CHAPTER V

THE OZONE EXPERIMENT

A. INTRODUCTION

The ozone experiment developed in this chapter was undertaken to satisfy a variety of objectives.

First, benefits of reducing ambient ozone concentrations are poorly understood apart from the overall value of reducing photochemical air pollution. Thus, development of a methodology for using the contingent valuation technique for valuing reductions in ozone exposure to households was one objective.

Second, the contingent valuation approach has been applied using mail surveys in some instances and interview surveying in other instances. However, the comparability of the two approaches has never been established. We accomplish that objective by employing both mail and interview surveying in valuing ozone reductions in six sample communities in the Los Angeles area. Overall, although response rates are substantially lower for the mail surveys, the two approaches give very similar results. This is quite surprising since we deliberately did no follow ups to increase the response rate for the mail surveys because we were interested in detecting non-response bias. This possible lack of apparent bias has a number of important implications. For example, the Bishop and Heberlein study (1979) used mail surveys, but included actual dollar payments for obtaining some bids. This study is important because it includes actual, as well as hypothetical attempts to repurchase hunting permits. However, the applicability of the results of this study have been limited because mail surveys might have differed substantially in bidding outcomes from interview surveys. Also, if mail surveys are valid, surveying for benefits of national environmental programs could be undertaken at a greatly reduced cost compared to in person interviews. Our results as presented in Section C suggest that further research in this area is warranted. We originally expected to reject mail surveying for bidding games as complex as the one used in this study.

The third objective was to obtain a better understanding of environmental preferences and how those preferences might affect the location decisions of individuals. As we show in Sections B and C, respectively, the theoretical and empirical linkages between survey responses and hedonic property values have not been explored, yet, this is a rich area for future research.

The fourth objective was to explore the consistency of daily bids for air quality levels with annual bids for a positively desired change in the frequency distribution of occurrence of those air quality levels. If annual bids (as perhaps capitalized in the property value study discussed below) are consistent with daily bids, as we show in Section E, then people are plausibly perceiving both the impact of daily changes in air quality on annual air quality, and of daily bids on annual bids, correctly. Also, this consistency, as shown in Section B, implies that individuals' utility functions are roughly separable over time in air quality.

Finally, the fifth objective was to attempt to validate the contingent valuation approach for ozone by comparison with a property value study, which we present in Section D. The property value study has been plagued by problems of multicollinearity. Distance to beach and the air quality variables of interest, ozone proxying for sub-clinical health effects and TSP (or extinction coefficient) proxying for aesthetic-visibility effects, are all highly collinear in the Los Angeles area. A variety of techniques were employed to attempt to solve this problem. The technique which <u>appears</u> to give the most stable results is the principal components approach. The precise economic-statistical implications of this approach are not well understood, so our results should be interpreted with caution. However, the objective of obtaining a health vs. aesthetics valuation split using a hedonic property value study is extremely important both for policy, since existing regulations are primarily health based, and to allow a comparison with the survey approach for valuing ozone. This comparison, which is quite favorable, is made in Section E.

B. THEORETICAL ISSUES IN INTERPRETING DAILY BIDS FOR AIR POLLUTION CONTROL

Two issues are of concern in analyzing individual daily bids for ozone reduction.

First, individuals will likely have very different tastes with respect to air pollution control. In a previous study (see Brookshire, Schulze, et al., 1982; and Schulze, et al., 1983) where individuals were allowed to bid for differing levels of pollution abatement for the Grand Canyon, some individuals had concave bid functions for reductions in air pollution willingness to pay increased at a decreasing rate for better air quality) while others had convex bid functions (willingness to pay increased at an increasing rate for better air quality). The latter.case is usually considered to be "pathological" in that nonconvex indifference curves are implied for individuals with convex bid functions. However, this case is not entirely implausible for environmental commodities. If individuals value a pristine environment very highly, but feel that a somewhat polluted environment is just as bad as a very polluted environment, then they will bid little for improvements in air quality to levels below pristine, but bid relatively large amounts to achieve pristine air quality. We analyze this type of behavior below, focusing on developing a simple measure of tastes to reflect the convexity of bid and indifference curves for analyzing the frequency of occurrence of individuals with what we will term "nonconvex environmental preferences" after the shape of the implied indifference curve. In addition, we show that with a well defined hedonic property value market for air quality, individuals with nonconvex preferences should cluster in the least and most polluted areas available and not be found in moderately polluted areas, Later, we examine this prediction in terms of the occurrence of nonconvex preferences as estimated from our surveys conducted in a highly polluted versus a moderately polluted area of Los Angeles County. We also compare the frequency of occurrence of nonconvex preferences as obtained from mail versus interview surveys to test for relative bias in sampling between the two approaches.

The second issue is the validity of obtaining daily bids for air quality improvements. Daily bids greatly simplify survey design, clarity and specificity, but imply a degree of separability over time which may not be entirely realistic. For example, an individual may wish to have clean air mostly on the day of a planned tennis game and care less if other days during the week are polluted. The validity of employing uniform daily bids for air quality improvements is evaluated below with a theoretical model specifying the degree of separability of utility functions which would be necessary to justify this approach.

To explore these issues, the following notation will be used:

Let
t = time in days ($t=1, 2, 3, \ldots$);
P_t = level of air pollution on day t;
R_t = reduction in pollution on day t;
Y = consumer income;
y_t = consumption on day t;
and B_t = daily bid for air pollution reduction.

To evaluate nonconvex preferences, time will initially be deleted from the analysis. Thus, consumer utility is taken to be a function of income and pollution.

$$U(y,P) \tag{5.1}$$

where the partial derivative U_y is positive and U_P is negative. If the initial pollution level is P°, the observed pollution level is given by

$$P = P^\circ - R \tag{5.2}$$

where R is the reduction in air pollution associated with the policy or standard to be valued. The bid, or willingness to pay for pollution reduction, denoted B, can then be defined using a compensating variation-measure by the following equation

$$U(y^\circ,P^\circ) = U(y^\circ - B,\ P^\circ - R). \tag{5.3}$$

The initial income and pollution levels y° and P° respectively give utility on the left-hand-side of (5.3) which is set equal to the utility on the right-hand-side determined by the new income level (which is reduced by the bid for pollution control to y°-B) and the new pollution level (which is lowered by the reduction in pollution to P°-R). Thus the maximum willingness to pay for pollution control is B.

Marginal willingness to pay can be obtained by totally differentiating (5.3) and solving for $\partial B/\partial R$ which yields:

$$\partial B/\partial R = -U_p/U_y > 0. \tag{5.4}$$

This expression is strictly positive given our assumptions on the signs of U_p and U_y. To obtain the curvature of the bid function implied by (5.4) with respect to pollution reduction, R, we take $\partial(\partial B/\partial R)/\partial R$ to obtain

$$\frac{\partial^2 B}{\partial R^2} = \frac{U_{pp}}{U_y} - \frac{U_p}{(U_y)^2} U_{yp}. \tag{5.5}$$

The usual assumption would be that the bid curve would increase at a decreasing rate in R so the expression in (5.5) would be negative. This would be true if $U_{pp} < 0$ and $U_{yp} < 0$ (or $U_{yp} > 0$ and sufficiently small) given that $U > 0$ and $U_p < 0$. Under these assumptions the indifference curve between y and R has the usual shape for positively desired commodities as shown in Figure 5.1 and the bid curve appears as shown in Figure 5.2. However, as indicated above, there is some evidence that bid curves for some individuals may increase at an increasing rate. This will occur if $U_{pp} > 0$ and $U_{yp} > 0$ for $U_{yp} < 0$ and $|U_{yp}|$ sufficiently small). Figures 5.3 and 5.4 show the indifference and bid curves respectively for the case of nonconvex preferences. Note also that the arrow in Figure 5.3 denotes the direction of preference, i.e., that y and R are desired commodities.

To test for nonconvex preferences among our respondents, we estimate individual bid curves as a function of pollution reduction using the following functional form

$$B = kR^{\eta} \tag{5.6}$$

where k and n are estimated as separate parameters for each respondent. Given this functional form, $\partial^2 R/\partial R^2$ takes the form

$$\partial^2 B/\partial R^2 = k\eta(\eta-1)R^{\eta-2} \begin{array}{l} < 0 \text{ if } \eta < 1 \\ = 0 \text{ if } \eta = 0 \\ > 0 \text{ if } \eta > 1. \end{array}$$

Thus, if the estimated parameter n is larger than unity for an individual respondent, we have an indication that the individual has nonconvex preferences as defined above. Further, we can treat n as a taste parameter reflecting the shape of respondents' indifference curves and plot frequency distributions of n among subsamples to see how tastes are distributed between our mail versus door-to-door surveys and how tastes are distributed spatially as well.

This last point deserves further elaboration. Our previous research suggests that a well defined property value gradient for air pollution exists in the Los Angeles area. This implies that the cost of a home or

Figure 5.1: Convex Indifference Curve

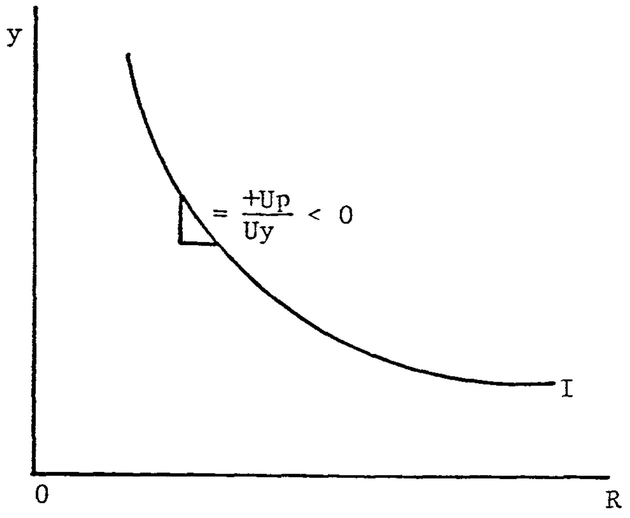


Figure 5.2: Concave Bid Function

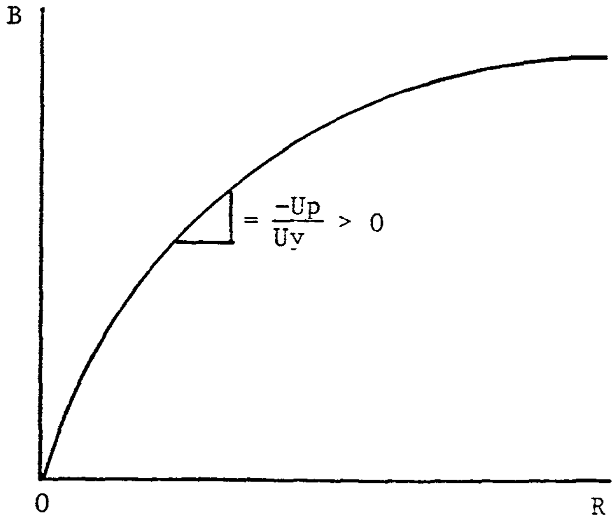


Figure 5.3: Concave Indifference Curve

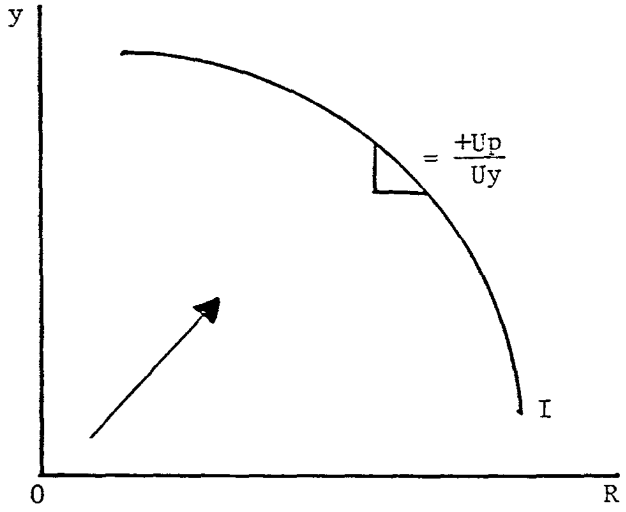
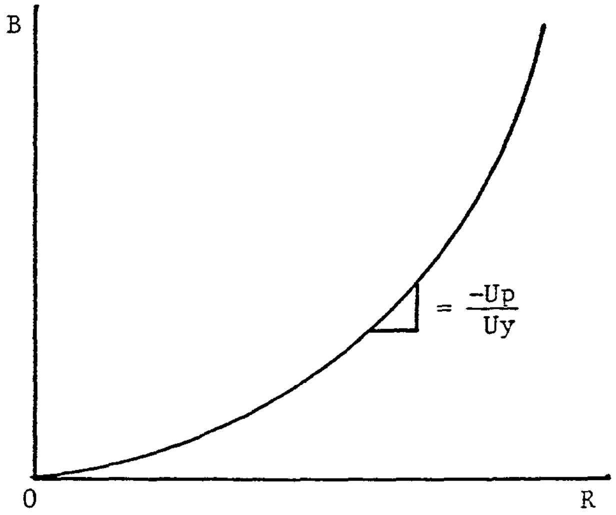


Figure 5.4: Convex Benefit Function



apartment varies with air pollution level. Where we denote this cost as C(P) where $C'(P) < 0$, consumers will choose a pollution "location" where they maximize utility,

$$U(y^\circ - C(P), P), \tag{5.8}$$

over choice of P. The first order condition for maximization of (5.8) implies

$$\frac{U_p}{U_y} = C'(P) \tag{5.9}$$

or that the slope of the indifference curve as shown in Figure 5.1 should lie tangent to the rent gradient which has a slope of $C'(P)$. The solution to this problem is shown graphically in Figure 5.5 for the case of normal preferences where $P = P^\circ - R$ is substituted into (9) above yielding

$$\frac{U_p}{U_y} = C'(P^\circ - R). \tag{5.10}$$

In Figure 5.5, R = 0 represents the worst air quality available in the region, where the air pollution reduction is zero. The vertical line at R_{max} denotes the best air quality available in the region, where the air pollution is reduced to the maximum extent possible. The cost of housing, $C(P^\circ - R)$, is subtracted from the horizontal line $y^\circ - y^\circ$, representing initial income before housing cost is subtracted, yielding the net income curve, $y^\circ - C(P^\circ - R)$. The indifference curve denoted I is tangent to the net income curve where pollution reduction is R* and the individual chooses to live at a pollution level $P = P^\circ - R^*$. The individual has chosen to reduce pollution by living in a less polluted area, but to pay a higher cost for housing than would have obtained in the most polluted area. Individuals with convex preferences would presumably have solutions like that shown in Figure 5.5 with tangencies distributed between R = 0 and R_{max} However, individuals with nonconvex preferences will likely locate only at R = 0 or at R_{max} as shown in Figure 5.6. Thus, for example, an individual with a preference direction A (and associated nonconvex indifference curves) would have a corner solution and locate at point a, an area of maximum pollution. An individual with preference direction B would also have a corner solution but locate at point b, an area of least pollution.

Thus, we have as a theoretical prediction that individuals with nonconvex preferences for air quality should cluster in the most and least polluted areas and that such individuals should be poorly represented in moderately polluted area. We test this prediction by examining the relative frequency of occurrence of nonconvex preferences (as indicated by n's greater than unity) in heavily versus moderately polluted areas in and around Los Angeles. Our empirical results presented in a following section show remarkable consistency with this prediction.

Figure 5.5

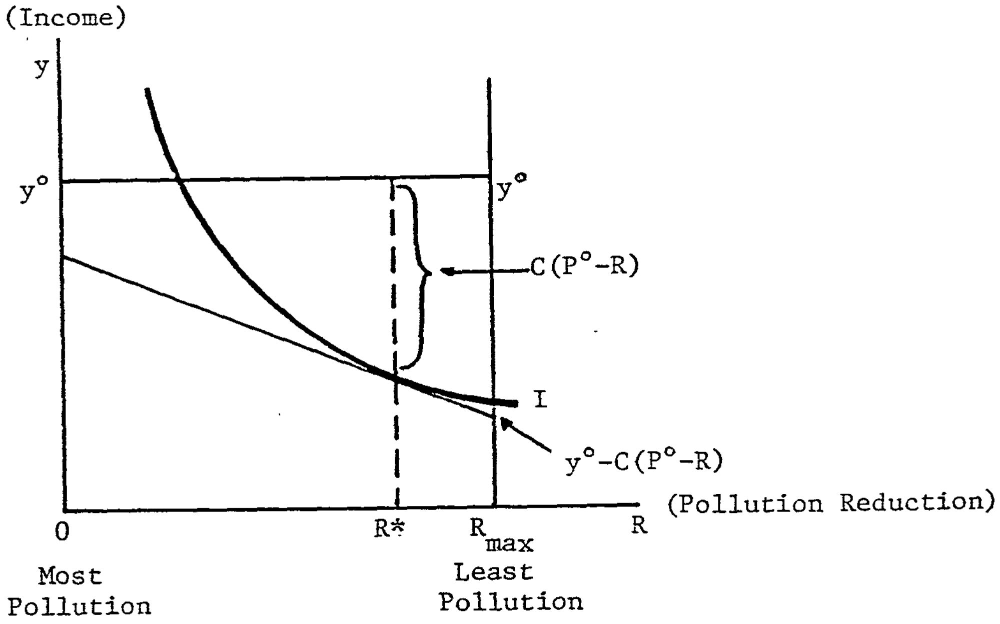
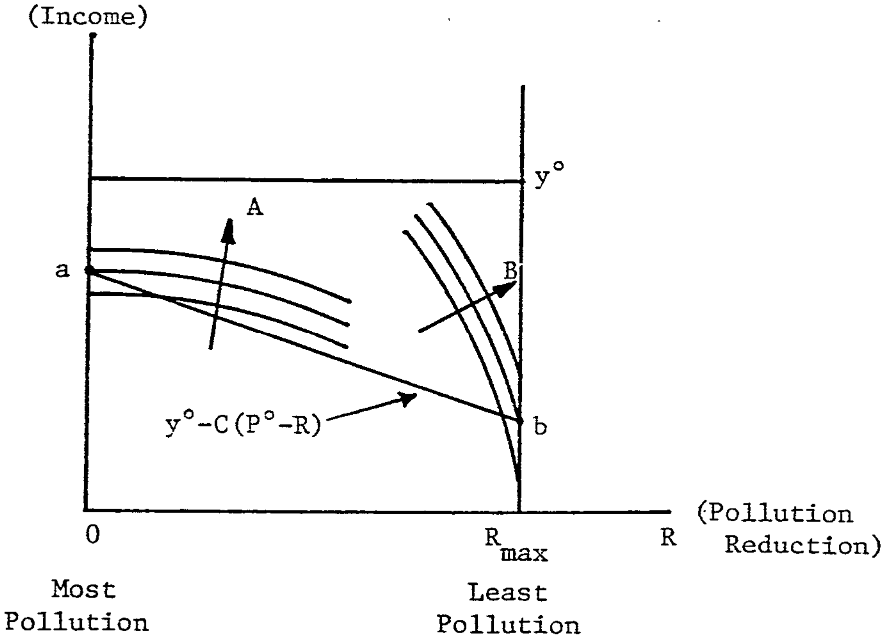

(Income)
y
y^o
y^o
$C(P^o-R)$
I
$y^o-C(P^o-R)$
(Pollution Reduction)
0
R*
R_{max}
R
Most Pollution
Least Pollution

Figure 5.6


(Income)
y°
A
a
B
y°-C(P°-R)
b
0
R_{max}
R
(Pollution Reduction)
Most Pollution
Least Pollution

The second theoretical consideration is that of uniform daily bids. In general, utility over time can be specified as

$$U(y_1, y_2, \ldots, y_T;\ P_1, P_2, \ldots, P_T) \tag{5.11}$$

a function of expenditures on day t, y_t, and pollution on day t, P_t, for all days over the planning horizon from t = 1 to t = T. If individuals could hypothetically purchase a reduction in air pollution on day t equal to Rt by paying a cost $c_t(R_t)$ then the budget constraint would be

$$y - \sum_{t=1}^{T} y_t - \Sigma\, c_t(R_t) \tag{5.12}$$

where we ignore the role of compound interest or assume the planning horizon is very short. Substituting $P_t = P^\circ_t - R_t$ into (5.11) where P°_t is the initial pollution level before reductions R_t are purchased, the consumer optimization problem is to choose y_t and R_t to maximize (5.11) subject to (5.12). Where λ is the Lagrange multiplier on (5.12) and L denotes the relevant Lagrangian, first order conditions are:

$$\partial L/\partial y_t = U_{y_t} - \lambda \leqq 0 \tag{5.13}$$

and

$$\partial L/\partial R_t = - U_{P_t} - \lambda c_t{}' \leqq 0. \tag{5.14}$$

Combining these we obtain (for noncorner solutions)

$$\frac{-U_{P_t}}{\lambda} = c_t{}'. \tag{5.15}$$

The left-hand-side of (5.15) is effectively identical to the marginal bid $B/\partial R$ defined earlier as $\partial B/\partial R = - U_P/U_y$ in (5.4) above. In both versions, the numerator is the marginal disutility of pollution while the denominator is the marginal utility of money (λ here is the shadow price on the budget constraint (5.12)). However, in this case $\partial B/\partial R$ is a fairly complicated expression since

$$\frac{\partial B}{\partial R} = \frac{-U_{P_t}(y_1, \ldots, y_T;\ P_1, \ldots, P_T)}{\lambda} \tag{5.16}$$

and as can readily be seen, the marginal disutility of pollution depends on expenditure levels over time, the date t, and on pollution levels over time. In terms of daily bids, λ is, most likely, practically fixed. However, daily marginal bids may well depend on whether the particular day is one on which high expenditures are planned, a long weekend occurs, or neighboring days are polluted or clear. This level of complexity would make surveying for bids difficult if not infeasible.

Thus, the approach taken has been to ask for an average daily bid. Another justification would be to assume that the utility function is separable as follows:

$$U = u(y_1, \ldots, y_T) - \sum_{t=1}^{T} D(P_t) \qquad (5.17)$$

so utility derived from daily expenditures, $u(y_1, \ldots, y_T)$, is separable from the disutility derived on any day from pollution, $D(P_t)$. Further, disutility from pollution on day t, $D(P_t)$ is separable from disutility on any other day t', $D(P_{t'})$, but the disutility function D(P) is the same for every day. In this case, marginal daily bids are of the form

$$\partial B/\partial R = \frac{D'(p_t)}{\lambda} \qquad (5.18)$$

where $P_t = P^\circ_t - R_t$. Except for some minor interdependence through effects on the marginal utility of money, λ, this implies separability of daily bid functions for air pollution control. This simplicity is of great use in survey design and also eases the task of calculating total benefits of changing the frequency distribution of occurrence of air pollution levels, which is the actual effect of air pollution control programs. However, as we have tried to point out above, the assumptions to allow this simplification are extreme indeed.

C. THE CONTINGENT VALUATION APPROACH

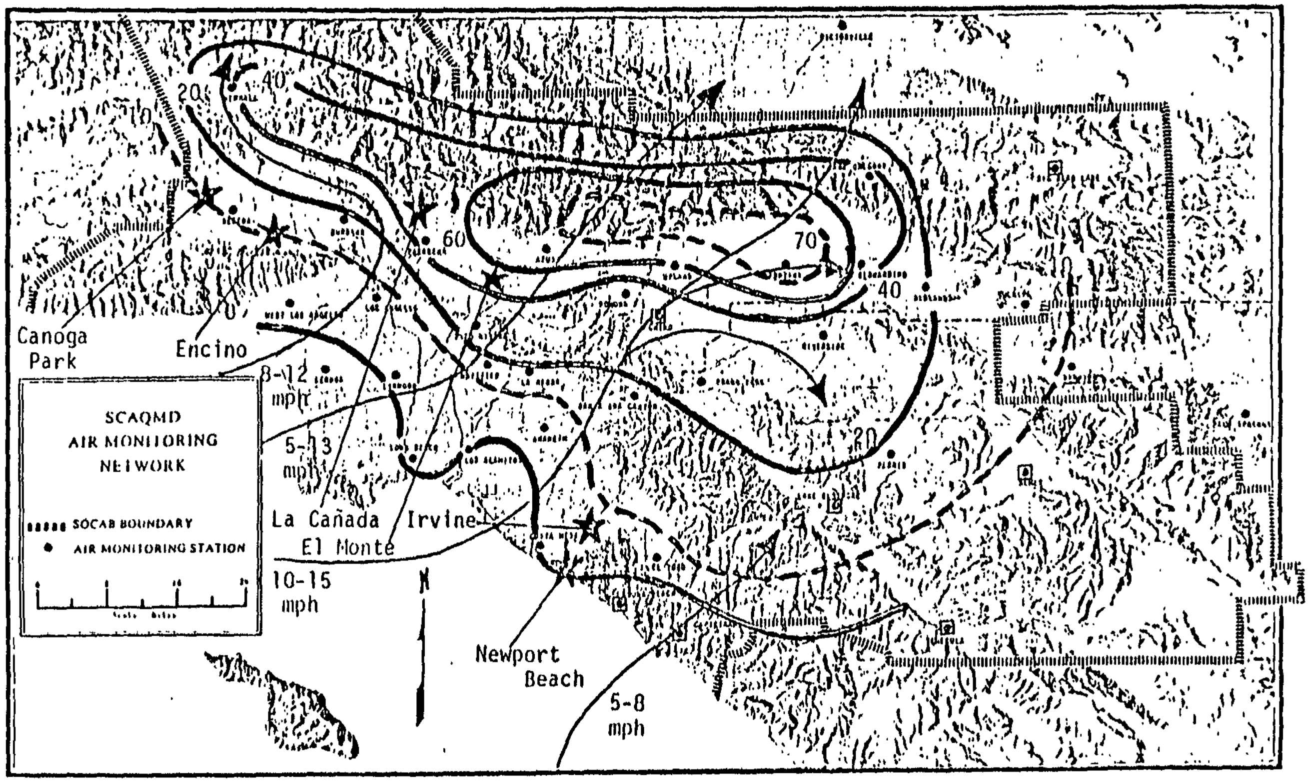
C.1 The Sample Plan

To provide a broad range of values for potentially relevant variables, six survey areas were selected that varied in peak ozone concentrations as well as in demographic characteristics.

The survey areas are in: La Canada and El Monte (in the West San Gabriel Valley); Canoga Park and Encino (in the San Fernando Valley); and Irvine and Newport Beach (in North Coastal Orange County).
Figure 5.7 shows the location of the survey areas in the South Coast Air Basin (SOCAB). The illustration also shows the number of Stage I Ozone Episodes during 1981 in the SOCAB.

It can be seen that La Canada and El Monte had approximately 50 such episodes during 1981, the San Fernando Valley communities had about 10 such days and in Orange County, Irvine had 5 and Newport Beach 0 Stage One Episodes. There is year-to-year variation in air quality measures apart from long-run trends but these figures provide a rough measure to indicate the diversity of ozone levels in the survey areas. Also shown on Figure 5.7 are typical daytime wind patterns. These winds are largely responsible for the intra-basin movement of airborne emissions.

Figure 5.7: Sample Areas, Number of State One Ozone Episodes in 1981 and Daytime Summer Wind Patterns in the South Coast Air Basin



◘ Air monitoring discontinued at this site.

→ Typical Summer Daytime Ocean Winds

★ Communities Surveyed

----- In 10-day intervals.

—— In 20-day Intervals.

* Source: "Season and Diurnal Variation in Air Quality in California's South Coast Air Basin" "1981 Summary of Air Quality in the South Coast Air Basin of California" Both published by South Coast Air Quality Management District

Various demographic traits of the survey areas are presented in Table 5.1.

When reviewing these traits, it should be kept in mind that no attempt was made to select a random sample of SOCAB residents. Rather, the intent was to provide sample communities which would provide the wide range of values sought in air quality and demographic measures.

This sampling technique is appropriate since the experiment was not an attempt to estimate aggregate benefits of ozone reduction across the SOCAB.

As can be seen, there is considerable variation among the sample areas in most characteristics. Mean household income (in 1979) ranged from $14,213 to $65,738. Further, within each air quality area there was variation in 1979 mean income: $14,213 and $65,738 in San Gabriel Valley; $16,028 and $58,675 in San Fernando Valley and $32,096 and $43,528 in Orange County. The desirability of low ozone levels made it virtually impossible to identify a neighborhood with high air quality and low incomes.

There was similar variation in other demographic variables: average number of persons per household varied within each air quality area although the variation was less in the San Fernando Valley.

The San Fernando Valley survey areas also showed relatively little variation in the fraction of the population that was more than 64 years old. In both these cases in which the San Fernando Valley showed relatively little variation, though, the values were intermediate. That is, there was no indication that the communities selected for any air quality area were extreme (except for the areas selected for extreme high or low ozone levels).

Within these broadly varied communities it might be possible to discern meaningful patterns in response rates or values of responses. The results are discussed in sub-sections C.3 and C.4 of this chapter.

C.2 Survey Design

Design Considerations

Survey-based bidding to estimate the value of nonmarket goods has been shown (Brookshire, et al., 1982) to be capable of producing estimates consistent with alternative evaluation techniques. Reliability in such estimates requires, however, that the object of the bid be a well-defined and understandable good and that the payment vehicle be plausible.

These are not trivial requirements in the case of basin-wide reduction of ozone concentrations.

Ozone is known to be among the most lethal of gases (National Research Council of the Rational Academy of Sciences, 1977) Even at the very low

TABLE 5.1

U.S. CENSUS INFORMATION FOR SAMPLE AREAS*

Community	Census Tract No.	Population	Avg. Persons	No. of Households	Mean Income	% > 64 Years	% White	Mean Travel Time to Work
La Cañada	4607	4903	3.03	1616	65,738	1.1	96.2	21 min.
El Monte	4334	9175	3.43	2673	14,213	7.1	72.7	21 min.
Canoga Park	1345	5645	2.40	2352	16,028	8.7	72.9	20 min.
Encino	1396	4319	2.60	1681	58,675	9.3	94.4	30 min.
Irvine	525.04	4340	3.16	1375	32,096	2.3	82.2	23 min.
Newport Beach	630.01	7528	2.25	3347	43,528	11.4	97.0	19 min.

*Source: 1980 Census.

concentrations (0 - 50 parts per hundred million) seen in SOCAB ozone has been shown to have significant effects on human health and comfort.

Ozone, however, exists as one of many irritants in photochemical smog. The effects of ozone in combination with these other pollutants is poorly understood. Even the effects of pure ozone have been difficult to examine: ethical and logistical difficulties inhibit the study of long-term intermittent exposure on human subjects while effects on experimental animals vary considerably among species.

The easily-identified effects of ozone exposure appear to be reversible, but are not always easily explained. In addition, some of the most common effects of smog (such as eye irritation) are typically caused by components other than ozone.

Ozone is produced when certain emissions (ozone precursors) are exposed to sunlight. In SOCAB daytime on-shore breezes move these compounds inland during the exposure period, resulting in higher ozone concentrations further inland (see Figure 5.7) with peak concentrations during late morning and afternoon (Hoggan et al.) Because of more intense solar radiation ozone, concentration tends to be higher in summer than winter.

The distribution of ozone concentrations within SOCAB varies with daily wind patterns, other meteorological phenomena and the level of human activity which produces ozone precursors. The issue of ozone reduction then is the issue of a probabilistic reduction of exposure to an agent with probabilistic effects.

Early consideration was given to the use of a downward shift in the annual distribution of daily maximum ozone concentrations as a bid object. While such a shift has the advantage of being the likely result of any feasible ozone reduction policy, it could not be presented in a manner suitable for a mail survey to the general population.

A specified ozone reduction on a specific day is more easily comprehensible but gives the choice of the day special significance. People might reasonably have very different preferences among weekends, holidays and other days and might even feel strongly about different weekends during any summer month.

A bid object was finally selected which was intended to be fully enough specified to elicit comparable responses from a wide range of individuals, but which avoided arbitrary specification of detail.

Identification of the good to be bid upon was accomplished by referring to a memorable day and using ozone levels on that day to define the base level for bids to reduce ozone concentration on an unspecified summer day.

Selection of the "memorable" day was straightforward: the summer of 1982 was one of generally low ozone levels, with a sharp increase just before and during the Labor Day weekend (see Figures 5.8-5.10). This last major holiday of the summer was also the time of a major outdoor concert (the US Festival). The coincidence of a severe deterioration in air quality and an entertainment spectacle caused widespread news coverage of both.

No such fortuitous event presented itself to aid in the designation of a payment vehicle.

A fee placed on the emission of ozone precursors would involve at least moderately intrusive monitoring of private vehicle use. A payment vehicle with substantial inconvenience would cause respondents' desire to avoid the inconvenience to mask their willingness to pay for ozone reduction.

The most workable payment mechanism seemed to be a generalized price increase with special attention drawn to increased operating costs for vehicles.

The specification of a good to be bid upon and the designation of a payment mechanism constitute the core of the experiment. The bid questions were supplemented with a number of other questions designed to provide information about the respondent.

The Survey Instrument

Separate (but similar) instruments were designed for each air quality area surveyed (San Gabriel Valley, San Fernando Valley, North Coastal Orange County).

Mail and Interview surveys differed only in that the Interview instrument included mechanical instructions to the interviewer to ensure a uniform survey procedure. The survey instruments are included in Appendix A.

Each survey instrument begins with a prologue which identifies the research team but not the sponsor. This is followed by a review of ozone effects and recent conditions in the survey area. After focusing attention on Labor Day weekend, 1982 (see sub-section C.2) the respondent is asked whether he (or she) or any family member experienced any of the described effects of ozone exposure. For each survey area the reference day is different because the ozone peak occurred on different days in different parts of the SOCAB. The questions for the San Gabriel Valley are:

1. Did you or any of the members of your immediate family experience any of the "ozone-induced" effects described above on Thursday, September 2?

_____ Yes _____ No (Please Check)

2. If you answered yes, which of these symptoms did you notice?

Figure 5.8

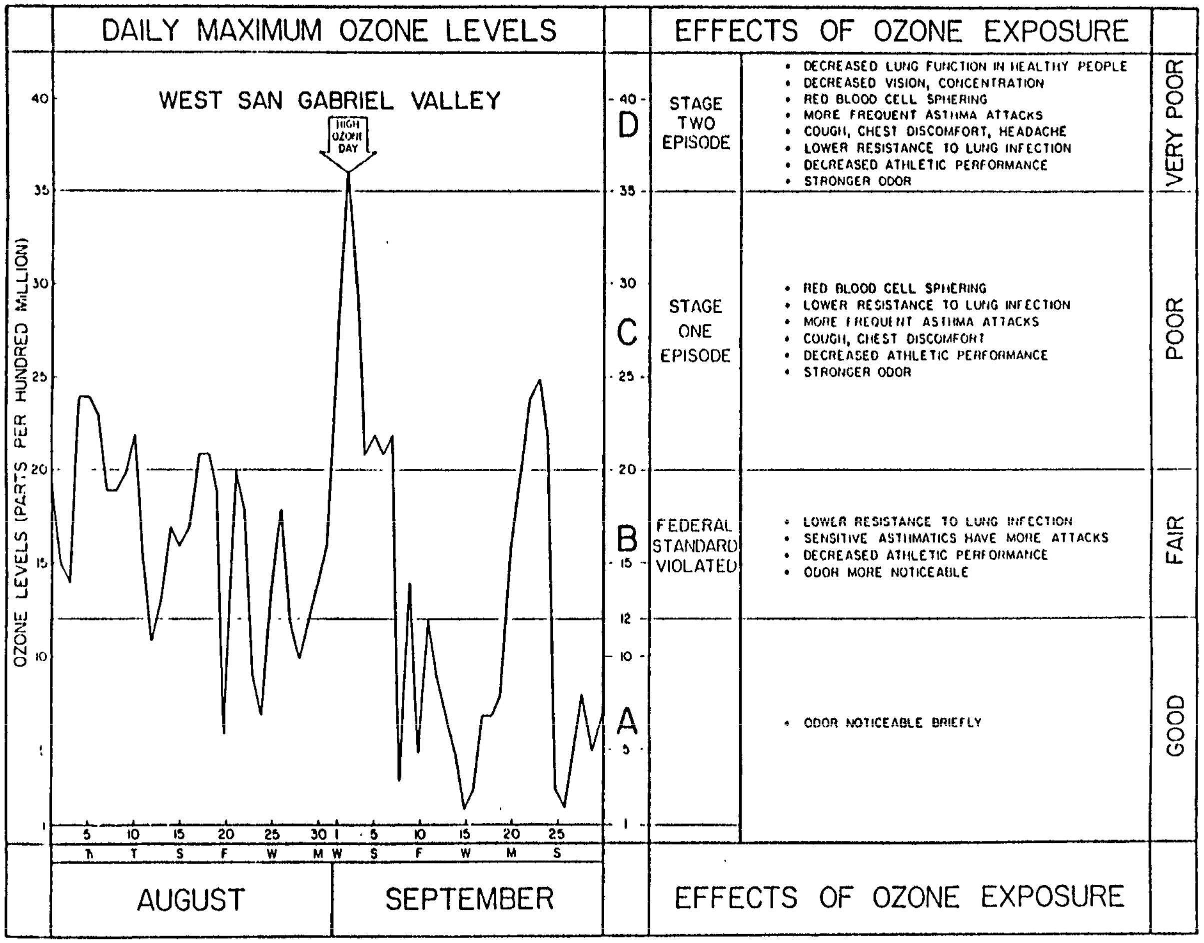

DAILY MAXIMUM OZONE LEVELS
WEST SAN GABRIEL VALLEY
HIGH OZONE DAY
OZONE LEVELS (PARTS PER HUNDRED MILLION)
40
35
30
25
20
15
12
10
5
1
5
10
15
20
25
30
1
5
10
15
20
25
Th
T
S
F
W
M
W
S
F
W
M
S
AUGUST
SEPTEMBER
D
C
B
A
EFFECTS OF OZONE EXPOSURE
STAGE TWO EPISODE
• DECREASED LUNG FUNCTION IN HEALTHY PEOPLE
• DECREASED VISION, CONCENTRATION
• RED BLOOD CELL SPHERING
• MORE FREQUENT ASTHMA ATTACKS
• COUGH, CHEST DISCOMFORT, HEADACHE
• LOWER RESISTANCE TO LUNG INFECTION
• DECREASED ATHLETIC PERFORMANCE
• STRONGER ODOR
VERY POOR
STAGE ONE EPISODE
• RED BLOOD CELL SPHERING
• LOWER RESISTANCE TO LUNG INFECTION
• MORE FREQUENT ASTHMA ATTACKS
• COUGH, CHEST DISCOMFORT
• DECREASED ATHLETIC PERFORMANCE
• STRONGER ODOR
POOR
FEDERAL STANDARD VIOLATED
• LOWER RESISTANCE TO LUNG INFECTION
• SENSITIVE ASTHMATICS HAVE MORE ATTACKS
• DECREASED ATHLETIC PERFORMANCE
• ODOR MORE NOTICEABLE
FAIR
• ODOR NOTICEABLE BRIEFLY
GOOD
EFFECTS OF OZONE EXPOSURE

Figure 5.9

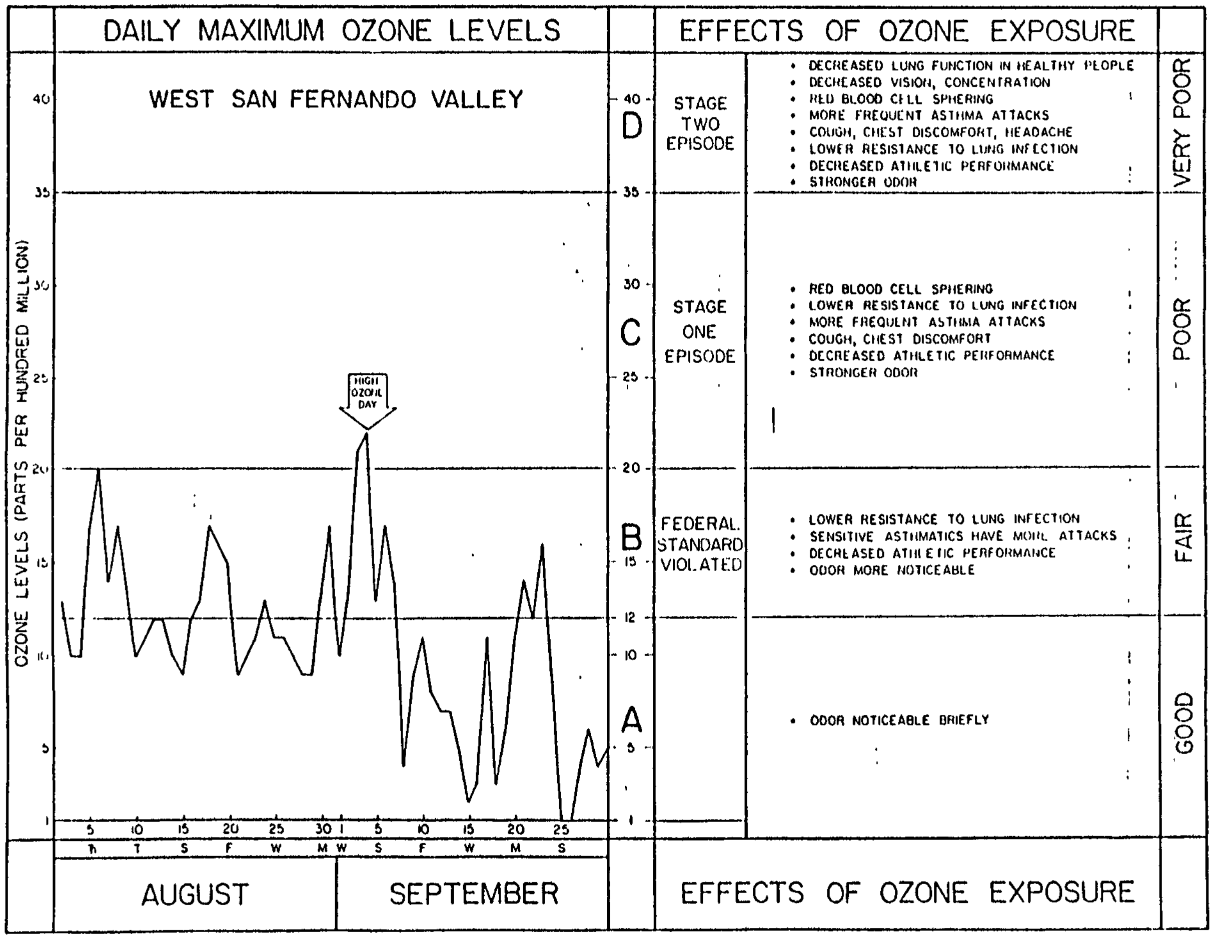
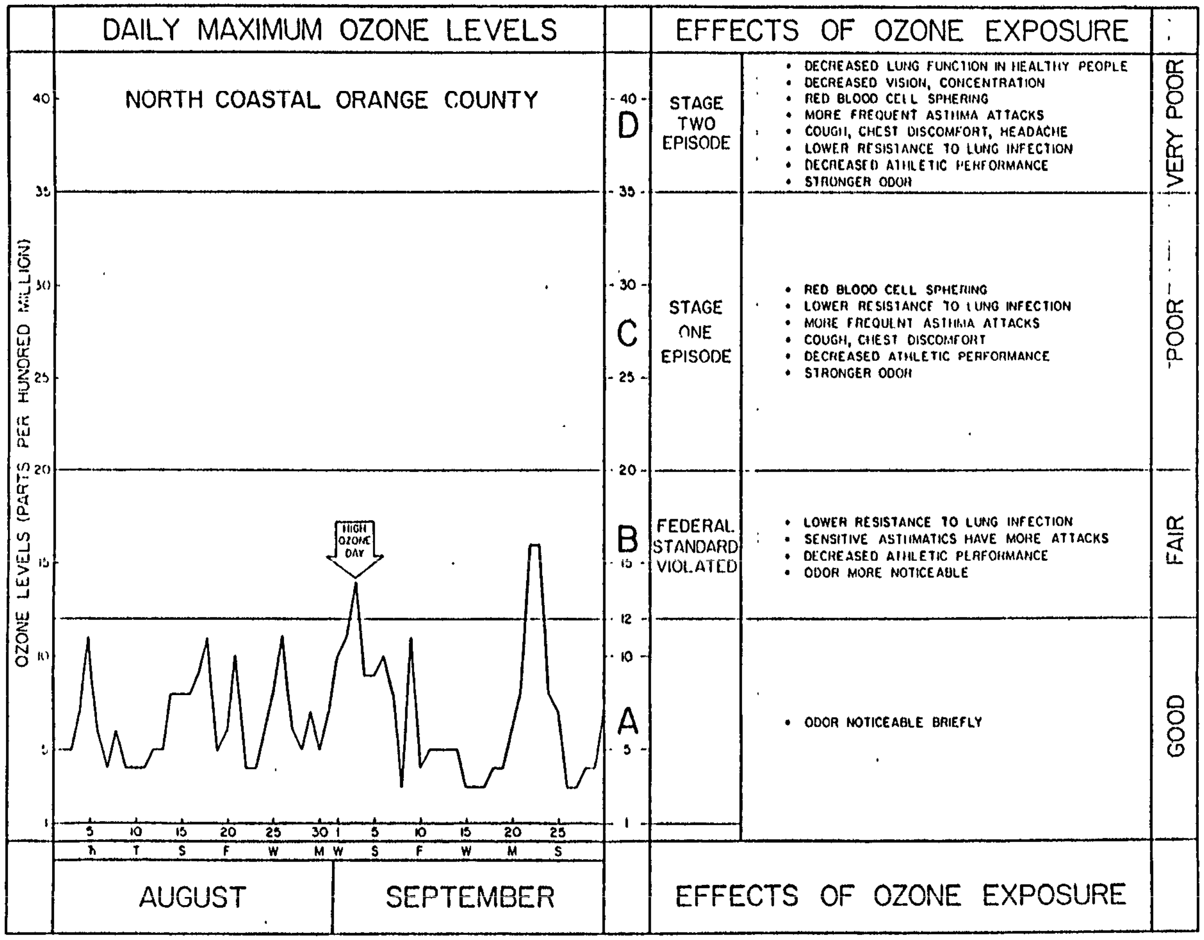

DAILY MAXIMUM OZONE LEVELS
EFFECTS OF OZONE EXPOSURE
WEST SAN FERNANDO VALLEY
OZONE LEVELS (PARTS PER HUNDRED MILLION)
40
35
30
25
20
15
12
10
5
1
HIGH OZONE DAY
D
C
B
A
STAGE TWO EPISODE
STAGE ONE EPISODE
FEDERAL STANDARD VIOLATED
• DECREASED LUNG FUNCTION IN HEALTHY PEOPLE
• DECREASED VISION, CONCENTRATION
• RED BLOOD CELL SPHERING
• MORE FREQUENT ASTHMA ATTACKS
• COUGH, CHEST DISCOMFORT, HEADACHE
• LOWER RESISTANCE TO LUNG INFECTION
• DECREASED ATHLETIC PERFORMANCE
• STRONGER ODOR
• RED BLOOD CELL SPHERING
• LOWER RESISTANCE TO LUNG INFECTION
• MORE FREQUENT ASTHMA ATTACKS
• COUGH, CHEST DISCOMFORT
• DECREASED ATHLETIC PERFORMANCE
• STRONGER ODOR
• LOWER RESISTANCE TO LUNG INFECTION
• SENSITIVE ASTHMATICS HAVE MORE ATTACKS
• DECREASED ATHLETIC PERFORMANCE
• ODOR MORE NOTICEABLE
• ODOR NOTICEABLE BRIEFLY
VERY POOR
POOR
FAIR
GOOD
5
10
15
20
25
30
1
5
10
15
20
25
Th
T
S
F
W
M
W
S
F
W
M
S
AUGUST
SEPTEMBER
EFFECTS OF OZONE EXPOSURE

Figure 5.10


DAILY MAXIMUM OZONE LEVELS
NORTH COASTAL ORANGE COUNTY
OZONE LEVELS (PARTS PER HUNDRED MILLION)
40
35
30
25
20
15
12
10
5
1
HIGH OZONE DAY
5 10 15 20 25 30 1 5 10 15 20 25
Th T S F W M W S F W M S
AUGUST
SEPTEMBER
D
C
B
A
EFFECTS OF OZONE EXPOSURE
STAGE TWO EPISODE
• DECREASED LUNG FUNCTION IN HEALTHY PEOPLE
• DECREASED VISION, CONCENTRATION
• RED BLOOD CELL SPHERING
• MORE FREQUENT ASTHMA ATTACKS
• COUGH, CHEST DISCOMFORT, HEADACHE
• LOWER RESISTANCE TO LUNG INFECTION
• DECREASED ATHLETIC PERFORMANCE
• STRONGER ODOR
STAGE ONE EPISODE
• RED BLOOD CELL SPHERING
• LOWER RESISTANCE TO LUNG INFECTION
• MORE FREQUENT ASTHMA ATTACKS
• COUGH, CHEST DISCOMFORT
• DECREASED ATHLETIC PERFORMANCE
• STRONGER ODOR
FEDERAL STANDARD VIOLATED
• LOWER RESISTANCE TO LUNG INFECTION
• SENSITIVE ASTHMATICS HAVE MORE ATTACKS
• DECREASED ATHLETIC PERFORMANCE
• ODOR MORE NOTICEABLE
• ODOR NOTICEABLE BRIEFLY
VERY POOR
POOR
FAIR
GOOD
EFFECTS OF OZONE EXPOSURE

Symptom	Yourself	Family Member
Decreased Vision	________	________
More frequent asthma attacks	________	________
Cough, Chest discomfort	________	________

Other (please name)________________________________

Following this, the payment mechanism is introduced and a bid is solicited for specified reductions in ozone levels from the designated peak. Three bids are solicited in the San Gabriel Valley, two in the San Fernando Valley and one in Orange County. Questions from the San Gabriel Valley are:

3. What is the most your household would be willing to pay to reduce the daily high ozone reading on that day from VERY POOR to POOR? Please circle your answer.

$ .00	$2.00	$4.00	$6.00	$8.00	$11.00	$15.00	$35.00
$ .50	$2.50	$4.50	$6.50	$8.50	$12.00	$20.00	$50.00
$1.00	$3.00	$5.00	$7.00	$9.00	$13.00	$25.00	$75.00
$1.50	$3.50	$5.50	$7.50	$10.00	$14.00	$30.00	$100.00

4. What is the most your household would be willing to pay to reduce the daily high ozone level on that day from VERY POOR to FAIR? Please circle your answer.

$ .00	$2.00	$4.00	$6.00	$8.00	$11.00	$15.00	$35.00
$ .50	$2.50	$4.50	$6.50	$8.50	$12.00	$20.00	$50.00
$1.00	$3.00	$5.00	$7.00	$9.00	$13.00	$25.00	$75.00
$1.50	$3.50	$5.50	$7.50	$10.00	$14.00	$30.00	$100.00

5. What is the most your household would be willing to pay to reduce the daily high ozone level on that day from VERY POOR to GOOD? Please circle your answer.

$ .00	$2.00	$4.00	$6.00	$8.00	$11.00	$15.00	$35.00
$ .50	$2.50	$4.50	$6.50	$8.50	$12.00	$20.00	$50.00
$1.00	$3.00	$5.00	$7.00	$9.00	$13.00	$25.00	$75.00
$1.50	$3.50	$5.50	$7.50	$10.00	$14.00	$30.00	$100.00

Immediately following the bid(s), the respondent is asked why they bid zero if they did.

The respondents are then asked the extent of their outdoor activities and how or if they change their behavior when ozone levels rise.

The survey is concluded with a series of demographic questions. Included in the series is a question asking whether or not air quality was considered in residential choice.

Survey Procedures - Mail

Execution of the mail survey was accomplished obtaining current street address telephone directories for each survey area. These documents, available from the local telephone utility, contain listed telephone service customers arranged by street address rather than alphabetically in each service area. From these were taken residential addresses within the preselected survey area. An initial goal of 500 mailings in each area was modified to accommodate somewhat fewer than anticipated customers with listed numbers in some of the areas.

The surveys were mailed during the first week of December, 1982. All responses received before January 15, 1983 were included in the sample if they were completed. Four responses not included in the sample were received between January 15 and February 15, 1983.

A series of mechanical and procedural errors resulted in a very small mailing to El Monte in December, 1982. To remedy this two additional mailings were required. The response rates were nearly identical in all three mailings. The results are treated as one group because of the small numbers in each mailing response.

No follow-up mailings or telephone calls were attempted. This strategy was adopted to examine the potential of a low-cost contingent valuation of environmental amenities. Such a device, if workable, would be useful in the conduct of policy research regarding national or regional rather than local amenities.

A possible extension of this approach could include a second mailing to increase response rates. Such an effort would have to be very carefully structured, though, since it would involve either the sacrifice of respondent anonymity or the possibility of dual responses from some respondents.

Survey Procedures - Interview

A field supervisor was retained in Los Angeles to recruit and manage interviewers and to review completed interview forms prior to their shipment to Laramie. The supervisor is an individual experienced in, among other things, hiring and training interviewers and managing fieldwork. He has considerable experience and has successfully completed similar assignments for other research groups. Interviewers were selected principally on the basis of successful experience in similar survey efforts. Other relevant criteria were availability of dependable transportation, perceived ability to deal effectively with at least one of the sample populations and interviewing skills.

A member of the project team traveled to Los Angeles to conduct a training session with the field supervisor and interviewers. The training session provided an opportunity for personal interaction with the interviewers as well as describing project objectives.

The session provided information to interviewers regarding the concept of benefit measurement, a review of previous related efforts and mock interviews. The interviewers were reminded not to provide additional information to respondents about the research sponsors or its applications.

The training session was a valuable part of the survey effort with interviewers gaining an understanding of the significance of the interview process as a part of benefits assessment.

Interviews were conducted during December 7-18, 1982, during the late morning and afternoon. Interviews were conducted on weekends, as well as weekdays to provide a full range of potential respondents.

Each interviewer was provided with a list of residents who had been sent mail surveys and a street map of the survey area. They were instructed to include all portions of the survey area in their attempts while avoiding residences to which a survey form had been mailed. In two of the survey areas (Canoga Park and Newport Beach) the interviewers were obliged to survey in adjacent areas of similar appearance to complete the desired number of interviews.

C.3 Survey Results

There was considerable variation in response rates among the five survey areas. Table 5.2 presents response information for both interview and mail survey efforts.

The interview response rate for resident contacts (those attempts when an adult-resident came to the door) varied from 24% in Canoga Park to 56% in La Canada. There is of course no comparable rate for the mail survey.

Survey response rates are plotted against mean household income in Figure 5.11. The most obvious pattern that emerges is that the contact response rate for interviews was in all cases higher than the mail response rate. This is hardly surprising. There is no consistent pattern within either the mail or interview groups. The Orange County communities had the highest mail response rates but were in the middle of the income range for the communities.

Within air quality areas, the higher income communities had lower mail response rates in Orange County and the San Fernando Valley, but higher in the San Gabriel Valley. The San Fernando Valley interview effort reversed this, with the higher income community having a higher response rate. The San Gabriel Valley communities had the highest response rates.

TABLE 5.2: RESPONSE RATES AND RELATED INFORMATION

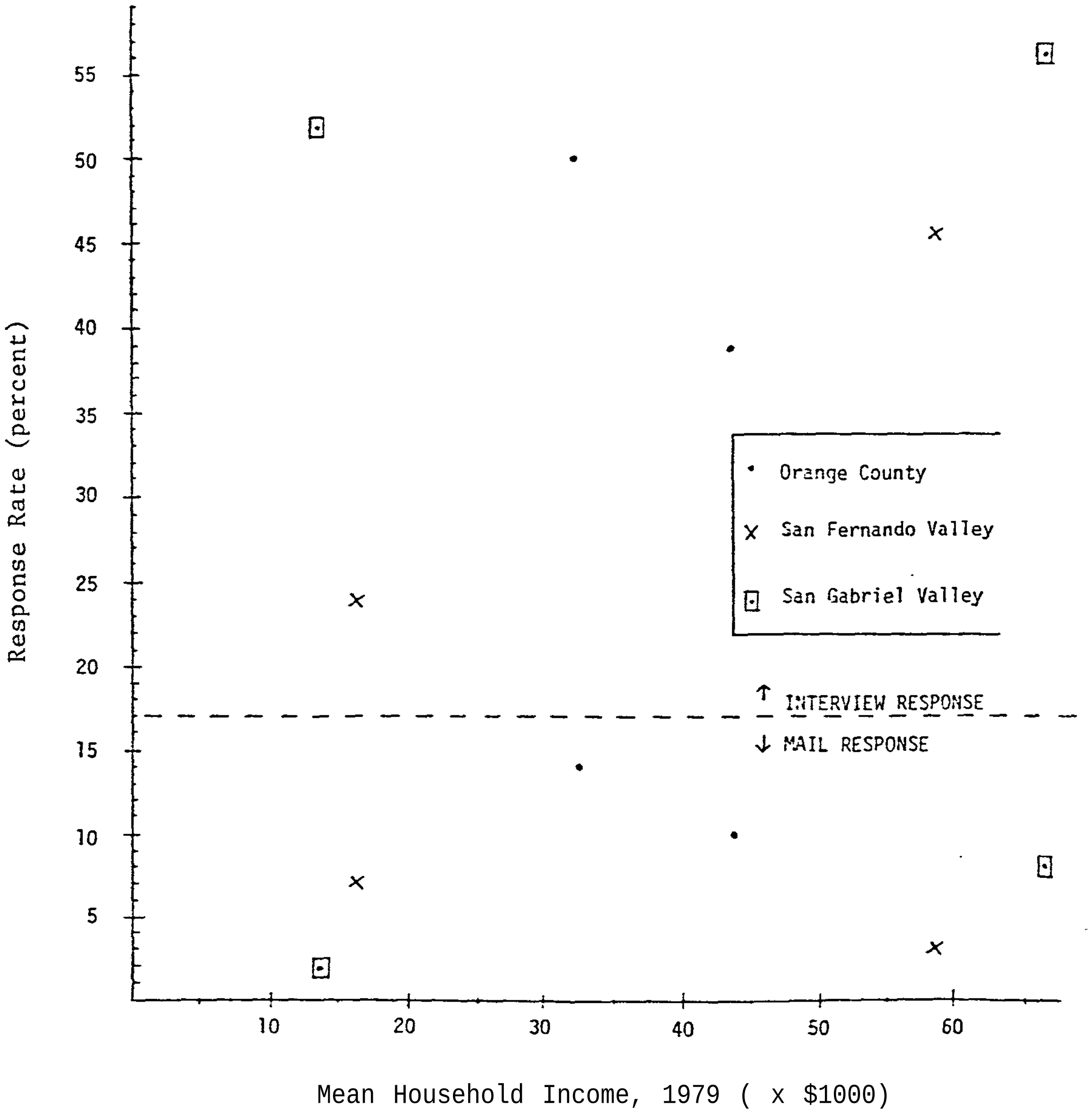
		A	B	C	D	E	F	G
					Reason for Refusal			
		Total Attempts	Resident Answered Door	Refusals	Do Not Consider Ozone to be a Problem	Too Busy	Other	No Reason Given
Communities Surveyed by Interview	El Monte	54	44	14	2	4	7	1
	La Cañada	58	32	8		5	2	1
	Canoga Park	175	90	65	6	26	29	4
	Encino	80	33	14		6	7	1
	Irvine	55	36	11	1	7	1	2
	Newport Beach	94	46	20		14	2	4
Communities Surveyed by Mail	El Monte	519						
	La Cañada	401						
	Canoga Park	295						
	Encino	616						
	Irvine	383						
	Newport Beach	408						

(Table 5.2, continued)

Table 5.2 (continued)

		H	I	J	K	I	M
		Flawed Surveys	Completed Surveys	Cross Response Rate (I÷A)	Non-Protest Surveys	Net Response Rate (K÷A)	Contact Response Rate (K÷B)
Communities Surveyed by Interview	El Monte	2	28	.52	23	.43	.52
	La Cañada	4	20	.34	18	.31	.56
	Canoga Park	1	24	.14	22	.13	.24
	Encino	5	19	.24	15	.19	.45
	Irvine	1	24	.43	18	.32	.50
	Newport Beach	1	25	.27	18	.19	.39
Communities Surveyed by Mail	El Monte	1	15	.03	11	.02	
	La Cañada	6	37	.09	32	.08	
	Canoga Park	15	22	.07	20	.07	
	Encino		23	.04	19	.03	
	Irvine		60	.16	53	.14	
	Newport Beach	18	52	.13	42	.10	

Figure 5.11: Response Rates and Income for Survey Areas



In short, neither mean household income nor air quality within a survey area has an obvious relationship to response rates for either mail or interview surveys.

The net response rate (percentage of survey attempts resulting in completed surveys that did not protest the fairness of a pollution-reduction charge) was as might be expected, higher for the interview survey than for the mail survey.

Responses to survey questions are summarized in Table 5.3. The responses are grouped by air quality area.

The responses to question 7 are scaled as 1, 2, and 3 respectively for Rarely, Occasionally and Often and summed for each respondent. This produces an index of outdoor activity with a potential range of 0-24.

Apart from the bids (which are examined more closely below) there appears to be a remarkable similarity between mail and interview respondents in each air quality area. Mean years in current residence (#9) and mean years in the Los Angeles area (#10) are very close for both mail and interview samples. Mail respondents tend to he somewhat older (#15) and more educated (#14) than interview respondents and are much more likely to be male (#16). This difference presumably reflects the fact that interviews were conducted on weekday afternoons as well as evenings and weekends.

Apart from these responses, no clear pattern emerges to differentiate mail and interview respondents across air quality areas: San Gabriel Valley (SG) mail respondents noticed ozone-induced symptoms more often but had lower mean bids; in the San Fernando Valley (SF) mail respondents in Encino noticed ozone-induced effects less often and bad higher mean bids while Canoga Park residents noticed the effects more often and had higher mean bids. Orange County (OR) mail respondents noticed the effects less often and had lower mean bids. Mean income was lower for mail respondents in SG, higher in SF and OR.

C.4 Analysis of the Data

The survey results are examined through three different techniques in an attempt to discern meaningful patterns in respondents' bids.

Tables 5.4-5.6 report the results of linear regression models of each bid level. That is, the bid of each specified ozone reduction is entered as the dependent variable in the regression. The bid is "explained" by the selection of independent variables: household income (INC), education (ED), an index of outdoor activities (ACT), and either years in current residence (YH) or years in the Los Angeles area (YLA). A separate equation is calculated for interview and mail respondents in each air quality area. While these equations have limited explanatory power, as measured by each equation's R^2, some of the results do warrant comment.

TABLE 5.3

Question #.	1	3	4	5	7	9	10	11
Community	Symptom % Yes (SD)	CBID ($) Mean (SD)	BBID ($) Mean (SD)	ABID ($) Mean (SD)	Activity (Index) Mean (SD)	Years in House (Yrs) Mean (SD)	Years in L.A (Yrs) Mean (SD)	Consider Air Quality (% Yes)
La Cañada								
Interview N=18	16.7	15.92 (31.18)	16.92 (31.05)	24.75 (36.08)	8.06 (5.05)	11.83 (9.62)	27.56 (16.92)	55.6
Mail N-32	46.9	9.70 (18.59	13.66 (19.83)	20.97 (26.24)	7.00 (4.33)	12.03 (10.09)	28.56 (19.45)	40.6
El Monte								
Interview N=23	21.7	3.61 (7.32)	5.17 (9.50)	11.30 (25.24)	3.09 (3.41)	10.00 (9.72)	26.17 (17.74)	17.4
Mail N=11	63.6	1 82 (2.05)	3.73 (2.90)	15.86 (28.71)	6.36 (4.99)	11.82 (12.67)	23.82 (15.32)	9.1
Canoga Park								
Interview N=22	27.3		4.82 (6.40)	8.59 (14.01)	7.77 (6.18)	5.64 (5.63)	18.77 (10.18)	13.6
Mail N=20	30.0		7.53 (22.15)	7.75 (22.10)	5.40 (2.52)	4.45 (3.36)	19.75 (16.28)	40.0
Encino								
Interview N=15	60.0		2.57 (4.17)	3.23 (4.79)	4.27 (3.86)	8.27 (8.07)	21.73 (14.34)	6.7
Mail N=19	31.6		8.18 (12.84)	12.21 (22.48)	7.21 (4.10)	10.37 (8.04)	24.11 (18.19)	26.3
Irvine								
Interview N=18	38.9			16.08 (31.37)	4.22 (3.19)	4.67 (2.97)	24.28 (17.75)	94.4
Mail N=53	22.6			4.46 (5.58)	9.04 (4.00)	4.79 (3.23)	14.02 (13.26)	71.7
Newport Beach								
Interview N=18	38.9			9.83 (25.63)	7.22 (4.49)	12.33 (6.61)	20.50 (11.76)	72.2
Mail N=42	19.0			4.77 (15.41)	6.55 (3.62)	12.81 (8.79)	31.67 (19.74)	73.8

(Table 5.3, continued)

Table 5.3 (continued)

Question #:	12	14	15	16	17	18	19	20	21
Community	Info Index Mean (SD)	Education (Years) Mean (SD)	Age (Years) Mean (SD)	Gender % Male	Household Size (Persons) Mean (SD)	Primary Earner %	Residence (% Detached) (SD)	Own or Rent (% Own) (SD)	Income ($000) (SD)
La Canada									
Interview N=18	1.28 (.83)	15.44 (2.26)	41.72 (14.64)	22.2	3.78 (1.52)	16.7	100.0	94.4	68.72 (20.29)
Mail N=32	1.53 (.80)	16.63 (1.56)	48.75 (10.85)	81.3	3.72 (1.55)	87.5	100.0	93.8	54.84 (18.67)
El Monte									
interview N=	1.44 (.90)	12.17 (2.08)	44.30 (13.53)	43.5	3.35 (1.70)	65.2	87.0	56.5	14.83 (9.44)
Mail N=	1.73 (.79)	13.27 (1.62)	35.46 (21.02)	36.4	2.73 (2.01)	63.6	72.7	63.6	18.09 (11.53)
Canoga Park									
Interview N=22	1.82 (.91)	13.64 (1.92)	31.32 (10.15)	54.5	3.32 (1.56)	40.9	77.3	31.8	23.68 (14.82)
Mail N=20	1.35 (.88)	15.00 (2.29)	36.70 (11.24)	90.0	2.10 (1.45)	75.0	25.0	25.0	28.30 (20.26)
Encino									
Interview N=15	1.00 (.66)	13.20 (1.66)	43.13 (16.61)	40.0	2.47 (1.30)	26.7	66.7	73.3	36.20 (20.55)
Mail N=19	1.63 (.68)	(1.66) (1.41)	41.47 (13.45)	68.4	2.53 (1.31)	63.2	42.1	52.6	52.68 (21.91)
Irvine									
Interview N=18	1.44 (.71)	13.89 (1.45)	35.11 (12.62)	38.9	3.33 (1.09)	38.9	100.	88.9	35.33 (11.11)
Mail N=53	1.40 (.91)	16.26 (1.76)	39.49 (9.68)	77.4	3.26 (1.24)	86.8	98.1	86.8	46.89 (16.80)
Newport Beach									
Interview N=18	1.22 (.65)	15.78 (1.80)	40.06 (13.67)	33.3	3.56 (1.20)	16.7	100.	94.4	53.17 (16.39)
Mail N=	1.60 (.89)	16.00 (1.71)	51.19 (11.00)	85.7	2.48 (1.04)	92.9	81.0	85.7	54.05 (.)

TABLE 5.4

REGRESSION RESULTS FOR BID ESTIMATES
SAN GABRIEL VALLEY SURVEY

	R^2	CONST	INC	YH	YLA	ED	ACT
Mean (Standard Deviation)	Beta Coefficients (t-Statistic)						
CBID							
INTERVIEWER RESPONSES:							
9.01 (21.93)	.30	-40.23 (-2.13)	-.23 (-1.57)	.05 (.16)		3.48 (2.18)	1.93 (2.41)
9.01 (21.93)	.31	-43.82 (-2.28)	-.24 (-1.64)		.12 (.66)	3.55 (2.24)	1.98 (2.51)
MAIL RESPONSES:							
7.69 (16.38)	.30	-27.32 (-1.55)	.19 (1.57)	.65 (2.98)		.62 (.47)	1.29 (2.38)
7.69 (16.38)	.25	-37.57 (-1.95)	.08 (.62)		.31 (2.37)	1.64 (1.20)	1.08 (1.96)
BBID							
INTERVIEW RESPONSES:							
10.33 (22.23)	.33	-37.52 (-2.00)	-.26 (-1.82)	.03 (.10)		3.36 (2.13)	2.27 (2.86)
10.33 (22.23)	.34	-41.28 (-2.15)	-.27 (-1.89)		.12 (.65)	3.43 (2.18)	2.33 (2.99)
MAIL RESPONSES:							
11.12 (17.64)	.31	-23.83 (-1.26)	-30 (2.30)	.66 (2.83)		.37 (.26)	1.12 (1.93)
11.12 (17.64)	.21	-29.04 (-1.36)	.20 (1.45)		.21 (1.44)	1.25 (.82)	.84 (1.38)
ABID							
INTERVIEW RESPONSES:							
17.21 (30.81)	.24	-.29 (-.01)	-.12 (-.56)	-.17 (-.36)		.47 (.20)	3.32 (2.83)
17.21 (30.81)	.24	-.26 (-.01)	-.12 (-.56)		-.06 (-.24)	.45 (.19)	3.36 (2.89)
MAIL RESPONSES:							
19.66 (26.64)	.15	-5.21 (-.16)	.42 (1.93)	.52 (1.34)		-.19 (-.08)	.34 (.36)
19.66 (26.64)	.12	-8.50 (-.25)	.35 (1.56)		.15 (.64)	.48 (.20)	.11 (.12)

TABLE 5.5

REGRESSION RESULTS FOR BID ESTIMATES
SAN FERNANDO VALLEY SURVEY

	R^2	CONST	INC	YH	YLA	ED	ACT
Mean (Standard Deviation)	Beta Coefficients (t-Statistic)						
BBID							
INTERVIEW RESPONSES:							
3.90 (5.65)	.12	.29 (.04)	-.01 (-.19)	-.04 -.31		.16 (.26)	.32 (1.76)
3.90 (5.65)	.13	1.12 (.15)	-.02 (0.33)		-.05 (-.64)	.17 (.27)	.33 (1.83)
MAIL RESPONSES:							
7.26 (17.85)	.04	-7.84 (-.32)	-.02 (-.11)	.42 (.85)		.69 (.42)	.31 (.33)
7.43 (18.07)	.06	16.92 (-.63)	-.04 (-.26)		.23 (1.79)	1.09 (.62)	.59 (.62)
ABID							
INTERVIEW RESPONSES:							
6.42 (11.43)	.11	-2.37 (-.16)	-.10 (-.84)	-.12 (-.42)		.70 (.55)	.49 (1.32)
6.42 (11.43)	.11	-1.76 (-.11)	-.11 (-.96)		-.06 (-.37)	.71 (.56)	.50 (1.35)
MAIL RESPONSES:							
9.66 (22.34)	.06	-21.02 (-.69)	.04 (.23)	.45 (.72)		1.55 (.75)	.29 (.24)
9.86 (22.61)	.08	-31.82 (-.951	.01 (.06)		.27 (1.15)	2.02 (.93)	.56 (.47)

TABLE 5.6

REGRESSION RESULTS FOR BID ESTIMATES
ORANGE COUNTY SURVEY

	R^2	CONST	INC	YH	YLA	ED	ACT
Mean (Standard Deviation)	Beta Coefficients (t-Statistic)						
<u>ABID</u>							
INTERVIEW RESPONSES:							
10.83 (25.49)	.26	24.30 (.57)	-.22 (-.78)	-.88 (-1.14)		-.19 (-.37)	3.43 (2.61)
13.53 (29.50)	.19	92.95 (2.01)	.009 (.03)		.20 (.56)	-7.09 (-2.22)	3.35 (2.09)
MAIL RESPONSES:							
4.60 (10.99)	.01	-5.35 (-.47)	.03 (.49)	.0009 (.006)		.52 (.72)	-.02 (-.06)
4.60 (10.99)	.02	-8.94 (-.77)	.02 (.33)		.06 (.86)	.67 (.91)	.05 (.15)

The outdoor activity index (ACT) is the only variable that has even modest statistical significance in most of the equations. This finding is not startling; it even provides modest comfort that a variable so closely tied to outdoor air quality is not generally irrelevant. A noteworthy feature of ACT's pattern is that the sign of the coefficient is positive wherever it has even modest significance (the exception in fact has t = -.06).

In each air quality area the t-statistic is higher for ACT in the interview sample than for the mail sample. This difference is most extreme in Orange County.

The Orange County samples also show the most extreme difference in magnitude for the estimated coefficient of ACT. In SG the mail and interview ACT coefficients diverge with the degree of ozone reduction. That is, the ACT coefficients for CBID are comparable in both forms of the equation. The differences are greater for BBID and extreme for ABID.

The coefficients for ACT are all roughly comparable in the SF samples.

The Orange County mail and interview equations differ to an extent that is disturbing. This is especially so since the two Orange County communities were more similar than those in other air quality areas and had much higher mail response rates.

The most extreme difference between the mail and interview responses (Table 5.3) were in ABID (with mail lower) and percentage of respondents who were household primary earners (mail lower). This latter difference was seen in SG and SF also, but mail respondent bids were generally higher.

This consistency, with typical expectations, is not shown in other variables. ED, for instance, shows moderate statistical significance with positive coefficients in SG, but in SF has statistical significance in only one equation, when the coefficient is negative.

This general inconsistency of sign and statistical significance suggests that considerable subtlety will be necessary to provide explanation of ozone reduction bids.

To determine the influence of "outliers" on the regression estimates, a technique developed by Belsley, Kuh and Welsch (1980) (B-K-W) and previously applied by Desvousges, Smith and McGivney (1982) (D-S-M) was adopted. The B-K-W statistic, DFBETA, measures the effect of an individual observation on the estimated coefficients in a regression model.

It is estimated by Equation:

$$\text{DFBETA} \equiv b - b(i) = \frac{(X^T X)^{-1} x_i^T e_i}{1-h_i}$$

where h_i is $x_i (X^TX)^{-1} x_i^T$ and the e_i's are the ordinary least squares residuals.

Following D-S-M, ±30 percent in any coefficient was taken as the standard for defining an outlier. The number of outliers detected was quite small: 1 each in 2 of the 12 SG equations; 1 each in 2 of the eight SF equations; and 1, 3 and 4 in 3 of the four OR equations. The re-estimated equations, with outliers removed, are presented in Tables 5.7-5.9. These revised equations differ substantially only in the constant term ,which was in all cases the term associated with a large DFBETA.

An examination of the difference between the mail and interview samples is presented in Table 5.10. The mean and standard deviation of each sample bid is presented for the complete sample and for the sample with outliers removed from each of the two regressions. For each pair of mail and interview bid samples, Student's t is calculated. This statistic tests the hypothesis that the two samples are drawn from the same population, with the difference in the means being a result of variation in the population.

In no case can this hypothesis be rejected at the .05 level, and even at the .10 level the hypothesis can be rejected only in Orange County.

This result is remarkable for a number of reasons. The large difference in response rates might have been suspected of being an indication of mail respondent self-selection and thereby causing sample bias. This possibility seemed especially troubling given the inherent complexity of both the substantive material and the survey instrument.

The interview respondents, with interviewers available to explain the material, had a less rigorous experience. This complexity may have substantially contributed to the self-selection of mail respondents with higher mean education than interview respondents. The mail respondents had mean years of education at least one year higher than interview respondents in all communities except Newport Beach, which had the highest interview respondent education level, 15.78 years.

The mean bids have a large standard deviation in all communities at all levels. This is to be expected for valuation of a public good.

Private goods, the benefits of which can be appropriated exclusively by one user, have large variations in quantity purchased at a price that is uniform for all buyers. Demand estimation is accomplished by estimating intended, desired or potential purchases by different individuals at varying prices.

Public goods cannot, by definition, be made available in different amounts to separate users; they are available in the same amount to all users, as is air quality in a given area.

The estimation of "demand" in this case is accomplished by estimating the prices different users would be willing to pay for a given amount of

TABLE 5.7

REGRESSION RESULTS FOR BID ESTIMATES
(With Outliers Removed)
SAN GABRIEL VALLEY SURVEY

	R^2	CONST	INC	YH	YLA	ED	ACT
Mean (Standard Deviation)		Beta Coefficients (t-Statistic)					
CBID							
INTERVIEWER RESPONSES:							
11.25 (24.41)	.29	-46.77 (-1.98)	-.24 (-1.41)	.05 (.09)		3.90 (2.02)	2.04 (1.99)
10.31 (23.52)	.31	-50.39 (-2.24)	-.25 (-1.59)		.15 (.74)	3.88 (2.17)	2.18 (2.33)
MAIL RESPONSES:							
8.15 (16.90)	.33	-29.10 (-1.60)	.19 (1.56)	.73 (3.16)		.73 (.52)	1.22 (2.05)
7.90 (16.94)	.26	-43.62 (-2.01)	.05 (.37)		.33 (2.32)	2.12 (1.34)	.99 (1.62)
BBID							
INTERVIEW RESPONSES:							
12.86 (24.63)	.32	-43.16 (-1.84)	-.28 (-1.62)	.02 (.05)		3.72 (1.95)	2.38 (2.35)
11.80 (23.78)	.34	-47.44 (-2.13)	-.29 (-1.82)		.15 (.73)	3.71 (2.10)	2.54 (2.74)
MAIL RESPONSES:							
11.33 (18.14)	.36	-24.83 (-1.30)	.30 (2.33)	.79 (3.27)		.33 (.23)	1.16 (1.85)
11.08 (18.23)	.23	-31.28 (-1.31)	.19 (1.25)		.26 (1.64)	1.34 (.77)	.83 (1.23)
ABID							
INTERVIEW RESPONSES:							
21.59 (33.67)	.19	2.10 (.06)	-.12 (-.47)	-.31 (-.42)		.44 (.15)	3.25 (2.16)
19.81 (32.68)	-21	-.19 (-.01)	-.12 (-.50)		-.06 (-.21)	.40 (.15)	3.41 (2.45)
MAIL RESPONSES:							
20.33 (27.37)	.19	-10.66 (-.33)	.42 1.90)	.67 (1.61)		.26 (.10)	.08 (.08)
20.08 (27.51)	.15	-4.58 (-.12)	.39 (1.63)		.21 (.87)	.21 (.07)	-.21 (-.20)

TABLE 5.8

REGRESSION RESULTS FOR BID ESTIMATES
(With Outliers Removed)
SAN FERNANDO VALLEY SURVEY

	R^2	CONST	INC	YH	YLA	ED	ACT
Mean (Standard Deviation)		Beta Coefficients (t-Statistic)					
<u>BBID</u>							
INTERVIEW RESPONSES:							
3.90	.12	.29	-.01	-.04		.16	.32
(5.65)		(.04)	(-.19)	(-.31)		(.26)	(1.76)
3.90	.13	1.12	-.02		-.05	.17	.33
(5.65)		(.15)	(-.33)		(.64)	(.27)	(1.83)
MAIL RESPONSES:							
7.26	.04	-7.84	-.02	.42		.69	.31
(17.85)		(-.32)	(-.11)	(.85)		(.42)	(.33)
7.43	.06	16.92	-.04		.23	1.09	.59
(18.07)		(-.63)	(-.26)		(1.19)	(.62)	(.62)
<u>ABID</u>							
INTERVIEW RESPONSES:							
6.42	.11	-2.37	-.10	-.12		.70	.49
(11.43)		(-.16)	(-.84)	(-.42)		(.55)	(1.32)
6.42	-11	-1.76	0.11		-.06	.71	-50
111.43)		(-.11)	(-.96)		(-.37)	(.56)	(1.35)
MAIL RESPONSES:							
9.66	-06	-21.02	.04	.45		1.55	.29
(22.34)		(-.69)	(.23)	(.72)		(.75)	(.24)
9.86	.08	-31.82	.01		.27	2.02	.56
(22.61)		(-.95)	(.06)		(1.15)	(.93)	(.47)

TABLE 5.9

REGRESSION RESULTS FOR BID ESTIMATES
(With Outliers Removed)
ORANGE COUNTY SURVEY

	R^2	CONST	INC	YH	YLA	ED	ACT
Mean (Standard Deviation)	Beta Coefficients (t-Statistic)						
<u>ABID</u>							
INTERVIEW RESPONSES:							
10.83 (25.49)	.26	24.30 (.57)	-.22 (-.78)	-.88 (-1.14)		-.19 (-.37)	3.43 (2.61)
13.53 (29.50)	.19	92.95 (2.01)	.009 (.03)		.20 (.56)	-7.09 (-2.22)	3.35 (2.09)
MAIL RESPONSES:							
4.60 (10.99)	.01	-5.35 (-.47)	.03 (.49)	.009 (.006)		.52 (.72)	-.02 (-.06)
4.60 (10.99)	.02	-8.94 (-.77)	.02 (.33)		.06 (.86)	.67 (.91)	.05 (.15)

TABLE 5.10

t-TESTS FOR DIFFERENCES BETWEEN MAIL AND INTERVIEW SAMPLES

				Years in House Outliers Removed			Years in L.A. Outliers Removed		
San Gabriel	N	Mean (Stan. Dev.)	t-Stat	N	Mean (Stan. Dev.)	t-Stat	N	Mean (Stan. Dev.)	t-Stat
CBID									
Interview	41	9.01 (21.93)		34	10.60 (23.81)		32	11.25 (24.41)	
Mail	43	7.69 (16.38)	.31	41	7.76 (16.76)	.59	41	8.00 (16.71)	.64
BBID									
Interview	41	10.33 (22.23)		34	12.13 (24.05)		32	12.86 (24.63)	
Mail	43	11.17 (17.65)	-.18	41	11.29 (18:05)	.17	41	11.54 (17.96)	.26
ABID									
Interview	41	17.21 (30.81)		34	20.37 (33.01)		32	21.59 (33.67)	
Mail	43	19.67 (26.65)	-.39	41	20.20 (27.18)	.02	41	20.44 (27.04)	.16
San Fernando									
BBID									
Interview	37	3.91 (5.65)		37	3.91 (5.65)		37	3.91 (5.65)	
Mail	39	7.85 (17.99)	-1.30	38	8.03 (18.19)	-1.33	39	7.84 (17.99)	-1.30
ABID									
Interview	37	6.42 (11.43)		37	6.42 (11.43)		37	6.42 (11.43)	
Mail	39	9.92 (22.10)	-.87	38	10.13 (22.36)	-.91	39	9.92 (22.10)	-.87
Orange County									
ABID									
Interview	36	12.96 (28.41)		33	13.53 (29.50)		32	10.83 (25.49)	
Mail	95	4.66 (10.99)	-1.72*	95	4.60 (10.99)	-1.70*	94	4.65 (11.04)	1.33

*Reject H_o at .10 level

$H_o : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

the good. Since there is variation in individual preferences, one would expect large variation in this bid estimate just as one would expect large variation in quantity estimates for a private good at a particular price.

The third technique applied to the data examines changes in individual bids over ozone-reduction intervals rather than aggregating individual bids for a specific reduction.

In this effort an equation of the form

$$B_i = kR_i^{\eta}$$

is estimated, where B_i is a household's bid for the ith ozone-reduction interval, R_i is the reduction and A and n are coefficients to be estimated. (See Section B for an examination of theoretical aspects of this bid equation).

For each respondent there are three observations in the San Gabriel Valley (from D to C, from D to B and from D to A) and two in the San Fernando Valley (from C to B and from C to A). With only one bid per respondent, an estimate of the equations in Orange County would be meaningless.

To estimate the equations, the ozone reductions were taken to be from the midpoint of the reference interval to the midpoint of succeeding intervals. That is, R_1 in SG is from 38.75 pphm (the midpoint of D as depicted), to 27.5 pphm (the midpoint of C), or a reduction of 11.25. Similarly R_2 in SG is 17.75 (from D to the midpoint of B, 14.5) and R_3 is 32.25 (38.75 to 6.5).

In SF, bids begin at the midpoint of C (27.5) so that R_1 is a reduction of 6.5 and R_2 is 21.

The results of these efforts are presented in Figures 5.12-5.13. The vertical axis is number of respondents in each category. The bar to the left of the origin shows the number of respondents who bid zero at all levels (This does not include "protest zeroes").

The numbers to the right of the origin are values of n.

The distribution of values for n of respondents has a pronounced pattern: In the intermediate ozone level area sampled (SF) the range stops at approximately 1.0 except for one observation. All three observation in the 1.0-1.1 range actually have estimated values for n of 1.026. In the high ozone level area sampled (SG) estimated values for n continue beyond unity ranging beyond 15.

The termination, at approximately 1.0 exists in both interview and mail samples in the San Fernando Valley (with the one exception); the continuation of the range in the San Gabriel Valley likewise exists in both samples.

Figure 5.12: Individual Bid Elasticity Estimates and Zero Bids

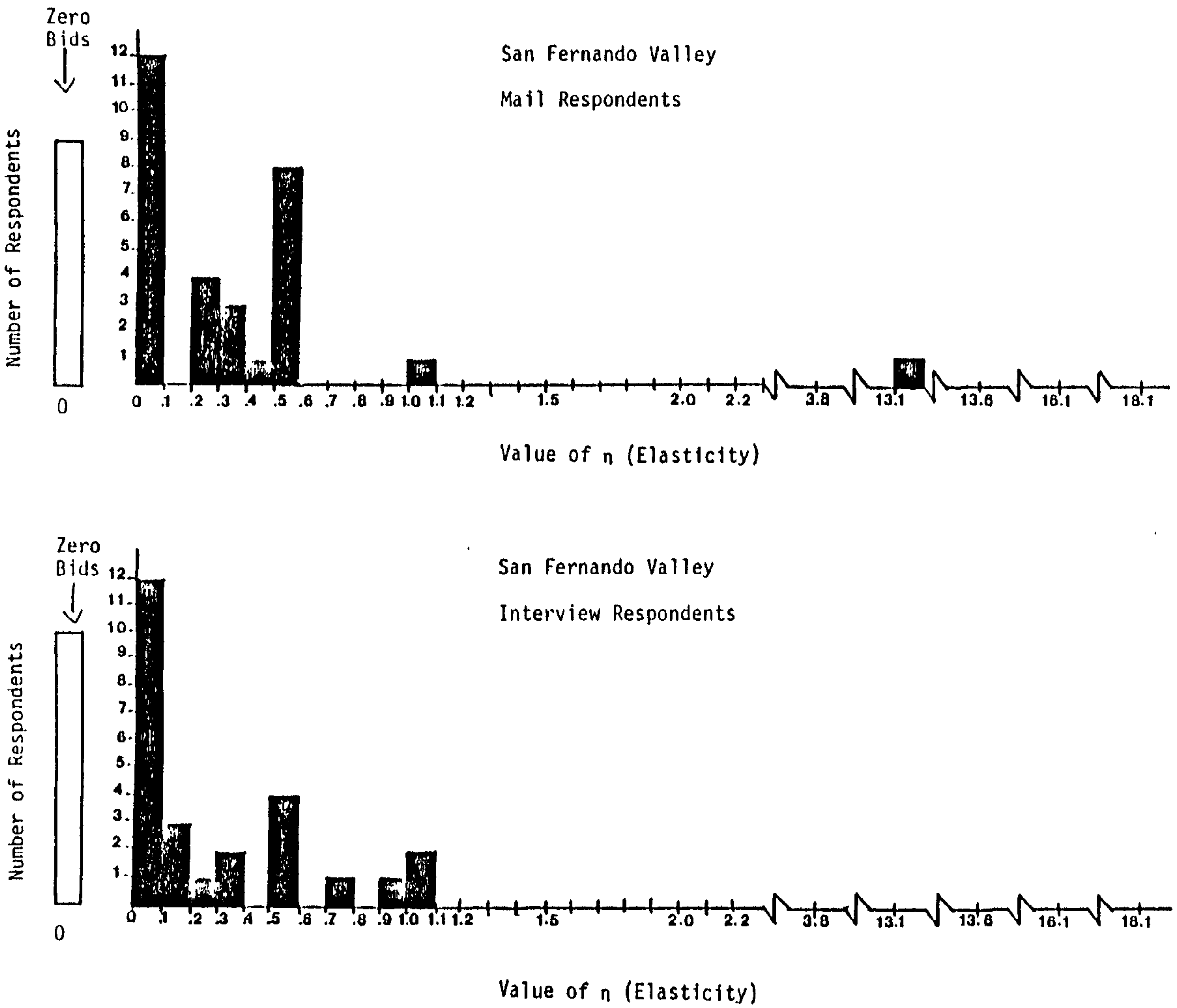
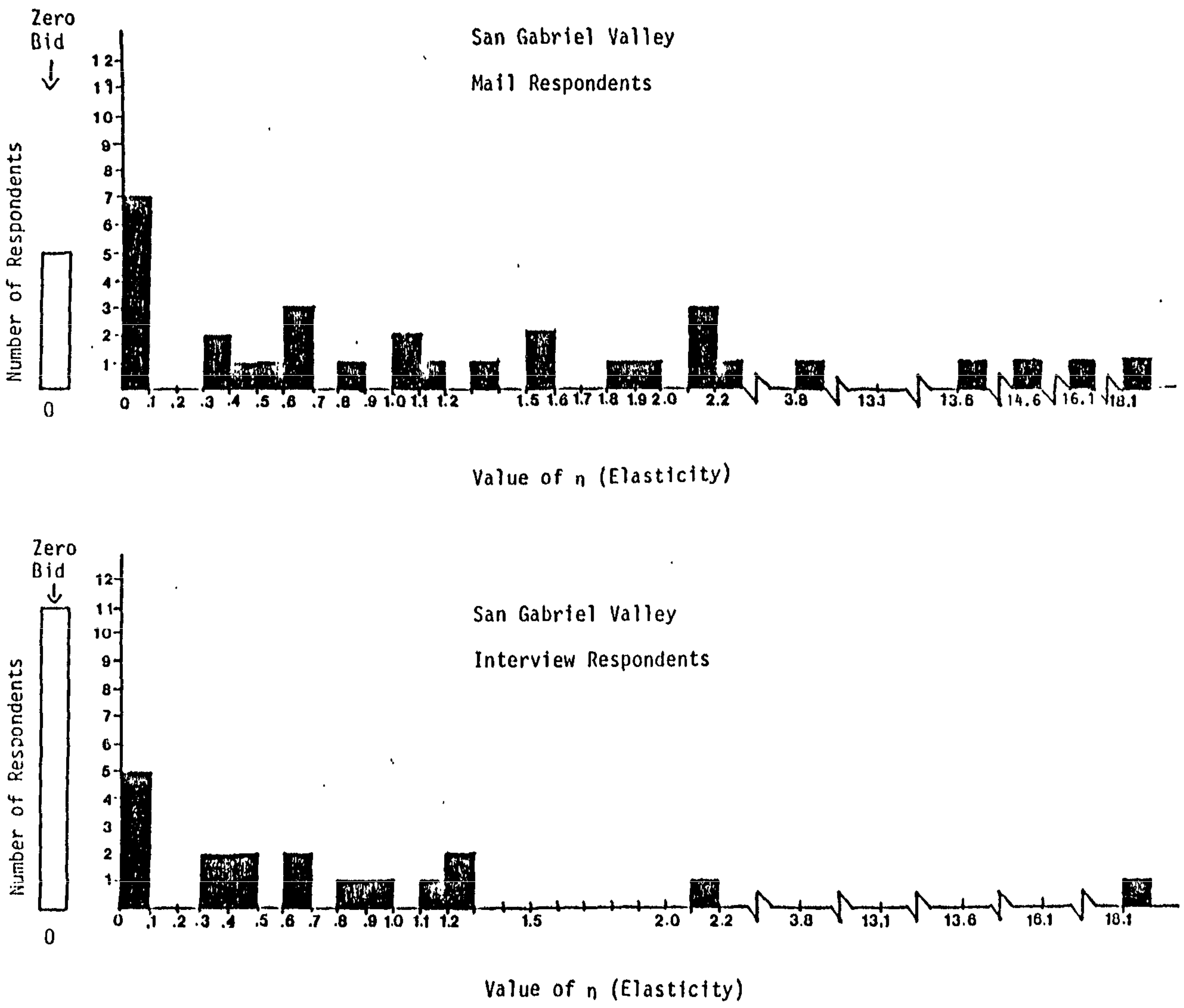


Figure 5.13: Individual Bid Elasticity Estimates and Zero Bids



As shown in Section B, values for this coefficient less than unity are consistent with the concave preference functions typically assumed by economists to exist. Values greater than unity indicate increasing marginal utility of ozone reduction. Individuals for whom n > 1 would be expected to locate themselves in areas of extreme air quality (whether high or low) unless there were a compelling preference unrelated to price and ozone levels in residential choice (a desire to be near one's job or one's childhood neighborhood for example).

This statement warrants some further elaboration, since it seems to suggest the existence of "extremists" who are little concerned with which extreme they choose.

A coherent description of the preferences of an individual with n > 1 would include the observation that such an individual places a relatively high value on preservation of air purity at a very high level. This person would place a lesser value on preservation of air purity if air quality had already been significantly degraded.

Conversely, a relatively low value would be placed on an incremental improvement in air quality unless the increment would "restore" pristine air. Each succeeding increment would have higher value. The final increment would have a higher value than any preceding improvement.

This person, with non-convex preferences, is to be contrasted with the typical person found in economic analysis who places ever smaller value on succeeding increments in availability of any good. The improvement that brings air quality to a pristine state from a slightly impaired condition would be valued less than a similar improvement in seriously degraded air. This parallels the expectation that a given ration of food would be valued more if a person had been deprived of food than if the same person were near full satiation.

Individuals with convex (i.e., "normal") preferences may have very different tastes regarding air quality. Some may place very high values on cleaner air and others may regard air quality as insignificant relative to all other considerations in residential location. The convex indifference curve shown as Figure 5.1 implies only that successive improvements in air quality have values that are less than earlier improvements. These early improvements may have very high as well as very low values.

These "normal" individuals can "purchase" a combination of air quality and other goods by choosing a location along the pollution-rent gradient depicted in Figure 5.5.

Individuals with non-convex preferences, though, would not be inclined to choose any intermediate level of air quality.

If, from a location with lowest air quality, such person were willing to "purchase" a small improvement (by moving to an area with slightly higher air quality), he or she would be willing to purchase more since each successive improvement has higher value.

With such a preference system, a person would be inclined to choose the highest possible air quality. If the premium for this level, though, were deemed to exceed the value, the second choice would not be some intermediate air quality location, but an area with low air quality.

These individuals differ from those with convex preferences not (necessarily) in the strength of their preferences for clean air as opposed to other goods but in the relative assessment of the value of improvements in air quality.

Thus we might find as neighbors in a low-pollution area one person with convex preferences who places a very high value on a small initial improvement and very small value on succeeding improvements and another person who places very small values on any improvement in air quality unless it brings pristine air.

The former would be little inconvenienced if local air were slightly degraded. The latter would protest vigorously or move.

Similarly, a high pollution area might contain some people who would make substantial sacrifices for a small improvement in air quality (but less than the housing-cost differential of such an area) and others who would make essentially no sacrifice unless it would bring pristine air.

These are of course the extreme cases. The important point is that persons with non-convex preferences would not generally locate in areas of intermediate air quality. The individuals are, of course, concerned with which extreme they choose.

San Fernando Valley respondents had, with one exception, convex preferences. San Gabriel Valley respondents included a number of people with non-convex preferences.

This distribution of preferences is that implied by the theoretical development in Section B. A very small number of individuals with non-convex preferences would be expected in intermediate air quality areas of other communities to exist with similar amenities differing only in air quality.

The Los Angeles area, with its very diverse mix of neighborhoods would be expected to offer very high or very low air quality locations with amenities similar to the San Fernando Valley communities in this study. Indeed, one suspects almost any conceivable amenity mix could be found.

The agreement between the pattern implied by a theoretical consideration of location choice and the estimated values of n in high and intermediate ozone level communities is rather dramatic.

The coefficient n can be thought of as a variable reflecting tastes. That it appears to be significant in residential choice suggests that examination of other variables reflecting tastes might be fruitful.

The activity index, ACT, used in regression models can also plausibly be interpreted as a taste variable. Given the broader range of "tastes", as measured by n, extent in the San Gabriel Valley than in the San Fernando Valley, one might expect a taste variable to have more significance in SG than in SF. This is so in the mail sample, but not in the interview sample. In fact, in SF and OR regressions ACT carries substantially more significance in interview than mail samples.

A greater relative importance of taste in explaining bids is, however, suggested by the much larger coefficients for ACT in SG and OR than in SF in cases where the coefficient has even a low level of statistical significance.

Opportunities for further research are indicated by the apparent complexity of the patterns involving survey response, respondent location decisions and other characteristics and bid levels.

D. THE PROPERTY VALUE APPROACH

D.1 Introduction

Previous research efforts have found survey results to be generally consistent with the hedonic housing value approach (Brookshire, et al., 1982), a hedonic wage analysis (Cummings, et al., 1982) and the recreation-based travel cost method (Desvousges, et al., 1982). In addition, surveys have been found to be internally consistent and compatible with demand theory (Schulze, et al., 1981). However, the debate over the validity of survey results continues in spite of these previous successes.

The purpose of the research reported in this section is to add to the literature concerning the validity of surveys designed to ascertain the value of environmental goods. This is accomplished by undertaking a detailed analysis of the relationship between housing values and ozone concentrations in the South Coast Air Basin. The objective was to develop an ozone based rent differential to compare to the survey results presented in the previous section. This is in accord with the theoretical treatment in Brookshire, et al. (1982).

The research described herein encompasses two separate but related housing value studies. First, the housing value analysis was conducted in Los Angeles County. Second, the study area was expanded to include the remainder of the South Coast Air Basin (Orange County, Riverside County, San Bernardino County). This was done to overcome empirical difficulties. The research was directed at determining whether housebolds actually pay for cleaner air in the form of higher housing values for homes in clean air communities and if this willingness to pay was comparable to the hypothetical willingness to pay expressed in the survey instrument.

Valuation of reductions in urban air pollution concentrations based upon housing value differentials is the most common form of the hedonic price procedure as developed by Rosen (1974), the basis of which is Lancaster's (1966) consumption theory. This procedure assumes that access to environmental (dis)amenities is capitalized in property values. This assumption is based on the premise that households are willing to pay a premium for an otherwise identical home located in a clean air area versus that located in a polluted area.

Among public goods which have been valued using the hedonic housing approach are air pollution (Anderson and Crocker, 1971; Harrison and Rubinfeld, 1978), social infrastructure (Cummings, 1978) and other community characteristics such as noise level (Nelson, 1979) and ethnic composition (Schnare, 1976).

The hedonic approach for assessing the benefits of environmental improvement is generally viewed as a multistage procedure (see Rosen, 1974; Freeman, 1979). The initial step is to estimate the hedonic price gradient which explains home sale price as a function of the house's structural characteristics as well as the characteristics of the community and neighborhood in which it is located, The second step is to determine'the implicit price of environmental change by differentiating the hedonic rent gradient with respect to the variable of interest. Subsequent steps include estimation of the inverse demand curve and integration to obtain benefit estimates.

The hedonic procedure as outlined above was generally well-received by the economics profession until just recently. However, a number of authors, including Brown and Rosen (1982), Mendelsohn (1981), and Palmquist (1982) have criticized the approach as not possessing sufficient information to identify the (inverse) demand curve in the subsequent steps. For this reason the methodology employed here is to follow Brookshire, et al. (1982) and conduct the validation test using the rent differential (second step) rather than actual benefit estimates.

Elimination of the theoretical problem of direct benefit estimation in the hedonic format does not, however, eliminate all potential difficulties. Estimation of the hedonic price gradient must be completed within the confines of the data. Problems which generally arise in housing value studies are misspecification and multicollinearity. The latter is especially problematical in this study. So much so that a large portion of the research reported herein is directed at attempting to solve this problem.

The central point is that the completion of a housing value study is not without theoretical and empirical difficulties. In this case the estimation problems are such that it is difficult to delineate explicitly the relationship between ozone concentrations and housing values. However, an estimated relationship between ozone and home sale price is obtained through the use of principal components analysis. In the next section this relationship is used to test the validity of the survey results. Preliminary indications are that surveys provide reasonable values for ozone reductions.

This section is organized as follows. In the following sub-section a discussion of the characteristics of the data is presented. Sub-section D.3 describes the empirical procedure and the base empirical results for Los Angeles County. As is described these results are beset with multicollinearity. Thus, a variety of solutions to this problem, with associated results, are presented in sub-section D.4. None of the solutions described in this section provide a satisfactory outcome. However, in sub-sections D.5 and D.6, two solutions which yield the expected relationship between home sale price and ozone concentrations are described. Sub-section D.7 offers summary remarks.

D.2 Data Specifics

The hypothesis to be tested is whether or not ozone concentrations are a significant determinant of housing sale price. The study area is first Los Angeles County and then the entire South Coast Air Basin, and is specifically confined to single family residences. Thus, not considered is the impact of ozone concentrations upon other structures (multiple family dwellings, mobile homes, commercial, etc.) or other ownership types (rental leasing, etc.). Therefore, within our sample, this research asks if households will pay a premium in the form of higher housing values for homes located in clean air areas and what is the magnitude of that willingness to pay.

The data base was constructed to enable the testing of hypotheses concerning the impact of ozone differences on housing sale price. The dependent variable in the entire analysis is the sale price of owner occupied single family residences.[1] The independent variable set consists of variables which correspond to three levels of aggregation: house, neighborhood, and community. Table 5.11 describes further the data employed in the study.

The housing characteristic data, obtained from the Market Data Center (a computerized appraisal service centered in Los Angeles), pertains to homes sold in the 1978-79 time period and contains information on nearly every important structural and/or quality attribute.[2] It should be emphasized that housing data of such quality (e.g., micro level of detail and over time) is rarely available for studies of this nature. Usually outdated data which are overly aggregate and not collected on a regular basis (for instance census tract averages only in census years) are employed. These data yield functions relevant for the "census tract" household but are only marginally relevant at the household (micro) level. Further, it is imperative that the rent differential is calculated at the household level for comparison with the survey results.

The initial empirical analysis was confined to Los Angeles County for the 1978-79 period. The Market Data Center provided computer data tapes

TABLE 5.11

VARIABLES USED IN ANALYSIS OF HOUSING MARKET FOR 1978-79

Variable	Definition (assumed effect on housing sale price)	Units	Source
<u>Dependent</u>:			
Sale Price	Sale price of owner occupied single family residences	($100)	Market Data Center
<u>Independent-Housing</u>:			
Sale Date	Month the home was sold (positive)	January 1978 = 1 December 1979 = 24	Market Data Center
Age	Age of home (negative)	Years	Market Data Center
Bathrooms	Number of bathrooms (positive)	Number	Market Data Center
Living Area	Square Feet of Living Area (positive)	Square Feet	Market Data Center
Pool	1 if pool, 0 if no pool (positive)	0 = no pool 1 = pool	Market Data Center
Fireplaces	Number of fireplaces (positive)	Number	Market Data Center
<u>Independent-Neighborhood</u>:			
Distance to Beach	Miles to nearest beach (negative)	Miles	Calculated
Age Composition	Percent Greater than 62 in Census Tract (positive)	Percent	1980 Census
Ethnic Composition	Percent White in Census Tract (positive)	Percent	1980 Census
Time to Work	Average time to Employment from Census Tract (negative)	Minutes	1980 Census
View	1 if view present, 0 if not (positive)	0 = no view 1 = view	Market Data Center
<u>Independent-Community</u>:			
School Quality	Community's 12th grade math score (positive)	Percent	California Assessment Program (1979)
Population Density	Population per square mile in surrounding community (negative)	Persons/square mile	1980 Census, Thomas Brothers Grid Maps
Pollution (TSP)	Total Suspended Particulates (negative)	μ_g/m^3, Annual Geometric Average	California Air Resources Board
Pollution (O_3)	Ozone Concentrations (negative)	PPHM, Annual Arithmetic Average of daily maximum	

listing all homes sold in Los Angeles County during this period. The number of entries was unmanageably large (approximately 50,000 observations) so the data set was reduced in size using a random number matching system, Thus, for the basic econometric work the number of randomly chosen observations was 5,921. Subsequent empirical analysis examined a region extended to include the other South Coast Air Basin counties. Again, a sample of approximately 5,000 observations was used.

In addition to the immediate characteristics of a home, other variables which could significantly affect its sale price are those that reflect the condition of the neighborhood and community in which it is located. Such variables include, school quality, ethnic composition, proximity to employment, distance to the beach, and measures of local population density. In order to capture these impacts and to isolate the independent influence of location vis-a-vis ozone differences, these variables were included in the econometric modeling.

The data base assembled for the housing value study is appropriate to test the hypothesis outlined above for two reasons. First, the housing characteristic data is extremely detailed at the household level of aggregation and extensive in that a relatively large number of observations are considered. Second, a variety of neighborhood and community variables which enable the isolation of ozone variation on housing values have been included.

D.3 Empirical Results - Single Equation Model for Los Angeles County

The underlying structure of the initial hypothesis test is a single equation empirical model which attempts to explain the variation in sale prices of homes located in Los Angeles County for the years 1978, 1979.[3] The estimated coefficients of these hedonic equations specify the effect a change in a particular independent variable has on sale price. In reference to the ozone variable, this procedure allows one to focus on its significance while separating out the influence of other extraneous variables. Therefore, this analysis yields two outputs concerning the relationship of ozone differentials to housing price. The relative significance of location variations is determined and the estimated coefficient pertaining to location implicitly measures its monetary value at the margin.

The estimated hedonic price gradient that best fits the data is presented in Table 5.12. A number of aspects of the equation are worth noting. First, both ozone and suspended particulate concentrations are included in the equation. The particulate measure is used as a proxy for the aesthetic component of air quality while ozone concentrations implicitly measure the health effects. Second, the nonlinear specification utilized is a significant improvement over linear forms. As Rosen (1974) pointed out, this is to be expected since consumers cannot always arbitrage by dividing and repackaging bundles of housing attributes. Third, approximately .82 of the variation in home sale price is explained by the variation in the independent variable set. Fourth, with the exception of the time to work and percent old variables, all coefficients are

TABLE 5.12

ESTIMATED HEDONIC EQUATION (SEMI-LOG) FOR LOS ANGELES COUNTY.
DEPENDENT VARIABLE = in (HOME SALE PRICE IN HUNDREDS OF 1978 DOLLARS

Variables	Coefficient	t-statistic
Site Specific Characteristics:		
Sales Date	.1664 * 10^{-1}	30.91
Age of Home	-.22998 * 10^{-2}	-12.01
Square Feet of Living Area	.3221 * 10^{-3}	42.77
Number of Bathrooms	.9720 * 10^{-1}	14.43
Number of Fireplaces	.8774 * 10^{-1}	15.61
Pool	.9977 * 10^{-1}	12.02
View	.1390	14.26
Community Characteristics:		
School Quality	.1674 * 10^{-3}	2.28
Population Density	-.1192 * 10^{-4}	-7.75
% White	.8583 * 10^{-2}	46.41
% Greater THan 62 Years Old	-.2182 * 10^{-3}	-.36
Pollution (TSP)	-.1148 * 10^{-1}	-32.67
Pollution (Ozone)	.1011 * 10^{-1}	7.30
Location Characteristics:		
Time to Employment	-.5349 * 10^{-3}	-.53
Distance to the Beach	-.1475 * 10^{-1}	-15.84
Constant	6.4380	147.45
R-Squared	.82	
Number of Observations	5921	

significantly different from zero at the one percent level and possess the expected relationship to home sale price. However, the most noteworthy aspect of the hedonic equation is that the ozone variable is positively related to home sale price.

The explanation for this unexpected result is found through examination of the correlation coefficient matrix. This indicates that ozone concentrations and distance to beach are highly collinear, with a simple correlation coefficient of .896. Whereas a high simple correlation coefficient warrants concern, it is not sufficient to claim collinearity as the cause of the problem with the ozone variable. However, the degree of harmful collinearity can be somewhat determined through a rule of thumb suggested by Klein. This rule indicates that multicollinearity would be regarded as a problem only if $R^2_{HSP} < R^2_{x_i}$ where R^2_{HSP} is the multiple correlation of home sale price versus the independent variable set and $R^2_{x_i}$ is the multiple correlation between ozone and the rest of the independent variables. In this case the Klein criterion is satisfied since $R^2_{HSP} = .82$ and $R^2_{x_i} = .83$. Thus, the degree of collinearity in the data is indeed harmful, preventing the estimation of an accurate relationship between ozone and home sale price.

In Los Angeles County the collinearity is especially problematical for the variables distance to beach and ozone for two reasons. First, the prevailing daytime wind patterns are essentially perpendicular to the beach meaning as one moves inland air pollution in general increases.[4] Secondly, the chemical reaction which causes ozone formation requires time and hence distance from the original discharge locations. Thus, the prevailing wind patterns plus the large stock of upwind pollutants yield significant increases in ozone concentrations as one moves inland from the beach areas. Each variable is then measuring exactly the same impact upon home sale price.

Finally, it should be noted that the collinearity problem in Los Angeles County is stable across both functional form and randomly drawn samples. To justify the former statement a variety of functional forms, which allow for variation in both dependent and independent variables, were estimated. Further, a number of random samples were drawn of varying size, including the limiting case of including all observations. In no instance was the collinearity between distance to beach and ozone concentrations broken. Given then that the collinearity could not be reduced through functional form or random sampling, a variety of other approaches were attempted. These are the subject of the next section.

D.4 Alternative Solutions *to* Multicollinearity

Given the multicollinearity between variables and the associated spurious ozone result as described above, the next task was to search for a reasonable solution. The econometrics literature contains a number of possibilities including: (i) dropping variables; (ii) using extraneous estimates; (iii) ridge regression; (iv) nonrandom sampling; (v) altering the model specification; (vi) increasing the spatial variation by increasing the study area; and, (vii) principal components. Each of these was considered. Most were eliminated either on theoretical grounds, lack of supporting information or statistical insignificance. Only the last two options provided any satisfactory solution.

Consider first the dropping variables solution. The problem with multicollinearity is that there is insufficient information in the sample to permit accurate estimation of the individual parameters. By dropping an independent variable (distance to beach in this case) one can derive estimates of the other parameters. However, these estimates are biased, even though they have smaller mean square errors than the original estimates. But it is precisely the unbiasedness that is desired in this case since the estimates are used to calculate the rent differential for comparison to the survey results. In this instance if distance to beach is excluded from the estimation, then the coefficient on ozone possesses the correct negative relationship to home sale price and is significant at the one percent level. However, the estimate is biased and includes the impact of both distance to beach and ozone concentrations. With no *a priori* method for determining the magnitude of the bias, dropping variables does not meet the criterion of reasonableness.

The use of extraneous estimates represents a means to control the collinearity by (i) using an estimate of the impact of distance to beach on home sale price taken from an exogenous estimation; and (ii) correcting home sale price for this impact and then estimating the independent influences of ozone on the dependent variable. However, to our knowledge, there exists no such truly extraneous estimate of distance to beach on home sale price. Furthermore, this method is somewhat questionable on the basis that the extraneous estimate may indeed be "extraneous" and not measure precisely what was intended (Meyer and Kuh, 1957).

The next solution, ridge regression (as used to solve collinearity) is a purely statistical solution without much basis in economic theory. Further, interpretability is oftentimes a problem with the parameter estimates from this procedure. Thus, this solution was not considered in detail.

The nonrandom sampling solution constitutes an attempt to break the collinearity by choosing the sample so as to control for one of the problem variables. Two separate nonrandom sampling procedures were tried in this study. First, sampling was completed along lines parallel (constant distance) to the beach. This was an attempt to control for beach distance yet allowing variation in the other explanatory variables. The primary problem of this procedure is control of beach distance effectively controlled the variation in other variables. The distance to beach variable is insignificant as is expected since it is being controlled. However, this does not solve the problem of the ozone variable since it too is not significantly different from zero even at the ten percent level. This is also to be expected given the degree of collinearity between the two explanatory variables; that is, controlling for one effectively controls the other.

In response to this problem, the second nonrandom procedure was conducted along lines possessing an approximate forty-five degree angle relationship to both the beach and the predominant wind direction. This constituted an intermediate sampling method by controlling somewhat for beach access yet allowing some variation. The results of this exercise were somewhat more promising in that ozone concentrations possess the correct relationship (negative, but not significantly different from zero) to home sale price. However this approach is beset by other limitations, which are also of concern in the first nonrandom sampling procedure. These limitations include the following.

First, there is insufficient variation in other variables to permit accurate estimation; that is, the sampling procedure reduces the inherent variation in the other variables. Second, there is induced multicollinearity as a result of this insufficient variation. Thus, whereas the simple correlation between ozone and distance to beach is reduced, the simple correlations between ozone and population density, ozone and TSP, TSP and population density, ozone and percent greater than 62 years old and others demonstrate marked increases. The total multicollinearity is therefore not reduced due to the non-random sampling. Third, without a specific sampling plan generalization outside the sample may not be justifiable.

In conclusion, the non-random sample experiments conducted were not completely successful. However, some hope remains, especially in light of the results concerning the second approach. It seems that a non-random sampling method could be devised that counters the arguments presented above. Thus, this solution is not without some merit and may warrant further investigation.

The failure of the previous experiments led these researchers to question the basic model specification. That is, rather than posit a single equation model, a simultaneous equation system was examined. The basis for this model is that ozone is a produced pollutant and is dependent upon its precursors (reactive hydrocarbons, oxides of nitrogen) plus some reaction time. If reaction tine is functionally dependent upon distance travelled then this would explain the high correlation between ozone and distance to beach in Los Angeles County. Note that distance to beach is essentially distance travelled (or reaction time) since the predominate daytime wind direction is perpendicular to the beach.

The structural equations of this simultaneous system can be formally stated as:

$$\text{HSP} = \beta_o + \beta_1 \cdot (\text{BD}) + \beta_2 \cdot (\text{O}_3) + \beta_3 (\text{NO}_x) + \sum_{i=1}^{n} \gamma_i X_i \qquad (5.19)$$

$$\text{Ozone} = \alpha_o + \alpha_1 \cdot (\text{BD}) + \alpha_2 \cdot (\text{NO}_x) + \alpha_3 \cdot (\text{HC}) \qquad (5.20)$$

where

HSP = home sale price

BD = distance to nearest beach

O_3 = ozone concentrations

NO_x = oxide of nitrogen concentrations

HC = reactive hydrocarbons concentrations

X_i = the rest of the independent variable set usually associated with a hedonic housing equation

$\alpha_i, \beta_i, \gamma_i$ = parameters to be estimated

The first equation is the standard hedonic housing equation. The second equation is the production relationship. Each equation could be specified as above (linear) or some other better fitting functional form. In this model the endogenous variables are home sale price and ozone concentrations. All other variables are exogenously determined. In addition, under the assumption that reactive hydrocarbons are not perceived directly by households (reactive hydrocarbons are omitted from the first equation) then the model is identified; that is, the rank condition for identification is satisfied.

Substituting the second equation into the first the model can be rewritten as:

$$HSP = (\beta_o+\beta_2\alpha_o) + (\beta_1+\beta_2\alpha_1) \cdot BD + (\beta_3+\beta_2\alpha_2) \cdot NO_x + \beta_2\alpha_3 HC + \sum_{i=1}^{N} \gamma_i X_i \qquad (5.21)$$

or where

$$\lambda_o = \beta_o + \beta_2\alpha_o$$

$$\lambda_1 = \beta_1 + \beta_2\alpha_1$$

$$\lambda_2 = \beta_3 + \beta_2\alpha_2$$

$$\lambda_3 = \beta_2\alpha_3$$

$$HSP = \lambda_o + \lambda_1 \cdot BD + \lambda_2 \cdot NO_x + \lambda_3 \cdot HC + \sum_{i=1}^{N} \gamma_i x_i \qquad (5.22)$$

Equations (5.20) and (5.22) are the reduced-form equations. The parameters of the model $(\alpha_i, \lambda_i, \gamma_i)$, can then be estimated using indirect least squares. In this method the reduced-form equations are estimated using ordinary least squares and then the structural equation parameters are obtained from the relationships specified above. Thus,

$$\beta_o = \lambda_o - (\lambda_3/\alpha_3) \cdot \alpha_o$$

$$\beta_1 = \lambda_1 - (\lambda_3/\alpha_3) \cdot \alpha_1$$

$$\beta_2 = \lambda_3/\alpha_3$$

$$\beta_3 = \lambda_2 - (\lambda_3/\alpha_3) \cdot \alpha_2$$

No transformation is required for the α_i and the γ_i.

Considering the ozone equation, estimation was completed as follows. Data at each of the air quality monitoring stations was utilized in the estimation. Ozone, NO_x and HC were specified as annual arithmetic averages of the daily maximum values. Distance to beach was measured in miles. The estimated equation in linear form is presented in Table 5.13.[5] As indicated the only significant variable is distance to beach. This implies that the proposed physical model is somewhat deficient.

Further investigation of the physical relationship between ozone and its constituent pollutants revealed that ozone peaks generally occurred downwind from the hydrocarbon and oxides of nitrogen peaks. Therefore, rather than use HC, NO_x and O_3 measurements from the same monitoring station, ozone concentrations at each station were related to the corresponding farthest upwind station. These results are presented in Table 5.14. Again, distance to beach is the only significant variable indicating rejection of the physical model of ozone formation. In this case the failure of HC and NO_x to appear as significant variables may be traced to the lack of sufficient variation in the upwind data on an annual average basis. A more reasonable approach would employ daily pollution data.

These experiments indicate that the proposed physical model is either incorrectly specified or the data is insufficient for the task. Without an accurate physical model the simultaneous equation approach as developed here lacks sufficient justification. Thus, as a solution to the multicollinearity problem the simultaneous equation method was abandoned. This does not imply that the methodology is inherently incorrect but rather that until further refinements are made the model holds little promise.

This section examined a variety of solutions to multicollinearity in the Los Angeles data set. Essentially, each proposed solution was unsuccessful. In the next two sections empirical results are presented for two solutions which do yield the expected relationship between ozone concentrations and home sale price.

D.5 Empirical Results - Single Equation Model, South Coast Air Basin

As is detailed above, there exists severe collinearity between

TABLE 5.13
ESTIMATED OZONE EQUATION (LINEAR) FOR LOS ANGELES COUNTY. DEPENDENT VARIABLE = OZONE CONCENTRATIONS IN PARTS PER MILLION

Variables	Coefficient	t-statistic
Beach Distance	.00426	3.10
Oxides of Nitrogen	.5233	1.05
Hydrocarbons	-.00464	-.834
Constant	-.0049	-.067
R-Squared	.60	
Residual Sum Squares	.0115	
Number of Observations	14	

TABLE 5.14

ESTIMATED OZONE EQUATION (LINEAR) FOR LOS ANGELES COUNTY UPWIND DATA. DEPENDENT VARIABLE = OZONE CONCENTRATIONS IN PARTS PER MILLION

Variables	Coefficient	t-statistic
Beach Distance	.0056	4.22
Oxides of Nitrogen	.962	.867
Hydrocarbons	.0021	.109
Constant	-.102	-.853
R-Squared	.55	
Residual Sum Squares	.0124	
Number of Observations	14	

ozone and distance to beach within Los Angeles County. However, in the areas adjacent to Los Angeles County the collinearity between these variables is much less apparent. Therefore, it was decided to increase the spatial variation in the data set through the addition of data from Orange, Riverside and San Bernardino Counties. The data addition was restricted to those areas of each county which borders Los Angeles County on the premise that data from long distances would constitute a separate housing market. The housing data was obtained from the SREA Market Data Center while the associated neighborhood and community data were obtained from the sources outlined in Table 5.1.

The data from the surrounding counties were pooled with the original Los Angeles County data. The new data set had approximately 68,400 observations. The relevant county breakdown was Los Angeles with 50,432, Orange with 12,117, Riverside with 1,452 and San Bernardino with 4,405. Due to this large size the data set was reduced to 4,951 observations using a random number matching system. In order to account for any variation in housing markets across county boundaries a set of zero-one variables for county location were constructed and added to the data set. Before proceeding to a discussion of the empirical results based on the new sample it should be noted that the additional data reduced the simple correlation coefficients between ozone and beach from .896 to approximately .66.

In addition to the data which increased the spatial variation, data which more closely approximates the aesthetic aspect of air quality became available. That is, a measure of actual visibility, or its reciprocal, light extinction was generated by a simultaneous California Air Resources Board project. The variable visibility is measured as median miles and was calculated for grid squares roughly four miles square for the study area.[6] This variable was entered into the data set as another explanatory or independent variable.

Given the data as outlined above, a single equation hedonic housing model was estimated. A particular example is presented in Table 5.15. Note that the Riverside County zero-one variable is the excluded variable so that the zero-one variables for the other counties are interpreted as deviations from Riverside County as depicted by the constant term. As is illustrated, the estimated equation seems to perform quite well on a number of counts. First, approximately 80 percent of the variation in home sale price is explained by the independent data set. Second, with few exceptions, the estimated coefficients possess the expected relationship to home sale price and are significant at the one percent level. Two exceptions are ozone and school quality. However, these variables are significantly different from zero at the ten percent level under the presumption of a priori information; that is, the sign of the variable is known in advance. Therefore, the only variable which is not significantly different from zero at the ten percent level is time to work. However, this is not totally unexpected since this variable is essentially constant, demonstrating a small variance around its mean. The indication is that most people travel about the same time to work. Thus, its insignificance is not particularly troublesome.