Market Transactions and Hypothetical Demand Data: A Comparative Study

MARK DICKIE, ANN FISHER, and SHELBY GERKING*

Empirical demand studies have been based on data from (a) actual market transactions or (b) hypothetical questions. Many social scientists are skeptical of the accuracy of responses to hypothetical questions, yet few studies assess the quality of this type of data. This article directly compares the demand relations obtained from actual market transactions and hypothetical survey responses using primary field data and limited dependent variable regression analysis. Using a log-likelihood ratio test, the null hypothesis that the two demand relations are statistically identical cannot be rejected at the 1% level of significance.

KEY WORDS: Accuracy of survey data; Demand data collection methods; Hypothetical questions; Tobit.

1. INTRODUCTION

Empirical demand studies have been based on data obtained from one of two sources: (a) actual market transactions or (b) hypothetical questions. Price and quantity data from actual market transactions have been most widely used by economists because of their accessibility from published and computerized sources as well as their obvious demand-revealing properties. Well-known gaps and problems with these data, however, have stimulated interest in using hypothetical questions to generate the required information. For example, for private goods that change hands infrequently, such as houses, actual market transaction data are limited. As a consequence, researchers have resorted to surveys, such as the Annual Housing Survey and the Census of Population, that ask how much money the respondent could get for his house if it were sold on today's market. In addition, for environmental goods, such as clean air or visibility, that are not separately traded in markets, actual transactions data do not exist. This situation has inspired the development of the contingent valuation method in which a survey respondent is directly asked how much money he would be willing to give up to enjoy a particular, but hypothetical, environmental improvement.

Hypothetical demand data, however, are subject to sevral sources of potential bias. An important type of bias, which might be termed payment bias, arises because hypothetical situations may not provide sufficient incentive for respondents to reveal their true preferences. This possibility alone is enough to arouse skepticism of results based on this type of data. Despite this skepticism, widespread use of hypothetical data continues. Yet only a few studies, including those by Brookshire, Thayer, Schulze, and d'Arge (1982), Kain and Quigley (1972), and Kish and Lansing (1954), have attempted to evaluate the quality of hypothetical data, and virtually no studies have directly compared demand relations based on hypothetical data with those obtained from actual market transactions data. Such comparisons, which are easiest to make in the context of a private good, would be of immediate value for assessing the relative usefulness of hypothetical data as well as for indicating ways to improve demand-revealing data collection methods generally.

This article provides a comparison of the demand equations for a private good estimated using actual market transactions data and hypothetical responses. These two types of data, which measure revealed and expressed preferences, respectively, were collected by means of a door-to-door, in-person survey. Thus the demand equation comparison not only tests for the payment bias described previously; it also represents a contribution to the broader literature on the accuracy or validity of survey data (e.g., see Dillman 1978; Sudman 1976). Since this comparison considers only the demand relation for one private good in one community, the results should be viewed only as suggestive. Nevertheless, these results are of interest because they illustrate the extent to which a demand equation based on actual market transactions data differs from one based on hypothetical responses.

2. RESEARCH DESIGN AND METHODOLOGY

Since this study focuses on the extent of payment bias in hypothetical demand data, the research was designed to control for other types of bias (considered at length by Cummings, Brookshire, and Schulze 1986) associated with hypothetical data. One potential source of bias in hypothetical response data, which is particularly relevant when dealing with public goods, is strategic misrepresentation of preferences. For instance, a respondent who has a strong desire for an environmental good may report more than his true willingness to pay if he feels that (a) his bid will influence the good's provision and (b) he will never actually have to pay this amount (either because the cost per person will be lower when spread across all taxpayers or because the payment per person will be based on the average response, which he expects to be lower than his own). Alternatively, a respondent may underreport his willingness to pay if he believes that others will reveal their true preferences, the good will be made available, and he cannot be excluded from consuming it even though he pays only the amount of his bid.

* Mark Dickie is Assistant Professor, School of Social Sciences, University of Texas, Dallas, TX 75083. Ann Fisher is Senior Economist, U.S. Environmental Protection Agency, Washington, DC 20460. Shelby Gerking is Professor, Department of Economics, University of Wyoming, Laramie, WY 82071. This research was supported by the U.S. Environmental Protection Agency under Cooperative Agreements CR-811077-01-0 and CR-812054-01-1. The results reported, however, do not necessarily reflect the views of that organization. The authors thank Donald Anderson, Scott Atkinson, Tayler Bingham, David Brookshire, Don Coursey, Jon Harford, William Schulze, V. Kerry Smith, Joseph B. Kadane, an associate editor, and two anonymous referees for valuable insights and suggestions and George Cutts, John Hovis, Chari Pepin, and Karen Radosevich for excellent research assistance.

Journal of the American Statistical Association
March 1987, Vol. 82, No. 397, Applications

Other biases may result if the individual is unfamiliar with the commodity or uninformed about relevant market conditions. For example, the first of these could occur if an individual were asked to value the hypothetical removal of toxic wastes from a dump site when no information is available concerning either the materials or the hazards present. The second situation might arise if an individual who has not recently paid attention to the market for residential property is asked to assign a rental or sales value to his home. In a related vein, answering questions about commodities that are intangible, unfamiliar, or complex may require time for preference research before an accurate assessment of value can be made. Other sources of bias include vehicle bias, where the method of payment may influence the results, and starting point bias, where an initial price suggested by the interviewer may influence the final value reported by the respondent.

In this study, these biases are controlled by minimizing the likelihood that they would occur or at least making this likelihood roughly equal in both the actual and hypothetical components. A private good, fresh strawberries, was chosen to minimize the incentives for strategic bias normally associated with public goods. Respondents in both components were asked to state the quantity of strawberries desired at given prices. Posing the question the other way around (i.e., asking for values at given quantities) would have immediately introduced the possibility for strategic behavior by respondents who participated in the actual market transactions portion. Because strawberries are a simple, tangible, and familiar commodity, and because the data were collected from individuals who regularly shopped for groceries, any biases associated with lack of familiarity with the commodity or the market should be minimal, as should the time and information needed for preference research. Finally, the method of payment presented to the respondents was identical in both the actual market transaction and the hypothetical response portions of the study.

Besides allowing greater focus on the single issue of payment bias, fresh strawberries were selected as the commodity for analysis for three additional reasons. First, since strawberries are nondurable, the demand for them can be viewed in a static framework. Second, because strawberries are a relatively inexpensive commodity, a large enough quantity to implement this study could be purchased on a limited research budget. In addition, since strawberries account for a small share of the household budget, there is no need to analyze income effects when prices change. Third, fresh strawberries are seasonal and normally exhibit price fluctuations even on a week-to-week basis. This characteristic makes it easier to estimate demand relations over a range of prices.

Primary field data were collected during the summer of 1984 in Laramie, Wyoming to generate information for both the actual market transactions and hypothetical response portions of the study. This community was chosen primarily on the basis of cost and convenience. Laramie is a town of approximately 25,000 residents and is the location of the University of Wyoming. The following multistage procedure, adapted from Sudman (1976), was used to obtain a sample of regular grocery shoppers from Laramie households. The City Planning Office has demarcated 19 divisions in Laramie, and these are the smallest neighborhood units for which 1980 census data are available. From these census data, the number of households and their average income in each division were determined, and approximately one-third of the city population was assigned to each of a low-, middle-, and high-income stratum. Then six divisions were randomly selected (two in each income stratum), with probability proportional to their population. These six primary sampling units (PSU's) were partitioned into clusters (a street or pair of adjacent streets containing six to eight city blocks) of approximately 40 households each so that any given cluster could accommodate 12 sample points. Next, two clusters were randomly selected from each of the six PSU's, with one cluster assigned to actual market transactions and the other to hypothetical response surveys. After a random start in each cluster, every third house was chosen until 12 sample points had been obtained from all six clusters. If a regular grocery shopper was unavailable at one of the chosen households, the survey team returned to the house at a later time. If this second attempt to contact a regular grocery shopper failed, or in the rare case in which this person refused to participate in the study, one of the two houses next door was chosen.

Thus 72 households were selected for inclusion in the actual market transactions portion and 72 additional households were drawn in a parallel manner for the hypothetical response portion. This sample size was selected in light of the range of strawberry prices used in the survey, the variation expected in the household income data, and the number of explanatory variables expected to be used in the statistical demand equations. Three survey teams of two persons each collected the data for both portions of the study over a 4-day period in July 1984. These data were collected during the late afternoon and early evening hours by means of a questionnaire administered in door-to-door, in-person interviews. The questionnaire had been pretested on five Laramie households to improve its design and to give the survey teams practice with administration procedures. With one exception, each team completed interviews with two clusters drawn from different income strata in both the actual and hypothetical portions of the study. The exception arose in the hypothetical portion when one member of the third survey team became ill and was unable to conduct the 12 assigned interviews in the low income stratum. Rather than delaying the survey for an unknown length of time or substituting an untrained interviewer, six of these interviews were conducted by the first team and six by the second.

To implement the actual market transactions portion, initial contact with the household identified the individual who regularly shopped for groceries. The interviewers gave a brief, standardized introduction, displayed the available strawberries, and then said, "Each pint is selling today for the price of $ ______ . How many pints would you like to purchase?" Six prices ($.60, $.80, $1.00, $1.20, $1.40, and \$1.60) were inserted in this statement, with 2 households in each cluster (for a total of 12 households) randomly assigned to each price. Fresh strawberry prices charged by the four major Laramie grocery stores ranged from \$.89 to \$1.29 per pint during July 1984. Thus the prices quoted to respondents more than spanned this range.

If the respondent desired to purchase at least 1 pint, an exchange of strawberries and money was completed. Immediately thereafter, the respondent was told that the purpose of the visit really was to collect market research information. The respondent then had his money refunded and was allowed to keep the strawberries in return for supplying the survey team with information needed to complete the questionnaire. (Copies of all questionnaires, the raw data, and tables and figures presenting supplementary results are available from Shelby Gerking on request.) On the other hand, if no strawberries were purchased, the interviewer offered them to the respondent at no charge in return for help in completing the questionnaire. The first items on the questionnaire called for the survey team to record the price and quantity data obtained. Other variables measured included the following: (a) number of household members (NUMBER), (b) total monthly household income (INCOME), (c) hours since last full meal was eaten (ATE), (d) days since household last shopped for groceries (SHOP), (e) respondent's years of age (AGE), (f) whether respondent is white (WHITE), and (g) respondent's years of formal schooling (SCHOOL).

The hypothetical response data were collected using the same procedure, except for two differences. First, instead of informing the respondent of the price at which strawberries would be sold, the survey team stated that they were gathering information for market research purposes. Second, after this introduction, the available strawberries were displayed, and the respondent was told, "Suppose each pint is selling today for \$ ______ . How many pints would you purchase?" The prices inserted in the preceding statement were the same as those used in the actual market transactions portion, and an identical procedure was used to match the prices to households. After obtaining the answer to the question, the respondent was offered strawberries at no charge and the interview commenced.

3. AN EMPIRICAL COMPARISON

The basic demand relation estimated using both the actual market transactions and hypothetical response data is shown in Equation (1).

$$Q_i = f_j(P_i, \text{INCOME}_i, \text{NUMBER}_i, \text{ATE}_i, \text{AGE}_i, \text{SHOP}_i, \text{WHITE}_i, \text{SCHOOL}_i), \quad (1)$$

where Q_i denotes the number of pints of strawberries that would have been purchased by the ith respondent at price P_i ($i = 1, \ldots, 144$). The observations on the variables in Equation (1) are ordered such that $i = 1, \ldots, 72$ corresponds to the actual market transactions data and $i = 73, \ldots, 144$ corresponds to the hypothetical response data. The function subscript j can take on two values depending on whether the actual market transactions data or the hypothetical response data are considered. Thus $j = 1$ if $i = 1, \ldots, 72$ and $j = 2$ if $i = 73, \ldots, 144$.

Table 1 shows that the dependent variable Q_i was 0 for 58% of the observations in the actual market transactions portion and for 47% of the observations in the hypothetical response portion. As a consequence, Equation (1) was estimated in a tobit framework [see Tobin (1958) and Judge, Griffiths, Hill, Lütkepohl, and Lee (1985) for details]. In addition, Table 1 indicates that one respondent in the hypothetical portion stated that 10 pints of strawberries would be purchased (at $P = \$.60$). This observation may appear to be an outlier and thus a candidate for either trimming or exclusion from the sample. Yet later in the interview, the respondent stated that half of this comparatively large quantity of strawberries would be frozen or canned. The influence of this observation was given special attention in interpreting the results presented here. For instance, trimming the value $Q = 10$ to $Q = 5$ or excluding the observation from the data set produces only minor changes in the tobit regression results. Consequently, the results shown include this observation without adjustment. A further numerical comparison, outside the framework of Equation (1) (discussed in connection with Table 4), however, is influenced by the treatment of this observation.

Tobit estimates of four versions of Equation (1) are presented in Table 2. These estimates are used to examine the differences between the actual market transactions and the hypothetical response demand functions. In the first and second columns of Table 2, separate regressions are presented for each type of data collected in the survey. A regression based on pooling the two types of observations is shown in the third column. The fourth column shows another pooled regression in which a dummy variable (MARKET) together with interaction variables between MARKET and all other explanatory variables are added to the covariates included in the Column 3 regression. MARKET equals unity if $i = 1, \ldots, 72$ and is zero otherwise. All four equations include dummy variables for two of the three survey teams (TEAM1 and TEAM2) as well as dummy variables for two of the three income strata (HIGH and MEDIUM). As a consequence, team effects are adjusted for stratum effects and vice versa. Attempts also were made to include the interaction variable TEAM1 × HIGH. (Because each team visited households in two strata, only one such interaction could be included without forcing an exact linear dependence between columns of the design matrix.) Convergence problems developed,

Table 1. Frequency Distribution for Q_i in the Actual and Hypothetical Samples

Q_i	Actual	Hypothetical
0	42	34
1	16	18
2	12	13
3	2	4
4	0	1
5	0	1
10	0	1
	72	72

Table 2. Comparison of Actual Market Transactions and Hypothetical Response Demand Equations

Independent variable	Unnormalized tobit regression coefficients*			
	Actual	*Hypothetical*	*Pooled*	*Pooled*
CONSTANT	4.738	7.623	6.612	7.064
	(3.430)	(3.283)	(4.342)	(3.778)
P	−2.247	−3.015	−2.446	−2.779
	(−4.235)	(−3.314)	(−4.360)	(−3.800)
INCOME	.0002933	.0004294	.000323	.0003921
	(2.041)	(1.431)	(1.993)	(1.620)
NUMBER	.3884	−.4604	.1319	−.4268
	(3.635)	(−1.396)	(.9086)	(−1.608)
ATE	−.01489	−.2588	−.1653	−.2505
	(−.1178)	(−1.542)	(−1.392)	(−1.850)
ATE2	.004758	.01149	.01034	.01105
	(.9671)	(1.504)	(2.036)	(1.794)
AGE	−.01729	−.03577	−.02218	−.03196
	(−1.441)	(−1.8647)	(−1.742)	(−2.062)
MALE	−.3077	.8274	.01375	.7463
	(−.8931)	(1.157)	(.03488)	(1.294)
SHOP	.006789	−.07131	−.04393	−.06452
	(.2236)	(−1.347)	(−1.380)	(−1.503)
WHITE	−.3483	−1.300	−.9186	−1.329
	(−.6431)	(−1.513)	(−1.632)	(−1.927)
SCHOOL	−.1946	−.1159	−.2083	−.09809
	(−2.813)	(−.9952)	(−2.703)	(−1.045)
TEAM1	.8537	1.311	.8666	1.209
	(1.804)	(1.363)	(1.700)	(1.557)
TEAM2	−1.209	1.042	−.01944	1.063
	(−2.378)	(.876)	(−.03423)	(1.109)
HIGH	.1238	.5146	.01848	.571
	(.2131)	(.6089)	(.03256)	(.837)
MEDIUM	−.0308	.8421	.1762	.919
	(−.0666)	(.7283)	(.3100)	(.987)
MKT				−1.284
				(−.452)
MKTPRICE				−.117
				(−.107)
MKTINCOME				−.0000543
				(−.1618)
MKTNUMBER				.9008
				(2.852)
MKTATE				.228
				(.909)
MKTATE2				−.004433
				(−.4335)
MKTAGE				.008734
				(.3459)
MKTMALE				−1.136
				(−1.398)
MKTSHOP				.07363
				(1.107)
MKTWHITE				.9904
				(.8692)
MKTSCHOOL				−.1530
				(−1.047)
MKTTEAM1				−.2655
				(−.2439)
MKTTEAM2				−2.835
				(−2.203)
MKTHIGH				−.2236
				(−.1881)
MKTMEDIUM				−.9388
				(−.7759)
Standard Error	.9522	1.940	1.809	1.575
Log of likelihood	−57.529	−98.932	−177.167	−162.130
Predicted probability of $Q_i > 0$	.4026	.5205	.4652	.4400
Observed frequency of $Q_i > 0$	.4167	.5278	.4722	.4722
Number of iterations	5	5	5	5
Number of observations	72	72	144	144

* *t* statistics are given in parentheses.

however, with the maximum likelihood procedure used in the estimation whenever this variable was included in the actual and pooled regressions. Finally, equations with dummy variables for PSU were estimated but are not presented since the coefficients of PSU seldom were significantly different from 0 at the 5% level. Thus, similar to the findings in larger scale surveys by Kish and Frankel (1970), the regression coefficients in Table 2 appear to have quite small design effects.

Table 2 reports estimates of the unnormalized coefficients. These values, which are estimates of the original coefficients in the regression model, are the normalized coefficients multiplied by the standard error of the estimate. In the equation estimated by using only the actual market transactions data (see Column 1), the coefficients of the key variables P and INCOME have the expected signs (negative and positive, respectively) and are significantly different from 0 at the $2\frac{1}{2}\%$ level, using a one-tailed test. The performance of the remaining explanatory variables listed in Equation (1), however, is not as strong. Less formally educated respondents in larger households tended to purchase larger amounts of strawberries; the coefficients of AGE, ATE, ATE2, MALE, SHOP, WHITE, HIGH, and MEDIUM, however, are not different from 0 at conventional significance levels. Finally, even though survey teams were trained to conduct interviews in a standardized manner, enumerator effects appear to be present. The coefficient of TEAM1 is positive and significant at the 10% level, and the coefficient of TEAM2 is negative and significant at the 5% level. These results might have been anticipated, since people seem to differ greatly in their natural abilities in salesmanship. Future investigators would be well advised to train enumerators extensively and perhaps, in addition, to send more than one team to the same houses at different times.

In the fitted tobit demand equation for the hypothetical response data (see Column 2 of Table 2) P and INCOME enter with negative and positive coefficients, respectively. The coefficient of P, but not of INCOME, is significantly different from 0 at the 5% level, using a one-tailed test. The negative coefficient of AGE also is significant using the same test procedure, and the coefficients of the remaining variables are not significant at conventional levels. An important difference between the actual market transactions and hypothetical response equations is that in the latter t statistics of TEAM1 and TEAM2 are small. This outcome is not surprising, since the actual market transactions data were collected during the first 2 days of the 4-day interview period and the hypothetical response data were collected during the last 2 days. Increased familiarity with interview procedures may have led to the smaller enumerator effects found in the hypothetical response data.

The third and fourth columns of Table 2 provide a basis for testing the null hypothesis of equality between the coefficients of the actual market transactions and hypothetical response demand equations. Since both equations are estimated using the tobit procedure after pooling the two types of data, the test examines the performance of the MARKET dummy variable (a constant term shifter) and the interactions of MARKET with all other covariates (the slope shifters). Except for the interactions between MARKET and TEAM2 and MARKET and NUMBER, none of the coefficients of these variables are significantly different from 0, even at the 10% level using a two-tailed test. A likelihood ratio test was made for the joint significance of the MARKET dummy variable and all interaction variables. This test fails to reject the null hypothesis of structurally identical actual and hypothetical demand equations at the 1% level. Notice that this outcome is obtained even though the test employs an underestimate of the error variance. The price by market by team interaction would be a more correct basis for the error mean square because a resampling of teams could cause slopes to shift.

The information obtained from this statistical test is augmented by comparing the values of the dependent variable predicted by the actual market transactions and the hypothetical response demand equations. These calculations make the results presented here easier to compare with those reported in the previously cited Kish and Lansing (1954) and Kain and Quigley (1972) papers. Figure 1 graphically depicts actual market transactions and hypothetical response demand equations in P, Q space. To obtain the curves labeled A and H, the estimated equations in the first two columns of Table 2 were evaluated at the overall sample means of all covariates except for P. The same procedure was used to obtain the A' curve, except that enumerator effects significant at the 5% level were controlled by reestimating the actual market transactions demand equation after eliminating the 24 observations collected by TEAM2. Three aspects of this figure warrant further discussion. First, it illustrates the functional form imposed by the tobit model. In the A curve, for example, the predicted value of Q is a negatively sloped linear function of P on the interval $0 \leq P \leq \$.98$, and at higher prices, predicted $Q = 0$. Second, the value of P at which predicted $Q = 0$ is higher for the H curve than for the A curve. This situation reflects the greater percentage of households in the actual market transactions portion to which no strawberries would have been sold. Third, the figure shows that eliminating the actual market transactions observations collected by TEAM2, which reflected a significantly lower sales volume, brings the two demand curves closer together. In fact, A' intersects H at the point $P = .76$, $Q = 1.12$, whereas H lies above A at all points on the interval $0 \leq P \leq \$1.13$.

Table 3 presents calculations of payment bias (PB), using the demand equations illustrated in Figure 1. The values of PB presented compare the H curve with both the A and A' curves at \$.10 intervals between $P = \$.60$ and $P = \$1.40$. For example, to compare A and H, PB_i is calculated using Equation (2).

$$\begin{aligned} \text{PB}_i &= ((P_iQ_{Ai} - P_iQ_{Hi})/P_iQ_{Ai}) \\ &= ((Q_{Ai} - Q_{Hi})/Q_{Ai}), \end{aligned} \tag{2}$$

where Q_{Ai} (Q_{Hi}) denotes the predicted quantity from the A (H) demand curve and P_i denotes a price shown in the left margin of Table 3. Multiplying PB_i by 100 measures

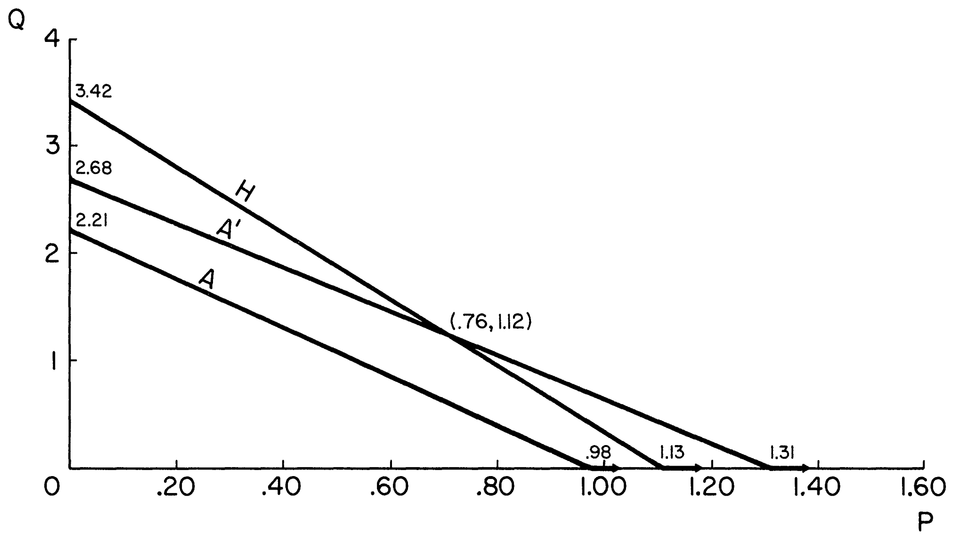

Figure 1. Actual and Hypothetical Demand Curves. The curves labeled A and H were derived by evaluating the estimated equations in the first and second columns, respectively, of Table 1 at the overall sample means of all covariates except P. Curve A′ was derived in a like manner except enumerator effects were controlled by reestimating actual market transactions demand after excluding TEAM2's observations. Excluding TEAM2 enhances considerably the similarity of the actual and hypothetical demand curves. The tobit method of handling the truncation of Q is also illustrated. When P exceeds \$.98, \$1.13, and \$1.31 for the A, H, and A′ curves, respectively, predicted Q = 0.

the percentage difference in total strawberry expenditures predicted by the A and H curves.

As shown in Table 3, there is considerable variation in values of PB_i. Of course, where the two demand curves compared both lie on the P axis, the absolute payment bias is 0, even though PB_i cannot be calculated. In addition, PB_i is small for values of P near the point of intersection of the A' and H curves. Table 3 also shows cases in which the difference in predicted total expenditures is 100% or more. The table illustrates the potential for PB_i to be large even though (a) the null hypothesis of structurally identical A and H curves was not rejected at the 1% level and (b) the significant (at the 5% level) enumerator effects associated with TEAM2 were controlled in obtaining the A' curve.

A final comparison can be drawn by examining the average expenditure for strawberries by respondents in the actual market transactions and hypothetical response portions of the study. Average expenditure is computed by adding the products of price and quantity for each respondent and then dividing by the number of respondents. Because the respondent's stated, rather than predicted, quantity is used in this calculation, the outlier observation previously noted in Table 1 exerts greater influence on the results. In particular, Table 4 shows values of $D = ((\bar{E}_A - \bar{E}_H)/\bar{E}_A) \times 100$, where $(\bar{E}_A)$ and $(\bar{E}_H)$ denote average expenditures in the actual and hypothetical portions of the study. Six values of D are presented that are classified by the treatment of (a) the outlier observation and (b) the actual market transactions data collected by TEAM2. Table 4 indicates that with the outlier and the actual TEAM2 data included, $D = -58.3\%$. After excluding the actual TEAM2 data and either trimming or excluding the outlier, however, the value of D rises substantially to values that are quite close to 0 ($D \leq -.8\%$). These latter values of D are smaller in absolute value than those found by Kain and Quigley and by Kish and Lansing. In their study of housing in St. Louis, Kain and Quigley found (a) an average absolute percentage difference of 21.2% between owner and professional appraiser estimates of value in 113 owner-occupied structures and (b) a percentage difference between the mean owner and appraiser values of 1.8%. In addition, Kish and Lansing found a 4% difference, roughly, between mean owner and mean appraiser house values, using a national probability sample of 568 home owners. This comparison with the housing studies, how-

Table 3. Percentage Differences Between Predicted Total Expenditures

	Percentage differences	
Price	A and H	A′ and H
\$1.40	a	a
1.30	a	100
1.20	a	100
1.10	b	76.9
1.00	b	36.5
.90	−351.4	15.9
.80	−153.7	3.5
.70	−96.1	−4.9
.60	−68.6	−10.9

NOTE: a = both demand equations lie on the P axis; b = actual market transaction demand equation lies on the P axis.

Table 4. Percentage Differences in Average Expenditures

	Treatment of TEAM2	
Treatment of outlier	Include all data collected	Include only hypothetical data
Left in sample	−58.3	−13.4
Trim $Q = 10$ to $Q = 5$	−5.1	−.8
Exclude observation	−4.7	−.5

ever, should not be overdrawn, because appraiser estimates of value may differ from the price received if the house were actually sold. In addition, the actual market transactions demand data (with or without the TEAM2 observations) may only approximate behavior at the grocery store.

4. CONCLUSIONS

This article has compared demand relations for fresh strawberries based on actual market transactions and hypothetical response data. The empirical analysis reveals that the null hypothesis of structurally identical demand equations obtained with these two data collection methods is not rejected at the 1% level of significance. At given prices inserted in the two demand equations, however, percentage differences in the predicted quantity of strawberries purchased can exceed 100%. A problem with the data collected is that at the 5% level one of the three interview teams sold significantly fewer pints of strawberries during the actual market transactions portion of the study. If these data, together with one possible outlier observation found in the hypothetical response portion, are set aside, then average strawberry expenditures by respondents in the two portions of the study differ by less than 1%.

The results of this study suggest that although demand equations based on actual market transactions and hypothetical response data may be similar from a statistical perspective, the latter type of data may be best utilized in aggregate form. In this situation, which characterizes measures of the average value of homes in a census tract using owner estimates or the average willingness to pay for a hypothetical environmental improvement elicited from a group of survey respondents, the payment bias from individual observations may tend to cancel out.

Further research would be useful in establishing whether the findings presented here can be extended to other circumstances, particularly those involving public goods. For example, are individuals better able to answer accurately hypothetical questions about what quantity to buy at given prices (the situation considered in this study) in comparison with questions asking for hypothetical valuations (the situation encountered in housing and environmental studies)? In addition, what is the effect on payment bias in instances in which less control can be exercised over other potential sources of hypothetical response bias? One generalization in this context would be to analyze a good of a more public character and thereby allow for the possibility of strategic bias. Other possible cases include consideration of goods with which subjects are less familiar, both in terms of the nature of the commodity and the prior valuation experience they have had with it.

[*Received July 1985. Revised April 1986.*]

REFERENCES

Brookshire, D. S., Thayer, M. A., Schulze, W. D., and d'Arge, R. C. (1982), "Valuing Public Goods: A Comparison of Survey and Hedonic Approaches," *American Economic Review,* 72, 165–177.

Cummings, R. G., Brookshire, D. S., and Schulze, W. D. (1986), *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method,* Totowa, NJ: Rowman & Allanheld Publishers.

Dillman, D. A. (1978), *Mail and Telephone Surveys: The Total Design Method,* New York: Wiley-Interscience.

Judge, G. G., Griffiths, W. E., Hill, R. C., Lütkepohl, H., and Lee, T. C. (1985), *The Theory and Practice of Econometrics* (2nd rev. ed.), New York: John Wiley.

Kain, J. F., and Quigley, J. M. (1972), "Note on Owner's Estimate of Housing Value," *Journal of the American Statistical Association,* 67, 803–807.

Kish, L., and Frankel, M. (1970), "Balanced Repeated Replications for Standard Errors," *Journal of the American Statistical Association,* 65, 1071–1094.

Kish, L., and Lansing, J. B. (1954), "Response Errors in Estimating the Value of Homes," *Journal of the American Statistical Association,* 49, 520–538.

Sudman, S. (1976), *Applied Sampling,* New York: Academic Press.

Tobin, J. (1958), "Estimation of Relationships for Limited Dependent Variables," *Econometrica,* 33, 382–394.