B	T
X61-DOCH	-0.52797D-01	-4.3493
X10-NONW	5.6276	4.5620
X11-MAGE	0.65893	11.540
X21-DENS	31.772	2.3469
X24-COLD	0.14436D-01	2.9089
X54-CI68	0.21968D-01	2.8120
X25-XPRO	0.19196D-02	3.5522
X27-XCAR	-0.79431D-04	-1.3612
X41-6SFT	0.39783D-03	1.4508
X31-NO69	1.6457	0.35799
X57-SO70	-0.31302D-02	-0.34850
X60-PA70	0.10744D-02	0.20059
CONSTANT	-79.296	-3.5115

R-SQUARE= 0.8205
SSR= 49.48 DF= 47

Table 4.14

Vascular Disease

VA70 2 -- 61 10 11 21 24 54 25 27 41

VAR	B	T
X61-DOCH	-0.70223D-02	-2.2186
X10-NONW	0.37954	1.2040
X11-MAGE	0.10936	7.2050
X21-DENS	1.7016	0.46071
X24-COLD	0.17695D-02	1.3504
X54-CI68	0.30001D-02	1.4865
X25-XPRO	0.18542D-03	1.2699
X27-XCAR	-0.13460D-04	-0.87885
X41-6SFT	-0.84316D-04	-1.3075
CONSTANT	-5.5672	-0.92769

R-SQUARE= 0.5657
SSR= 3.998 DF= 50

VA70 2 -- 61 10 11 21 24 54 25 27 41 31 57 60

VAR	B	T
X61-DOCH	-0.88088D-02	-2.6671
X10-NONW	0.63816	1.9014
X11-MAGE	0.11671	7.5127
X21-DENS	2.7540	0.74768
X24-COLD	0.20961D-02	1.5524
X54-CI68	0.42167D-02	1.9838
X25-XPRO	0.23976D-03	1.6306
X27-XCAR	-0.17117D-04	-1.0781
X41-6SFT	-0.18874D-04	-0.25297
X31-NO69	-0.86198	-0.68915
X57-SO70	-0.38911D-02	-1.5922
X60-PA70	-0.63071D-03	-0.43278
CONSTANT	-8.3456	-1.3583

R-SQUARE= 0.6022
SSR= 3.663 DF= 47

Table 4.15

Heart Disease

HA70 3 -- 61 10 11 21 24 54 25 27 41

VAR	B	T
X61-DOCH	-0.22340D-01	-3.4204
X10-NONW	1.7509	2.6917
X11-MAGE	0.29627	9.4592
X21-DENS	9.3338	1.2247
X24-COLD	0.43566D-02	1.6113
X54-CI68	0.13129D-01	3.1525
X25-XPRO	0.66878D-03	2.2197
X27-XCAR	-0.17969D-04	-0.56857
X41-6SFT	0.42380D-03	3.1849
CONSTANT	-35.183	-2.8412

R-SQUARE= 0.7517
SSR= 17.03 DF= 50

HA70 3 -- 61 10 11 21 24 54 25 27 41 31 57 60

VAR	B	T
X61-DOCH	-0.19800D-01	-2.9132
X10-NONW	1.3603	1.9694
X11-MAGE	0.28122	8.7965
X21-DENS	7.4177	0.97861
X24-COLD	0.43300D-02	1.5583
X54-CI68	0.10868D-01	2.4847
X25-XPRO	0.57165D-03	1.8893
X27-XCAR	-0.22071D-05	-0.67551D-01
X41-6SFT	0.39645D-03	2.5822
X31-NO69	4.8498	1.8842
X57-SO70	0.82843D-03	0.16473
X60-PA70	0.14052D-02	0.46855
CONSTANT	-32.823	-2.5960

R-SQUARE= 0.7737
SSR= 15.51 DF= 47

Table 4.16

Pneumonia and Influenza

PN70 4 -- 61 10 11 21 24 54 25 27 41

VAR	B	T
X61-DOCH	-0.12925D-02	-1.2098
X10-NONW	0.18381	1.7274
X11-MAGE	0.19920D-01	3.8880
X21-DENS	2.9692	2.3816
X24-COLD	0.13030D-02	2.9460
X54-CI68	0.61524D-03	0.90310
X25-XPRO	0.70654D-04	1.4336
X27-XCAR	-0.11329D-05	-0.21914
X41-6SFT	0.38010D-05	0.17462
CONSTANT	-3.0478	-1.5046

R-SQUARE= 0.4108
SSR= 0.4556 DF= 50

PN70 4 -- 61 10 11 21 24 54 25 27 41 31 57 60

VAR	B	T
X61-DOCH	-0.38115D-03	-0.37068
X10-NONW	0.13820	1.3226
X11-MAGE	0.16194D-01	3.3484
X21-DENS	2.6677	2.3264
X24-COLD	0.10428D-02	2.4807
X54-CI68	0.61816D-03	0.93415
X25-XPRO	0.46048D-04	1.0060
X27-XCAR	-0.10027D-05	-0.20286
X41-6SFT	-0.49261D-05	-0.21208
X31-NO69	0.55215	1.4180
X57-SO70	-0.47719D-03	-0.62722
X60-PA70	0.14272D-02	3.1456
CONSTANT	-2.1183	-1.1075

R-SQUARE= 0.5409
SSR= 0.3550 DF= 47

Table 4.17

Emphysema and Bronchitis

EM70 5 -- 61 10 11 21 24 54 25 27 41

VAR	B	T
X61-DOCH	-0.13851D-02	-2.3989
X10-NONW	0.22533D-01	0.39184
X11-MAGE	0.49451D-02	1.7859
X21-DENS	0.92106D-01	0.13670
X24-COLD	0.18996D-03	0.79468
X54-CI68	0.53419D-04	0.14509
X25-XPRO	0.65261D-04	2.4502
X27-XCAR	-0.95199D-05	-3.4072
X41-6SFT	-0.17700D-04	-1.5046
CONSTANT	-0.90721	-0.82870

R-SQUARE= 0.3559
SSR= 0.1331 DF= 50

EM70 5 -- 61 10 11 21 24 54 25 27 41 31 57 60

VAR	B	T
X61-DOCH	-0.13761D-02	-2.2419
X10-NONW	0.46333D-01	0.74278
X11-MAGE	0.51065D-02	1.7687
X21-DENS	0.17013	0.24853
X24-COLD	0.14392D-03	0.57352
X54-CI68	0.24382D-03	0.61719
X25-XPRO	0.66687D-04	2.4404
X27-XCAR	-0.10427D-04	-3.5335
X41-6SFT	-0.13868D-04	-1.0001
X31-NO69	-0.11515	-0.49533
X57-SO70	-0.47320D-03	-1.0419
X60-PA70	0.29115D-03	1.0750
CONSTANT	-0.96354	-0.84382

R-SQUARE= 0.3876
SSR= 0.1265 DF= 47

Table 4.18

Cirrhosis

CI70 6 -- 61 10 11 21 24 54 25 27 41

VAR	B	T
X61-DOCH	0.51923D-04	0.65875D-01
X10-NONW	0.22721	2.8943
X11-MAGE	0.17323D-01	4.5831
X21-DENS	2.1612	2.3497
X24-COLD	0.52436D-03	1.6070
X54-CI68	0.61784D-03	1.2293
X25-XPRO	0.73982D-04	2.0347
X27-XCAR	-0.77437D-05	-2.0303
X41-6SFT	-0.11197D-04	-0.69726
CONSTANT	-2.3252	-1.5559

R-SQUARE= 0.6258
SSR= 0.2479 DF= 50

CI70 6 -- 61 10 11 21 24 54 25 27 41 31 57 60

VAR	B	T
X61-DOCH	0.31818D-03	0.37741
X10-NONW	0.18469	2.1558
X11-MAGE	0.16077D-01	4.0543
X21-DENS	1.9819	2.1080
X24-COLD	0.49340D-03	1.4315
X54-CI68	0.39600D-03	0.72987
X25-XPRO	0.65122D-04	1.7351
X27-XCAR	-0.67906D-05	-1.6755
X41-6SFT	-0.19734D-04	-1.0362
X31-N069	0.24209	0.75825
X57-S070	0.51883D-03	0.83173
X60-PA70	0.69311D-04	0.18632
CONSTANT	-1.9446	-1.2399

R-SQUARE= 0.6399
SSR= 0.2387 DF= 47

Table 4.19

Kidney Disease

NE70 7 -- 61 10 11 21 24 54 25 27 41

VAR	B	T
X61-DOCH	-0.67302D-03	-2.9834
X10-NONW	0.90661D-01	4.0353
X11-MAGE	0.34110D-02	3.1531
X21-DENS	0.75879	2.8825
X24-COLD	0.73289D-04	0.78479
X54-CI68	0.10132D-03	0.70439
X25-XPRO	0.15723D-04	1.5110
X27-XCAR	0.47349D-07	0.43376D-01
X41-6SFT	0.63724D-05	1.3866
CONSTANT	-0.72766	-1.7013

R-SQUARE= 0.5419
SSR= 0.2031D-01 DF= 50

NE70 7 -- 61 10 11 21 24 54 25 27 41 31 57 60

VAR	B	T
X61-DOCH	-0.55265D-03	-2.3310
X10-NONW	0.79829D-01	3.3134
X11-MAGE	0.30851D-02	2.7665
X21-DENS	0.71889	2.7189
X24-COLD	0.29075D-04	0.29997
X54-CI68	0.80286D-04	0.52619
X25-XPRO	0.12995D-04	1.2312
X27-XCAR	-0.20721D-06	-0.18181
X41-6SFT	0.12112D-05	0.22615
X31-NO69	-0.66609D-01	-0.74186
X57-SO70	0.26956D-03	1.5366
X60-PA70	0.97647D-04	0.93342
CONSTANT	-0.51502	-1.1677

R-SQUARE= 0.5743
SSR= 0.1887D-01 DF= 47

Table 4.20

Congenital Birth Defects

C\B%   8 -- 61 10 11 21 24 54 25 27 41

VAR	B	T
X61-DOCH	-0.11009D-02	-0.95740
X10-NONW	-0.10484	-0.91545
X11-MAGE	0.90212D-02	1.6360
X21-DENS	0.51667	0.38505
X24-COLD	0.41358D-03	0.86881
X54-CI68	0.11415D-02	1.5569
X25-XPRO	0.32274D-04	0.60844
X27-XCAR	-0.15016D-06	-0.26987D-01
X41-6SFT	0.15405D-04	0.65758
CONSTANT	-1.3390	-0.61416

R-SQUARE= 0.1867
SSR= 0.5277 DF= 50

C\B%   8 -- 61 10 11 21 24 54 25 27 41 31 57 60

VAR	B	T
X61-DOCH	-0.10806D-02	-0.88051
X10-NONW	-0.62417D-01	-0.50047
X11-MAGE	0.96362D-02	1.6693
X21-DENS	0.68174	0.49810
X24-COLD	0.29908D-03	0.59609
X54-CI68	0.15099D-02	1.9116
X25-XPRO	0.36136D-04	0.66141
X27-XCAR	-0.25284D-05	-0.42856
X41-6SFT	0.16833D-04	0.60717
X31-NO69	-0.47663	-1.0255
X57-SO70	-0.44966D-03	-0.49518
X60-PA70	0.46820D-03	0.86461
CONSTANT	-1.3072	-0.57256

R-SQUARE= 0.2205
SSR= 0.5058 DF= 47

Table 4.21

Early Infant Diseases

I\B% 9 -- 61 10 11 21 24 54 25 27 41

VAR	B	T
X61-DOCH	0.39159D-03	0.13339
X10-NONW	0.89283	3.0537
X11-MAGE	-0.44748D-02	-0.31786
X21-DENS	5.8110	1.6964
X24-COLD	0.35758D-03	0.29424
X54-CI68	-0.13312D-02	-0.71117
X25-XPRO	-0.10427D-03	-0.76995
X27-XCAR	0.17641D-04	1.2419
X41-6SFT	0.15367D-03	2.5695
CONSTANT	0.68864	0.12373

R-SQUARE= 0.4741
SSR= 3.439 DF= 50

I\B% 9 -- 61 10 11 21 24 54 25 27 41 31 57 60

VAR	B	T
X61-DOCH	0.21825D-02	0.73588
X10-NONW	0.77575	2.5739
X11-MAGE	-0.79126D-02	-0.56722
X21-DENS	5.4464	1.6467
X24-COLD	-0.47371D-03	-0.39069
X54-CI68	-0.12146D-02	-0.63634
X25-XPRO	-0.13747D-03	-1.0412
X27-XCAR	0.98932D-05	0.69390
X41-6SFT	0.67486D-04	1.0073
X31-NO69	-2.0458	-1.8215
X57-SO70	0.43841D-02	1.9978
X60-PA70	0.17731D-02	1.3549
CONSTANT	4.1264	0.74792

R-SQUARE= 0.5484
SSR= 2.954 DF= 47

Table 4.22

Cancer

CA70 53 -- 61 10 11 21 24 54 25 27 41

VAR	B	T
X61-DOCH	-0.71271D-02	-3.8356
X10-NONW	0.63382	3.4249
X11-MAGE	0.13235	14.854
X21-DENS	3.9772	1.8342
X24-COLD	0.17859D-02	2.3216
X54-CI68	0.50032D-02	4.2228
X25-XPRO	0.21038D-03	2.4545
X27-XCAR	-0.13166D-04	-1.4643
X41-6SFT	0.51683D-04	1.3652
CONSTANT	-9.5112	-2.6997

R-SQUARE= 0.8556
SSR= 1.378 DF= 50

CA70 53 -- 61 10 11 21 24 54 25 27 41 31 57 60

VAR	B	T
X61-DOCH	-0.70763D-02	-3.4976
X10-NONW	0.64712	3.1475
X11-MAGE	0.13242	13.915
X21-DENS	4.0247	1.7837
X24-COLD	0.17298D-02	2.0913
X54-CI68	0.51394D-02	3.9471
X25-XPRO	0.21071D-03	2.3394
X27-XCAR	-0.14074D-04	-1.4471
X41-6SFT	0.51447D-04	1.1257
X31-NO69	-0.16472	-0.21498
X57-SO70	-0.16143D-03	-0.10784
X60-PA70	0.23494D-03	0.26317
CONSTANT	-9.4472	-2.5101

R-SQUARE= 0.8560
SSR= 1.374 DF= 47

However, differences between our estimated air pollution effects as opposed to the Lave and Seskin (1977) work are profound. Lave and Seskin (1977) did not find a significant association between particulates and pneumonia. More importantly, Lave and Seskin (1977) found positive associations between air quality (specifically sulfate) and a cardiovascular disease mortality variable and between air quality and cancer mortality. Whether we use SO_2 or the highly collinear sulfate measure, we cannot accept the hypotheses that air pollution has any association with heart and vascular disease or with cancer mortality. Further, our estimated total effects of air pollution on human mortality are about one order of magnitude smaller than those shown by Lave and Seskin (1977).

We can summarize the results of our analysis as follows. When we <u>increase</u> each of the. following significant variables by one percent over their mean values in our sample, from the estimated total mortality equation the following percentage change in mean total mortality rate results: (1) for doctors per capita a 0.76 percent <u>decline</u> in mortality rate; (2) for per capita cigarette consumption a 0.32 percent <u>increase</u> in mortality rate; and (3) for per capita protein consumption a 6.7 percent <u>increase</u> in mortality rate. These results suggest several observations. First, medical care, smoking, and diet appear to be enormously important factors in human health. Second, if one looks to a 100% decrease from mean levels for these variables, i.e., the impact on average total mortality of setting these variables to zero, one obtains a 76% increase in mortality for a zero level of doctors per capita, a 3.2% decrease in mortality for no smoking and a 670% decrease in mortality for no protein in diet. Obviously, the last of these effects is impossible and suggests that we may only have linear approximations of highly non-linear effects. Further, some protein is required to sustain life. Thus, the estimates of mortality effects are likely to be valid only for relatively small changes in explanatory variables. Finally, the air pollution variables are insignificant in the total mortality equation -- as one might suspect if air pollution has only a small effect. on mortality rates. This is verified by the fact that the significant estimated effects of particulates on pneumonia and influenze, and of SO_2 on infant diseases are very small in terms of total mortality as compared to the effects of doctors, smoking, and diet.

Given these results, it is important to test the sensitivity of the model to changes in specification of included variables and structure. Two alternative formulations have been specified and tested. First, a version of the model which: (1) uses lagged diet (1955 dietary variable) as opposed to 1965 diet); (2) employs a two-stage doctors per capita variable which includes air pollution in the reduced form equation; and (3) adds lead and sulfate to the air pollution variables, produces essentially identical results both for the impact of medical care and, air pollution on mortality. Sulfate air pollution is statistically insignificant across all diseases. The second alternative formulation is identical to the one presented in detail above but the air pollution variables are again included in the reduced form equation for doctors per capita. The results are consistent for the effect of medical care and for the positive associations between sulfur oxides and infant diseases and for particulates and pneumonia. More interesting, however, is a significant <u>negative</u> association which appears between doctors per capita and air pollution in the reduced form equation. It appears that doctors may choose <u>not</u> to live in polluted cities (perhaps for aesthetic reasons). If this is the case, one can easily explain false positive associations between air pollution and mortality where medical care is excluded as an explanatory variable. If doctors avoid polluted cities, and if doctors do reduce mortality rates, then pollution could well be associated with higher mortality rates; but <u>not</u> because of any direct health effect of air pollution on mortality. Rather, failure to account for the locational decisions of doctors (supply and demand for medical care) may well bias estimated epidemiological relationships. In fact, the negative association between doctors per-capita and pollution is so strong, that when pollution is included in the reduced form equation for doctors, the estimated doctors variable used in the two-stage procedure becomes collinear with the pollution variables. This collinearity in some cases produced negative coefficients on the pollution variables in estimated dose-response relationships for some disease categories where pollution is used in the reduced form equation for docotrs per capita. Thus, it is important that, in spite of this collinearity, stable <u>positive</u> associations are retained between pneumonia and influenza and particulates and between infant diseases and sulfur oxides. The inclusion or exclusion of air quality from the reduced form equation has little impact on the conclusions of this study. In part, this occurs because air pollution is collinear with diet. In fact, saturated fats and sulfur oxides are reasonable proxy variables for each other. It has been shown by McCarthy (1971) that the exogenous variables which are collinear with included exogenous variables may be excluded from estimated reduced forms with little loss in consistency in a two-stage least squares procedure.

Another important question for analysis is the possibility that heteroskedasticity is present. At this point, we have only examined one disease category -- cancer mortality -- for this problem. An examination of the residuals plotted against several important explanatory variables (age, for example) showed no evidence of heteroskedasticity.

Finally, in interpreting the results, it should be observed that the associations we have found between mortality and air pollution are principally for diesases of the very young and very old -- particularly susceptible groups within the population. Further, these effects are those which one would perhaps associate with short-term as opposed to long-term air pollution exposures. It may well be that aggregate epidemiology may be incapable of revealing the long-term consequences of air pollution exposures. Two problems are particularly significant here. First, lagged data or data on air pollution histories is not available for such studies. Second, it is nearly impossible to control for population mobility in such studies. Thus, even if one accepts the hypothesis that air pollution levels show enough persistence over time to reveal long-term effects, population mobility will still distort and confound attempts at estimating such effects. A partial remedy for these problems is, of course, to use data on individuals as opposed to aggregate data. The next chapter provides a preliminary exploration of just such a data set.

We now turn to an economic evaluation of the value of air pollution control in reducing mortality based on the value of safety approach described

Table 4.23

Methodology for Health Benefits Assessment

Benefits = (Population at Risk) x (Value of Safety) x
(Reduction in Health Risk)

<u>Value of Safety Based on Consumer's Willingness to Pay</u>

Low estimate:	$340,000
Source:	Thaler & Rosen (1975)
High Estimate:	$1,000,000
Source:	Robert Smith (1974)

above.

4.6 A Tentative Estimate of The Value of Safety from Air Pollution Control

Given all of the caveats discussed above concerning the validity of the estimated effects of air pollution on mortality, it is possible to construct benefit measures using the methodology outlined in Section 4.2 above. The methodology is briefly summarized in Table 4.23.

First, to obtain national estimates, we must know the population at risk. Since our sixty-city sample is entirely urban, and since air pollution is principally an urban problem we will use a population risk for 1970 of 150 million urban dwellers. As a range for the value of safety, we will employ Thaler and Rosen's (1975) estimate of $340,000 (in 1978 dollars) as a lower bound and Smith's (1974) estimate of $1,000,000 (in 1978 dollars) as an upper bound. Finally, to provide an estimate of reduced risk from air pollution control, we will assume an average 60% reduction in ambient urban concentrations both for SO_2 and particulates. Then, using the mean concentration of these pollutants in our sixty-city sample as a basis for calculation, we can derive the average reduction in risk of pneumonia mortality for a 60% reduction in particulates and the average reduction in risk of infant diseases for a 60% reduction in SO_2 from our estimated dose response functions for these diseases.

Multiplying the population at risk by the assumed value of safety, and then by the <u>average</u> reduction in risk, gives a crude approximation of the benefits for a 60% reduction in national urban ambient concentrations of particulates and SO_2, respectively. National urban totals and the value of the average individual risk reduction are shown in Table 4.24.

The value estimates as shown in Table 4.24 agree surprisingly well with those developed by Lave and Seskin (1977) for national air pollution damages. However, the dollar value is similar only because we use a range for the value of safety (derived from observed market behavior of consumers) which is about an order of magnitude larger than the “value of life" figure based on lost earnings which is employed by Lave and Seskin (1977). We, of course, reject the value of life notion, instead focusing on the measurable concept of value of safety. Since there is no evidence to suggest that society puts less value on safety for children, the aged or women than on employed heads of households, we feel that the best measures available now for the value of safety should be employed for all individuals. Eventually, more refined measures of the value that different individuals place on safety may become available. However, for the time being, these are the best valuations of the social worth of safety we can employ.

Table 4.24

Urban Benefits from Reduced Mortality: Value of Safety for 60% Air Pollution Control

Disease	Pollutant	Average Individual Safety Benefit (1978 Dollars/Year)	National Urban Benefits (1978 Billion Dollars/Year)
Pneumonia	Particulates	29 - 92	4.4 - 13.7
Early Infant Disease	SO_2	5 - 14	.7 - 2.2
Total		34 - 106	5.1 - 15.9

CHARTER V

THE MICHIGAN SURVEY EXPERIMENT

5.1 Objectives of the Experiment

The data set employed in this chapter refers to the health status and the time and budget allocations of each of several thousand household heads over a nine-year period. Its highly disaggregated form therefore avoids many of the estimation problems associated with the aggregate data used in Chapter IV. This avoidance is not our only purpose, however. The richness of detail in the data set allows us to extend the range of phenomena that we study. Most important, we are able to investigate the morbidity effects of air pollution, considering acute effects and chronic effects separately. The detail of the data set allows us to identify much more readily those variables that are not current determinants of health status, thus providing a means of avoiding the simultaneity problems that plagued the aggregate dose-response functions of the previous chapter. It is important to note that the results reported here reflect a preliminary attempt to evaluate the usefulness of Michigan Survey Data in estimating morbidity (sickness) effects of air pollution and consequent economic losses. As a result of the preliminary nature of the research, many highly desirable transformations of the variables as defined in the Michigan Survey Data set have not yet been made. However, in spite of the preliminary nature of the results they do represent the first attempt to qualify the economic losses due to morbidity as opposed to mortality resulting from air pollution.

With the richness of the data available to us, we need not terminate our efforts after having estimated a set of dose-response expressions for the morbidity effects of air pollution. We are able to ascertain the labor productivity effects and the impact on willingness to pay to avoid chronic and/or acute illness as well. Both of these additional efforts are undertaken in this chapter.

5.2 The Sample and the Variables

Our analysis is based on yearly interviews conducted by the University of Michigan's Survey Research Center with a nationwide random sample of 4,802 to 5,862 families from 1968 through 1976. No families with living members were ever intentionally deleted from the sample, and, as families broke apart, the adult components were added to the sample as distinct families. The cumulative interview response rate over the nine-year period declined from 76 percent in the 1968 and first interview year to 55 percent in the 1976 interview year, implying an average yearly reinterview response rate of nearly 95 percent. From 1970 through 1976, this yearly response rate averaged 97 percent. Of special interest to us is that, in addition to substantial detail on household head time and budget allocations, the sample contains generalized measures of the head's health states as well as information on lifestyle and biological and social endowment variables that might plausibly contribute to the health states.

Information from the interview has been combined with data on a limited set of environmental variables, particularly information on air pollution concentrations, to establish imperfect measures of the environment in which each family head has lived during the nine-year period. To the best of our knowledge, the Survey Research Center data set is the only one currently available that combines, for the same set of individuals over a substantial number of years, information on places of residence, states-of-health, and time and budget allocations. The sample thus raises the prospect of our being able to value, through empirical applications of the economic theory of consumer behavior, the contributions of environmental pollution exposures to states-of-health.

The major characteristics of our sample and the variables we employ in our empirical efforts are presented in Tables 5.1 and 5.2. All variables refer to household heads. Table 5.1 gives complete definitions of variables, their scalings, and their assigned acronyms; Table 5.2 provides representative arithmetic means and standard deviations of variables used. Because we employ various partitions of the sample throughout the chapter, we do not use the Survey Research Center sample weights. Cur samples are therefore not entirely representative of the national population.

In Table 5.2a, so as not to make worse the already considerable and cumbersome length of the listing, only the two health variables, LDSA and ACUT are listed as dependent variables. The geometric means of the air pollution variables have their means and standard deviations entered for the various sample partitionings indicated at the bottom of the table. The means and standard deviations for the other variables are listed in Table 5.2b. This latter table refers only to the samples used for the chronic illness expressions, while the former refers to the acute illness expressions. Whether reference is to the partitioned or unpartitioned samples, the means and standard deviations represent only those samples used to estimate dose-response functions involving geometric mean measures of the air pollution variables. All estimates employing different combinations of variables, whatever the combination might be, were established using a random drawing from the entire Survey Research Center population sample for a particular year. Therefore, the means and standard deviations listed in Table 5.2, although extremely representative, are not the exact values for each of the samples used in the estimation effort.

The definition and measurement of most of the variables listed in Tables 5.1 and 5.2 is standard, and we shall comment here only on those that pose definitional and measurement problems for the major focus of this report. This criterion immediately directs attention to the air pollution variables.

Table 5.1

Complete Variable Definitions

Health State Variables

Acute illness (ACUT) -- workdays ill times 16 for the first 8 weeks and times 12 thereafter. Only individuals who are currently employed or unemployed and looking for work could have positive values for this variable.

Degree of disability (DSAB) -- complete limitation on work = 1; severe limitation on work = 2; some limitation on work = 3; otherwise = 0.

Length of disability (LDSA) -- ≤ 2 years = 1; 2 - 4 years = 2; 5 - 7 years = 3; ≥ 8 years = 4; otherwise = 0. This is a follow-up question to inquiries about whether the respondent has any physical or nervous condition that limits the amount or kind of work or housework he can do.

Biological and Social Endowment Variables

Age of family head in years (AGEH)

Grew up in city (CITY) = 1; otherwise = 0. This variable, as transformed, is binary.

Education attainment (EDUC) -- 6 - 8 grades = 2; 9 - 11 grades = 3; 12 grades = 4; 12 grades plus non-academic training = 5; college, no degree = 6; college degree = 7; advanced or professional degree = 8; otherwise = 1.

Father's educational attainment (FEDU) -- same scaling as for EDUC.

Family size in number of persons in housing unit (FMSZ).

Length of present employment (LOCC) -- ≤ 1 year = 1; 12 - 19 months = 2; 1-1/2 - 3-1/2 years = 3; 3-1/2 - 9-1/2 years = 4; 9-1/2 - 19-1/2 years = 5; ≥ 19-1/2 years = 6; otherwise = 0.

Marital status (MARR) -- married = 1; otherwise = 0. This variable, as transformed, is binary.

Income level of parents (POOR) -- poor = 1; otherwise = 0. This question asked whether the respondent's parents were " . . . poor when you were growing up, pretty well off, or what?" The variable, as transformed, is binary.

Race of family head (RACE) -- white = 1; otherwise = 0. This variable, as transformed, is binary.

Sex of family head (SEXH) -- male = 1; otherwise = 0. This variable, as transformed, is binary.

Member of a labor union (UION) -- Yes = 1; otherwise = 0. This variable, as transformed, is binary.

Life Style Variables

Practices absenteeism from work (ABSN) -- absent once or more a week from work = 1; otherwise = 0. This refers to a question in which the respondent is asked if there are times when he doesn't go to work at all, even if he isn't sick. The variable, as transformed, is binary.

Table 5.1
(continued)

Frequency of church attendance (CHCH) -- once a week or more = 1; otherwise = 0. This variable, as transformed, is binary.

Annual family expenditures on cigarettes in dollars (CIGE) -- this variable is not indexed for differences in prices among locales.

Participates in energetic activities (EXER) -- first mention = 1; otherwise = 0. This question asks the family head what he usually does in his spare time. Energetic activities include fishing, bowling, tennis, camping, travel, hunting, dancing, motorcycling, etc.

Family food consumption relative to food needs standard in percent (FOOD) -- family food consumption refers to food expenditures in dollars and includes amounts spent in the home, school, work, and restaurants, as well as the amount saved in dollars by eating at work or school, raising, canning or freezing food, using food stamps, and receiving free food. The food needs standard is in dollars and is based on USDA Low Cost Plan estimates of weekly food costs as published in the March 1967 issue of the Family Economics Review. The standard itself is calculated by multiplying the aforementioned weekly food needs by 52 and making a series of adjustments according to the size of the family.

Is often late to work (LTWK) -- late once or more a week to work = 1; otherwise = 0. This question asks the respondent if. there are times when he is late getting to work. The variable, as transformed, is binary.

Daily number of cigarettes smoked per adult family member (NCIG) -- ≤ 3 =1; 3- 17 = 2; 18 - 22 = 3; 23 - 35 = 4; 2 - 3 packs = 5; ≥ 4 packs = 6; otherwise = 0.

Fundamentalist religious preference (RELG) -- Mormon, United Church of Christ, Disciples of Christ, Quaker, etc. = 1; otherwise = 0. This variable, as transformed, is binary.

Degree of risk aversion (RISK) -- a weighted index devised by the survey team in which the individual's degree of risk aversion increases if he drives the newest car in good condition, does not own a car, has all cars insured, uses seat belts, has medical insurance, smokes less than a pack a day, has some liquid savings, and has more than two month's income saved. Nine is the greatest degree of risk aversion that can be exhibited.

Head's annual hours working for money (WORK).

Precuniary Variables

Cost-of-living in 1970 country of residence (BDALO) -- an index of comparative costs for a four-person family living in various areas as published by the U.S. Bureau of Labor Statistics in the Spring 1967 issue of Three Standards of Living for an Urban Family of Four Persons. The lowest living standard was employed. This index is published for the thirty-nine largest SMSA's and by region for the nonmetropolitan areas. For the remaining SMSA's, the regional average of the metropolitan indices was used.

Table 5.1
(continued)

Has hospital or medical insurance (INSR) -- Yes = 1; otherwise = 0. This variable, as transformed, is binary.

Family income in dollars not due to current work effort (ICTR) -- this variable includes assorted welfare payments, pensions, and annuities, as well as earnings from assets.

Family net real income in dollars (RINC) -- this variable is the sum of money income plus value of goods and services received at less than market prices less the cost of earning income.

Savings in dollars equal or greater than two month's income (SVGS) -- Yes = 1; otherwise = 0.

Head's marginal hourly earnings rate in cents (WAGE) -- in circumstances where the head neither has a second job nor commands overtime pay, this variable is simply total annual earnings from labor divided by annual hours worked for money. Where he has two or more jobs, it is his hourly earnings in the last job he names. If he has only one job, can and does work overtime if he wishes, and receives overtime pay, the variable is his average overtime hourly earnings.

Environmental Variables

Works in chemicals or metals manufacturing industries (CHEM) -- Yes = 1; otherwise = 0. The chemicals industry includes chemicals and allied products, petroleum and coal products, and rubber and miscellaneous plastic products. The metals industry includes steel, aluminum, foundaries, etc.

Number of days in 1972 when temperatures were below freezing at some time during the day (COLD). This data was obtained from USNOAA, Climatological Data, National Summary 1972.

Number of persons per room in family dwelling (DENS).

Distance from nearest city of 50,000 or more people (MILE) -- $\leq$ 5 miles or outside continental United States = 1; 5 - 15 miles = 2; 15 - 30 miles = 3; 30 - 50 miles = 4; $\geq$ 50 miles = 5.

Nitrogen dioxide: annual 24-hour geometric mean (M), ninetieth percentile (N), and 30th percentile (T) in micrograms per cubic meter as measured by the Gas Bubbler TGS Method-Frit before 1974 and the Saltzman method for 1974 and after (NOX). This data was obtained from the annual USEPA publication, Air Quality Data -- Annual Statistics.

Sulfur dioxide: annual 24-hour geometric mean (M), 90th percentile (N), and 30th percentile (T) in micrograms per cubic meter as measured by the Gas Bubbler Pararosaniline-Sulfanic Acid Method (SUL). This data was obtained from the annual USEPA publication, Air Quality Data - Annual Statistics.

Total suspended particulates: annual 24-hour geometric mean (M), 90th percentile (N), and 30th. percentile (T), in micrograms per cubic meter as measured by the Hi-Vol Gravimetric method (TSP). This data was obtained from the annual USEPA publication, Air Quality Data -- Annual Statistics.

Table 5.1
(continued)

Ultraviolet radiation in microwatts per square centimeter (ULTV). This data was taken from Pazand, R., Environmental Carcinogenesis -- An Economic Analysis of Risk, unpublished Ph.D. dissertation. University of New Mexico (June 1976).

Explanation of Table

Unless otherwise stated, all data is taken from tapes described in Survey Research Center, A Panel Study of Income Dynamics, Ann Arbor: Institute for Social Research, University of Michigan (1972, 1973, 1974 1975, 1976).

All variables referring to an individual person refer only to the family head.

On occasion, definitions for the same phenomenon will differ from year to year. If this occurs, a single integer indicating the year to which reference is made is attached to the end of the variable acronym. Thus 1967 = 7; 1968 = 8; . . .; 1976 = 6.

Table 5.2a

Representative Means and Standard Deviations of Health and Air Pollution Variables for Samples Involving Family Heads Currently Employed or Actively Looking for Work*

Variable Acronym	Year 1967	1968	1969	1970	1971	1972	1973	1974[b]
<u>Health States</u>								
ACUT	100.414 (183.594)	120.486 (214.759)	133.657 (332.171)	113.750 (277.022)	113.323 (266.274)	149.845 (427.983)	112.530 (259.120)	
LDSA[a]	0.953 (1.720)	0.645 (1.326)	0.337 (0.979)	0.363 (0.971)	0.268 (0.888)	0.290 (0.921)	0.260 (0.874)	0.348 (0.952)
<u>Environmental</u>								
NOXM						157.043 (51.070)	118.045 (72.230)	
SULM	24.475 (19.098)	25.113 (18.714)	27.220 (25.013)	16.286 (12.150)	17.657 (9.449)		2.051 (4.188)	7.435 (11.728)
TSPM	100.403 (35.469)	99.917 (30.628)	98.713 (29.609)	95.534 (18.943)	87.213 (27.920)	99.157 (30.941)	35.310 (42.183)	71.1.08 (36.085)

[a] Except for 1970, all samples refer to family heads who have never lived in more than than one state. In 1971, the reference is to family heads who currently live within walking distance of relatives.

[b] Includes housewives, retirees, and students.

*Standard deviations are enclosed in parentheses.

Table 5.2b

Representative Means and Standard Deviations of All Other Variables[a]

Variable Acronym	1967	1968	1969	1970	1971	1972	1973	1974	1975
Health State									
DSAB	0.493		0.111	0.426	0.488	0.470	0.304	0.800	0.624
	(1.291)		(0.315)	(1.634)	(1.011)	(0.949)	(0.754)	(2.159)	(1.854)
Biological and Social Endowment									
AGEH	43.558		40.323	43.745	44.218	44.305	45.155	37.322	37.925
	(12.337)		(11.841)	(13.451)	(13.649)	(15.276)	(16.158)	(15.421)	14.749
CITY	0.646		0.451	0.678	0.678	0.632	0.655	–	–
	(0.481)		(0.498)	(0.468)	(0.468)	(0.459)	(0.476)	–	–
EDUC	3.680		3.683	3.878	3.923	7.705	3.720	3.912	3.659
	(1.696)		(1.747)	(1.862)	(1.866)	(1.851)	(1.844)	(1.672)	(1.685)
FEDU	2.391		2.300	2.313	2.360	2.458	2.395	–	–
	(2.254)		(2.036)	(1.442)	(1.473)	(1.609)	(1.451)	–	–
FMSZ	3.812		4.586	3.993	3.930	3.508	3.233	–	–
	(2.401)		(2.542)	(2.376)	(2.412)	(2.202)	(2.126)	–	–
LOCC	2.257		3.271	–	–	2.283	2.168	–	–
	(2.234)		(1.869)	–	–	(2.168)	(2.188)	–	–
MARR	0.617		0.617	–	–	–	0.525	0.468	0.540
	(0.489)		(0.487)	–	–	–	(0.500)	(0.500)	(0.500)
POOR	0.578		0.543	–	0.520	0.490	0.520	0.551	0.615
	(0.496)		(0.499)	–	0.500	(0.501)	(0.500)	(0.499)	(0.488)

(continued)

Table 5.2b
(continued)

Variable Acronym	1967	1968	1969	1970	1971	1972	1973	1974	1975
RACE	0.469		0.917	0.410	0.500	0.443	0.475	–	0.346
	(0.501).		(0.276)	(0.099)	(.0.501)	(0.497)	(0.500)	–	(0.477)
SEXH	0.629		0.677	0.635	0.635	0.573	0.603	0.640	0.631
	(0.468)		(0.496)	(0.482)	(0.482)	(0.495)	(0.490)	(0.382)	(0.417)
UION	–		0.354	0.233	0.198	–	0.198	–	–
	–		(0.479)	(0.423)	(0.399)	–	(0.399)	–	–
Lifestyle									
ABSN	–		–	–	0.108	–	–	–	–
	–		–	–	(0.310)	–	–	–	–
CHCH	–		0.440	–	–	–	–	–	–
	–		(0.448)	–	–	–	–	–	–
CIGE	–		93.146	–	–	–	–	–	–
			(124.022)	–	–	–	–	–	–
EXER	0.144		0.189	0.225	0.198	–	–	–	–
	(0.352)		(0.392)	(0.418)	(0.399)	–	–	–	–
FOOD	505.830		757.669	822.500	840.990	–	–	1030.976	1145.150
	(380.977)		(372.594)	(716.450)	(716.100)	–	–	(574.163)	(707.099)
LTWK	–		–	0.070	0.209	–	–	–	–
	–		–	(0.255)	(0.407)	–	–	–	–

Table 5.2b
(continued)

Variable Acronym	1967	1968	1969	1970	1971	1972	1973	1974	1975
NCIG	1.851		–	–	–	–	–	–	–
	(1.912)		–	–	–	–	–	–	–
RELG	0.018		–	–	–	–	–	0.054	0.062
	(0.136)		–	–	–	–	–	(0.226)	(0.242)
RISK	4.489		4.503	4.658	4.673	–	–	–	–
	(1.605)		(1.452)	(1.545)	(1.540)	–	–	–	–
WORK	1245.875		1989.649	1560.895	1527.732	1333.540	1354.137	–	–
	(1059.780)		(674.723)	(1001.253)	(982.381)	(1030.346)	(1056.153)	–	–
<u>Pecuniary</u>									
BDALO	99.638		99.220	100.413	100.266	100.618	100.736	–	–
	(4.720)		(4.297)	(4,625)	(4.788)	(4.925)	(4.819)	–	–
INSR	0.889		0.794	0.708	0.695	–	–	–	–
	(0.316)		(0.404)	(0.455)	(0.461)	–	–	–	–
ICTR	1096.22		508.249	1238.392	1013.846	1342.585	1366.702		
	(1314.401)		(1124.259)	(1198.698)	(1721.377)	(1874.235)	(1993.720)	–	–
RINC	9148.605		8902.377	10852.230	10875.650	9556.803	11077.950	–	–
	6511.900		(6100.167)	(7833.473)	(7439.632)	(7274.871)	(8337.711)	–	–

(continued)

Table 5.2b
(continued)

Variable Acronym	1967	1968	1969	1970	1971	1972	1973	1974	1975
SVGS	0.342 (0.475)		0.289 (0.454)	0.333 (0.472)	0.371 (0.484)	– –	– –	– –	– –
WAGE	292.119 (405.985)		314.440 (221.346)	322.500 (316.450)	358.258 (331.738)	298.230 (319.890)	336.525 (337.425)	– –	– –
<u>Environmental</u>									
CHEM	0.022 (0.147)		– –	0.008 (0.086)	– –	– –	0.003 (0.050)	0.049 (0.216)	0.045 (0.206)
COLD	81.502 (52.684)		– –	– –	– –	– –	– –	– –	– –
DENS	– –		3.420 (1.797)	– –	– –	0.870 (1.198)	0.725 (0.414)	– –	– –
NOXN	– –		– –	– –	– –	246.573 79.826	104.860 (75.994)	97.429 (44.564)	90.717 (22.716)
NOXT	– –		– –	– –	– –	132.045 (37.087)	31.536 (23.964)	32.931 (31.761)	48.597 (13.911)
SULN	107.687 (134.484)		– –	74.663 66.016	61.768 (38.495)		42.625 (31.115)	34.566 (42.841)	25.650 (41.603)

(continued)

Table 5.2b
(continued)

Variable Acronym	1967	1968	1969	1970	1971	1972	1973	1974	1975
SULT	26.041			10.798	11.190	–	9.551	5.006	7.836
	(37.369)			(10.663)	(5.875)	–	(9.305)	(9.955)	8.233
TSPN	176.986			248.965	156.185	170.768	147.960	126.702	120.580
	(78.097)			(339.668)	(63.787)	(58.121)	(39.684)	(43.086)	(56.438)
TSPT	77.605			74.837	74.088	82.995	56.232	67.122	62.779
	(23.661)			(43.932)	(20.772)	(26.627)	(9.650)	(22.200)	(27.046)
ULTV	1494.75			–	–	–	–	–	–
	(634.638)			–	–	–	–	–	–

[a] All samples include housewives, retirees, and students.

*Standard deviations are in parentheses.

If one has detailed and real-time information on changes in health states, ideally one would like to have real-time records of all air pollution exposures. The coarse yearly indicators of acute and chronic illness in the Survey Research Center (SRC henceforth) data could not support such detail. We therefore chose to collect outdoor air pollution data averaged over a time period corresponding to the time interval employed in the SRC data. In addition, we wished to ascertain whether representations of moments of the outdoor air pollution frequency distribution other than measures of central tendency might contribute to ill-health. The result of these deliberations was a decision to acquire data on the geometric mean (because outdoor air pollution tends to be log-normally distributed over time), 30th percentile, and 90th percentile of the annual concentrations of five pollutants: nitrogen dioxide; ozone; total oxidants; total suspended particulates; and sulfur dioxide. Although the ozone and total oxidant data has been combined with the SRC data, the number of monitoring locations and the monitoring time intervals were inadequate to allow other than minor variations in the exposures of the sample individuals. Thus the empirical results to be reported neglect these two possibly important pollutants.

Matching the thousands of outdoor air pollution monitoring stations in the United States to the hundreds of counties where the SRC sample families resided could be a complex combinatorial problem. The matching was achieved for each of the nine years at the cost of not having outdoor air pollution information for some SRC sample families during some years and of assigning somewhat inappropriate air pollution exposures to some sample individuals. The full extent of this information loss is presently unknown.

The matching process started by listing all the counties in the United States where one or more SRC sample families had resided during the nine year interval. Separately for each of the five previously mentioned air pollutants, a yearly listing of the counties having outdoor air pollution data for one or more of the three frequency distribution measures being considered was constructed. Of the 301 counties in 50 states where sample families resided during the nine year interval, outdoor air pollution monitoring data for one or more of the measures of one or more of the five air pollutants existed at least for one year in 118 of the counties in 50 states. No attempts were made to extrapolate air pollution data from one county to another, nor were any switches between monitoring stations in a single county ever made. In counties where multiple outdoor monitoring stations were present, the data from the single station that had operated for the greatest portion of the nine years was used. If two or more stations in a county had operated for equal portions of the nine years, the station having the most complete (in terms of numbers of pollutants and pollutant measures) was employed. When air pollution data were available in a family's residence county for a particular year, these criteria served to assign outdoor air pollution exposures to all sample families. For most years, somewhat more than 3,000 families had some sort of outdoor air pollution data assigned them. Because of our reluctance to adopt a new monitoring station location in a county whenever the activities of a station we had previously used were terminated, we undoubtedly missed a few opportunities to assign air pollution data to a few sample families. This issue pales, however, beside the issue of the extent to which the assigned data represent actual outdoor air pollution exposures.

The SRC family data sample provides only the family's county and state of residence: it does not give the home town or city. Thus, for large urban counties such as Cook County, Illinois, or Los Angeles County, California, or occasional rural counties such as San Bernardino County, California, where there exist major locational differences in potential air pollution exposures within the county, substantial error could exist in the air pollution assignations. This important source of measurement error could perhaps be substantially reduced if all counties having this property were identified and if all families residing in the identified counties were excised from the sample. We have made no attempt to perform this excision.1/

This criteria employed to select pollution monitoring stations probably result in the assignment of downtown urban locations, where pollution concentrations have historically tended to be greatest and where the most extensive monitoring has been done. Since relatively few of the SRC sample families actually live in downtown areas, the constructed data set generally exaggerates family outdoor air pollution exposures, implying that the health effects, if any, of air pollution will tend to be under-estimated.2/

Outdoor air pollution at the place of residence is not the only plausible environmental source of deleterious health effects. Indoor air pollution at home and in the work place, outdoor air pollution at other locations, contaminants in diet, and water pollution are additional widely acknowledged possible sources. We introduce measures (albeit imperfect) of some of these plausible alternative sources in our empirical efforts and fail to give any attention to others such as water pollution. If these excluded types of pollution have health effects of their own, and if their extent tends to be positively correlated with the extent of outdoor air pollution, then the included air pollution variables will capture some of their contributions to ill-health, causing the measured contribution of the outdoor air pollution variables to be exaggerated. The extent of this upward bias will vary directly with the degree of correlation between the included and the excluded variables and the extent to which the excluded variable actually contributes to the effect of interest. For this study, of the previously mentioned alternative environmental pollution sources of health effects, the utter exclusion of any measures of water pollution is perhaps the most serious, At various points in the empirical effort, rather crude measures of indoor home air pollution (family smoking habits), diet (a dietary adequacy index), and indoor air pollution at the work place (employment in the chemicals or metals manufacturing sector) are included.3/

The issue of excluding possibly relevant variables from the analysis included outdoor air pollution as well. Oxidants and ozone, Because of insufficient variation in apparent exposures among sample families, have been disregarded, even though exposure values are present in the constructed data set. Other important air pollutants, for which data were available such as carbon monoxide, were not even considered because of the large variations in their instantaneous concentrations within a few hundreds of feet. Some pollutants that have attracted recent regulatory and public concern, such as acid sulfates, had no data readily available. Finally, of the pollutants that were included in the constructed data set and exploited in the empirical effort, the time series for all except total suspended particulates were incomplete. Thus, for example, no information was available on sulfur dioxide concentrations in 1972.

Measurement error is not only an issue in the outdoor air pollution variables. What some might choose to interpret as measurement error is a prime feature of the two dependent variables, number of days annually ill and length of time disabled.[4/] Although we have no basis other than seemingly sensible intuitive interpretations of the form of the questions asked the respondents (see the explanations for ACUT and LDSA in Table 5.1), we choose to interpret the former as acute illness and the latter as chronic illness. Definitional problems of the distinction between acute and chronic illness aside, it must be remembered that what is an illness to one individual is not an illness to another individual. Even the same individual may differ over time in what he considers to be a state of illness. Illness is, in part, an idiosyncratic and subjective phenomenon only partly susceptible to consensus standards of definition. Therefore, if one prefers a reductionist perspective and wishes to have all phenomena collapse to, say, a chemical measurement, then the values of the variables we are trying to explain in this study indeed leave a great deal to be desired.[5/] Economic analysis, however, presumes that illness and its costs lie in the eye of the beholder. No laws whatsoever governing choices are innate in the material objects of ordinary cognition. As has been emphasized in the introduction to this section of the report, the degree of illness that afflicts an individual is, in part, often a matter of purposive choice. Economic principles relate to the subjective desires motivating individuals to become aware of and perhaps to alter their environments. Thus no object or status becomes relevant in economic analysis until humans perceive it can be used for or defeats some subjective purpose. Illness that is defined in clinical terms but which is never subjectively realized by the individual who is said to be clinically ill is of little interest except to clinicians. It is certainly arguable whether their standards of what constitutes illness should prevail over those of the individual who professes illness. For this study, we are forced by circumstances to adopt the latter's perspective. Fortunately, it fits readily into economic analysis.

In spite of the preceding argument a type of measurement error does persist in the two dependent health variables. This type of error is inherent in the use of any fairly encompassing measure of health status. Kinds of debilitating acute illness for an individual may range, for example, from headaches to heel blisters. Chronic illnesses may show similar variations over body sites and implied debilitating effects for the same individuals. In effect, therefore, an individual's response to a question about the number of days he has been ill or the length of time he has been disabled involves an aggregation of several attributes perhaps sampled from some larger population of attributes. The weights the respondent employs to combine these attributes to obtain the encompassing health measures may differ among individuals. Furthermore, they may not be those weights that correspond to the contribution of the attribute to some other parameter of interest, such as hours of work or money wages. Recognition of the possibility that individuals may employ different weights to aggregate to the encompassing health measure serves perhaps to deepen the reader's perception of the subjectivity of our measures of ill-health. It says only that there may be as many unique measures of ill-health employed as there are respondents in the sample. The import for our empirical efforts of discrepancies between the contributions of attributes to ill-health and to other parameters of interest is greater, since we shall try to ascertain the impact of air pollution-induced ill-health upon labor supply and productivity. In particular, the use of the encompassing measures of ill-health rather than the specific attributes may attenuate cur estimate of the effect of air pollution-induced health effects upon labor supply and productivity.

As Table 5.1 indicates, all SRC sample individuals not currently employed or seriously looking for current employment had no information recorded about the number of days they professed to be acutely ill. Furthermore, those individuals for whom information on ACUT was recorded were never sick on weekends: their accute illneses occurred, according to the data, only on workdays. The ACUT variable may thus be confounded by the wish of some respondents to legitimatize for the sake of social appearance or internal self-respect their failure to go to work. In the empirical efforts regarding ACUT therefore, an actual choice of leisure over labor could thus be falsely attributed to ill health. Marquis (1978), however has been unable to discover any basis for this source of bias.

The rather long list of other variables considered can be divided, somewhat imperfectly, into health state, biological and social endowment, lifestyle, pecuniary, and environmental variables. For the moment, we will limit our discussion of the variables not already discussed to the parts they are expected to play in dose-response functions, reserving the discussion of labor supply and productivity impacts to a later section. Only those variables actually used in the estimated dose-response functions are therefore discussed in this section. A summary table of expected signs is presented in Table 5.3.

DSAB, the degree of disability is the only included health state variable not employed as a dependent variable. Since it is ordinally scaled, its meaning when used as a dependent variable is arbitary. Any four or five monotonically increasing numbers would have no more and no less meaning. When entered as an explanatory variable in the chronic illness production function, its expected sign is unclear. If the individual continues to live in spite of having a chronic disability, one would expect the period of recovery, if any, to be lengthier the more severe the disability. However, in the general population, severe disabilities perhaps are more likely to lead to earlier death. Thus, those sample individuals who are severely disabled might be expected to have been disabled only for a relatively short time span. This would lead one to expect a negative association between DSAB and LDSA. Which effect would dominate in any particular sample must be conjectural. In contrast, since disabilities, both in terms of length and severity, probably cause the

Table 5.3

Expected Signs for Explanatory Variables
in Estimated Dose-Response Functions

	Acute Illness	Chronic Illness
Health States		
DSAB	+	?
LDSA	+	X
Biological and Social Endowments		
AGEH	+	+
CITY	?	?
EDUC	?	?
FEDU	-	-
FMSZ	-	?
MARR	-	?
POOR	+	+
RACE	-	-
SEXH	?	?
Lifestyles		
CHCH	-	-
EXER	-	X
FOOD	-	-
NCIG	+	+
RELG	-	-
RISK	-	-
Precuniary		
INSR	-	-
Environmental		
CHEM	+	+
COLD	?	?
DENS	+	+
All NOX	+	+
All SUL	+	+
All TSP	+	+
ULTV	?	?

? ≡ unknown
X ≡ irrelevant

individual to be more susceptible to common temporary illness, we expect LDSA and DSAB to contribute positively to ACUT. However, because the values for DSAB are not monotonically ordered, the magnitudes of the coefficients for DSAB in both the LDSA or the ACUT expressions should be disregarded.

No one holds that health states improve with adult age. The adult human organism suffers natural decay, making the investment necessary to maintain a given health state progressively more costly. The inclusion of two additional irrevocable biological attributes, race and sex, can be justified on at least two grounds. First, susceptibilities to some diseases differ by race or sex. Men, for example, don't have breast surgery and whites don't contact sickle cell anemia. The implications of this for the signs of RACE and SEXH are unclear, however. Second, and probably most important with respect to race, minorities have frequently had less preventive and ameliorative medical care available to them and have perhaps had less opportunity to learn how to use what is available wisely. The RACE variable might therefore capture some fair portion of past and present differences in the availability of medical services to individuals. If this speculation is correct, RACE, which has a value of 1 if the individual is white and 0 otherwise, should have a negative sign attached for both illness types.

CITY, FEDU, and POOR are intended to represent differences among individuals in their childhood environments. If one grew up in a city, he probably had better access to medical care. On the other hand, he was probably exposed to more toxics in his everyday environment. The sign to be expected for CITY is therefore ambiguous. In contrast, the proper signs to expect for FEDU and POOR are relatively unambiguous. Educated parents, in addition to their other knowledge about worldly affairs, will perhaps be more sensitive to the implications of childhood health practices for future adult health status of the child. In addition, they might tend to be better at interpreting signals of health distress and choosing the medically most effective course of action. If adult health states are positively influenced by childhood health practices, then the sign attached to the FEDU coefficients in either acute or chronic illness dose-response functions should be negative. For similar reasons, the POOR coefficients are expected to have positive signs.

With one ambiguous exception, EDUC, FMSZ, and MARR contribute to good health. Many recent studies indicate that among socioeconomic variables, years of formal schooling completed is frequently the most important predictor of good health. Grossman (1975) has found empirical evidence of a causal relationship running from past schooling to current health. The individual who is married has his wife's time available, as well as his own, for the protection of his health. At least for acute illness, increasing family size also implies that certain individuals within the family can specialize in the production and the protection of other family members' health. This implies that over some interval there exist increasing returns to health production specialization within the family, a proposition that accords neatly with casual observation but for which no strong empirical evidence appears to exist.

The expected sign for FMSZ in a chronic illness dose-response function is ambiguous because the number of children a family has is, in part, an investment decision.[6/] Older children provide more opportunities for family members to specialize in health production and protection; however, if a state of chronic disability was suffered by the family head before the accumulation of a large family, it would seem that the investment process in children would be made more costly. The latter statement implies that fewer children and chronic disability are positively associated, while the former says that children, once they are able to assume some responsibilities for family production, contribute to good health. Put in terms of our concerns in the introduction to this portion of the report, an observed association between an individual's state-of-health and his family size could reflect causality running both from family size to health and from health to family size. This issue could, of course, be resolved by building an analytical structure in which family size is made a decision variable. To do so would take us beyond the immediate scope of this research effort. We have therefore employed family size as an explanatory variable in our estimated chronic illness dose-response functions without imposing any sign expectations upon it and recognizing that its presence could bias the air pollution coefficients.

All of the lifestyle explanatory variables are standard entries in epidemiological studies of air pollution. There are, however some special features worthy of note for each variable. NCIG, for example, is not the number of cigarettes smoked by the individuals but rather the number smoked per adult family member. It is assumed this serves as a reasonable proxy for the smoking habits of the individual head. For the cigarette variable therefore, its estimated coefficient is best considered as an indicator of the health effects of smoking <u>or</u> not smoking. Little, if any, credence should be assigned these coefficients as indicators, in the neighborhood of the average smoking habits of the respondent sample, of the incremental health effects of smoking an additional cigarette; that is, the sign of the coefficient rather than its magnitude is the result to inspect.

Biomedical wisdom says that continuing participation in energetic activities and an adequate diet contribute to good health. Since the SRC data set contained no information on the respondent's exercise habits before he became disabled, we have not included EXER in the chronic illness dose-response function. Otherwise one must face the two-way causality problem with inadequate data resources to handle it. In neglecting this variable, however, which proves to be consistently statistically significant in the acute illness dose-response function, we raise the spectre of biasing the air pollution coefficients in the estimated chronic illness dose-response functions. Since, <u>a priori</u>, energetic activities are expected to reduce the incidence of chronic illness, the absolute magnitudes of the air pollution coefficients will be biased downward, causing the effect of air pollution on chronic illness to be underestimated. However, for those years in which EXER is available in the SRC data set, the absolute value of the simple correlation between it and the air pollution variables is generally less than (0.15. The bias its exclusion introduces is probably therefore minor unless it intrinsically has a very strong influence on the <u>magnitude</u> of the chronic illness variable.

So as to enhance the creditability of the dietary habits variable, FOOD, we quote from Survey Research Center (1972a, p.75):

> "Since expenditure on food is a relatively easy to measure proxy for adequate nutrition and is one of the study's more important variables, much care has been taken to improve the technique of asking these questions; several refinements, but no added objectives, have resulted in a few changes to these questions over the five waves of the survey."

Accepting the assertion that the amount of food expenditures was one of the most carefully treated questions in the entire SRC survey effort, the issue remains as to whether these expenditures, even when stated relative to food "needs," are capable of providing useful information on the etiology of illness. Certainly they can provide no information on dietary contributions to particular diseases unless expenditures on particular food groups are known. But then we are dealing in any case only with generalized measures of self-reported health status. As for the use of expenditures on food rather than actual food consumption, one's comprehension of this measure is aided if it is viewed as a proxy for a stock variable relating to the history of the individual's investments in diet. Real capital in the hospital industry is not measured in terms of gadgets and buildings but rather as the discounted value of the accumulated investments. Similarly, dietary adequacy may be measured as the discounted value of the individual's accumulated expenditures on food. FOOD, which is simply current expenditures on food relative to a "needs" standard, will generally tend to be positively related to this discounted value.

The intent of including the CHCH, RISK and RELG variables is to capture acquired behavioral traits consistent with an out-of-the-ordinary aversion to health-endangering activities. We hope at least some of those forms of health-enhancing everyday behavior not otherwise available in the data set collapse into these variables. Among these forms would be regulatory getting six to eight hours sleep, a tightly-knit and emotionally supportive family life, a healthy mix of foods consumed, and the many other lifestyle factors to which assorted medical commentators variously attribute the production and protection of good health.

INSR, a dummy variable referring to whether or not the individual is covered by medical insurance, should be correlated with the individual's consumption of medical care. The variable should be negatively related to the price of medical care that the individual faces and therefore positively related to the quality of medical care he has consumed. If medical care improves health or maintains good health, then the medical insurance variable should have a negative coefficient in both the acute and the chronic illness dose-response functions. Our use of this variable in a dose-response function might be criticized on grounds that it is serving as a proxy for the quantity of medical care consumed, where this quantity and the proxy are the consequence of current period decisions. We admit the possible validity of this view but nevertheless chose to retain INSR as our only available proxy likely to be strongly associated with the individual's adult history of medical service consumption. In short, we assume that the benefits to estimation from including a plausibly relevant variable (a history of the individual's adult consumption of medical services) outweigh the losses to estimation incurred by employing a current period decision variable as an explanatory variable in a single equation structure.

Among the environmental variables, all the air pollution variables, as well as DENS and CHEM, are expected to have positive signs for both acute and chronic illnesses. People who live in crowded conditions are in closer contact with other individuals, making personal sanitation more difficult, and increasing the probabilities of contracting whatever communicable illnesses plague others. The contacts of workers in the chemicals and metals manufacturing sectors are not so much with carriers of communicable illnesses, but rather with exposures to toxic substances in the work place. These exposures are thought to exceed those of the rest of the population.

Hippocrates, 460-337 B.C. (1939) and the writers of a large literature descending from those ancient times have asserted a sort of climatic determinism with respect to health.[7/] We briefly acknowledge this literature by considering two climatic variables, COLD, to represent the extent of freezing weather, and ULTV, to indicate the amount of sunshine. Although the literature in this area says that climate has an influence on health, any advice it gives as to whether these climatic parameters are harmful or beneficial is unsettled. We therefore prefer not to make assertions about the signs to be expected for the coefficients of these variables.

A great many more variables for each of our variable classes is available on the SRC survey tapes. In addition, since the county of residence is known for each individual respondent for each year of the survey, additional environmental and general area information could be combined with the SRC tapes. Many more variables could be constructed from the basic SRC information. We did initially consider some other definitions and versions of the variables in Table 5.3, but this list should provide a reaonable description of the data we had available.

Before proceeding to the presentation and discussion of the dose-response functions, there are several salient characteristics of the constructed data set that do not necessarily have clear implications for the results but which nevertheless provide form and a setting for them. Tables 5.2 and 5.4 are thus worthy of some attention. The reader is reminded, however, that these tables are incomplete: they are only representative of the samples used to estimate the dose-response functions.

Note that three of the characteristics of Table 5.4 are consistent with a high proportion of the individuals in the sample having lived for long periods in one locale. People who live within walking distance of relatives, have always lived in one state, and have never moved to take a job elsewhere have likely had a long history of exposures to the outside air pollution of one municipality. In short, the SRC data allow one to compensate somewhat for the lack of a long data series on the pollution

Table 5.4

Proportions of Entire Survey Research Center Sample Processing a Particular Characteristic During 1971

Characteristic	Percent
Asset income $\leq$ \$500	81.1
Children $\leq$ 25 years in family unit	51.3
Has relatives living within walking distance	42.6
Employed head	72.7
Unemployed head	2.2
Retired head	16.6
Housewife head	6.7
Student head	1.6
Working wife	33.3
Disabled person in family other than head	3.8
Neighborhood of detached single-family homes or lesser density	65.9
Rents dwelling unit	37.8
Always lived in one state (1970 data)*	40.4
Never moved from a community for a job change (1970 data)*	57.9
Disabled head	21.8

*These proportions are not indicated in the code book describing the 1971 data. It is highly unlikely that they differ significantly from the 1971 proportion.

exposures of sample individuals. If one is willing to assume that relative pollution concentrations among locations have been reasonably constant over time, then he can at least loosely grasp the effects of cumulative exposures on differences in health states. These cumulative exposures might not be terribly relevant with respect to acute illness, but they can be highly important with respect to chronic illness, Therefore in all our empirical efforts dealing with chronic illness, we deal only with sample individuals who have always lived in one state or who have never moved for a job change. Even though this partitioning by no means guarantees that we fully capture the cumulative air pollution exposures of the sample individuals, we believe that it does so to a substantially greater degree than do most air pollution epidemiology data sets.

The proportion of sample individuals who profess disabilities consistently approximates one out of every five. Over the nine year interval of the data set, it ranges from a low 18.2 percent in 1974 to a high of 23.6 percent in 1969. In fact, only for the 1974 and 1975 entire SRC population samples was the proportion disabled below 20 percent (in 1975, the proportion was 18.4 percent). These lower proportions for 1974 and 1975 are probably due to the rather drastic drop in the mean age of the sample population occurring between 1973 and 1974, which is reflected in the mean values for the AGEH variable in Table 5.2. The drop causes the proportion of the SRC sample that reports being disabled to better approximate the proportion disabled in similar area probability samples of the U.S. civilian non-institutionalized population. These other samples generally tend to have ten to fifteen percent of their individuals suffering from self-reported disabilities.

A glance at Table 5.2 shows that the number of individuals employed in the chemicals and metals manufacturing sector is usually too small, given sample sizes of about 400, to estimate reliably the extent to which the exposures associated with this employment generate illness. As earlier noted, the 1973 SRC data include information on three-digit occupational codes by three-digit industry for the sample individuals. If, after having carefully perused the data to ascertain exactly which occupations in which industries involve substantial exposures to toxics, the entire SRC population sample were to be used to estimate an acute or chronic illness dose-response function, one might have sufficient degrees of freedom available to obtain reliable coefficients for these manufacturing sectors. At best, one or two of the samples we employ here have enough sample individuals employed in these sectors to be slightly suggestive about an association between exposures in them and acute or chronic illnesses.

Finally, when evaluating the empirical results reported in this study, one must face the question of the accuracy of respondent recall. Since there exists no data base referring to contemporaneously observed sample individual behavior and status, one's judgments about accuracy must necessarily be more-or-less personal and introspective. The following pair of facts can aid in the formation of this judgment. First, all respondent interviews were conducted within 12 months of the year for which respondent behavior and status was to be reported. Thus the longest interval that could pass between some respondent event and his reporting of that event was 23 months. In all years, however, the great bulk of the interviewing was completed by June of the year following the year that was to be reported. For these respondents, the greatest time lag that occurred between an event and its reporting was 17 months. The smallest lag that could occur, since interviewing started in early March of the year following the year to be reported, was two months. 8/

Perhaps more relevant to the recall issue than the question of lags is the incentive respondents had to make mental or written note of their behavior and status to ensure accurate answers when the appointed time for their interviews arrived. Several points relevant to this incentive issue can be made. First, as reinterviewing "waves" (this is the SRC's term) passed, those original respondents who were hostile to the interviewing process and purpose probably removed themselves from the sample. We speculate that those who voluntarily stayed in the sample possessed a substantial incentive for accurate recall. This implies that data from later years is perhaps more reliable than data from earlier years. Second, those families that did remain in the sample became more familiar with what would be asked them with each reinterviewing wave and would therefore take more care to make mental or written note of events so that they could be accurately reported. This too implies that data from later years tends to be more reliable. Third, the respondents were paid a small sum ($5.00 - $7.00) for participating in the interview. Finally, after having completed an interview, the respondent was left a postcard that he was asked to send to the SRC in early January of the following year. This card informed the SRC of the respondent's current address. Those who did and did not return the cards were sent a reminder and a postcard in January, along with a summary explanation of empirical results from the interviewing of the preceding year. All who returned the postcards, whether or not reminded, were rewarded with an additional payment of $5.0O. The SRC does not report the proportion of those who returned postcards, but, given the reinterview rate, one can reasonably conclude, that the return rate must have been fairly high. We judge from this that respondent interest in the survey must have been substantial, resulting in an incentive to keep rather careful track of behavior and status.

Aside from the detail of its information, the SRC sample and its combination with the air pollution data contain little that is remarkable relative to other data sets that have been used in air pollution epidemiology. Judging from the general sociodemographic attributes depicted in Tables 5.2 and 5.4, the sample in spite of our disregard of the SRC sample weights, appears to be close to a random sample of the U.S. civilian non-institutionalized population. The high proportion of non-whites does, however, raise some doubt about its exact representativeness. 9/ The increasingly better control of sulfur dioxide emissions is clearly registered in Table 5.2, although control of particulates and nitrogen dioxide appears not to have exhibited much improvement over the nine-year interval. Table 5.2, by its failure to show data for variables in some years that appear in other years, exhibits both changes in the SRC interview formats as well as our deletion of variables in expressions estimated for some years when they were not statistically significant in expressions estimated for samples drawn from other years.