Nonprofit and Voluntary Sector Quarterly

The Quality and Comparability of Survey Data on Charitable Giving

Mark O. Wilhelm

The online version of this article can be found at:
http://nvs.sagepub.com/cgi/content/abstract/36/1/65

Published by:
$SAGE Publications

http://www.sagepublications.com

On behalf of:

Association for Research on Nonprofit Organizations and Voluntary Action

Additional services and information for *Nonprofit and Voluntary Sector Quarterly* can be found at:

Email Alerts: http://nvs.sagepub.com/cgi/alerts

Subscriptions: http://nvs.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations (this article cites 11 articles hosted on the
SAGE Journals Online and HighWire Press platforms):
http://nvs.sagepub.com/cgi/content/abstract/36/1/65#BIBL

The Quality and Comparability of Survey Data on Charitable Giving

Mark O. Wilhelm

Indiana University Purdue University–Indianapolis

This article examines six major household surveys of charitable giving and attempts to trace differences in estimates to underlying differences in survey methodology. The main result is that surveys that cue respondent recall by directing attention to charities' outputs and use interviewers experienced in obtaining information about dollar amounts measure larger amounts of giving. Even so, it is very difficult to estimate giving at the top of the distribution without a high-income oversample: Only one of the surveys not containing a high-income oversample produces estimates around the 90th percentile similar to that obtained with a high-income oversample.

Keywords: *charitable donations; charitable contributions; relative distributions*

Research interest in voluntary giving to charitable purposes crosses disciplinary boundaries (Berking, 1999; Komter, 1996; Lee, Piliavin, & Call, 1999; Vesterlund, 2006). The tax deductibility of charitable giving and the recently proposed Charity Aid, Recovery, and Empowerment Act are evidence of the government's long-standing and ongoing attempts to encourage charitable giving. New releases of giving statistics routinely receive substantial media attention.

With this level of research, government, and media interest it may come as a surprise that household-level survey data on charitable giving were scarce until recently. After the National Study of Philanthropy, an extensive household survey fielded in 1974, survey data on giving were not collected until Independent Sector's 1988 initiation of its biennial series of cross-sections

© 2007 Association for Research on Nonprofit Organizations and Voluntary Action

65

Giving and Volunteering in the United States. Since that time there has been a flurry of activity. The General Social Survey placed the core questions from Giving and Volunteering on its 1996 survey. In 1997, the Canadian Centre on Philanthropy took a different methodological approach to the design of giving questions and Statistics Canada fielded the new instrument as part of its labor force survey. Researchers at the University of San Francisco followed that new approach in a survey of giving in California. Because none of these surveys were producing panel data, the Center on Philanthropy at Indiana University and the University of Michigan's Survey Research Center launched a module of giving and volunteering questions in the 2001 wave of the Panel Study of Income Dynamics (PSID).

The recent availability of so many data sets is a welcome change but social scientists planning research on charitable giving must now choose which data set to use in their projects. That choice would not matter if each data set produced similar giving statistics but preliminary indications are that they do not. For instance, the percentage of households making donations is 68.5 in Giving and Volunteering (Hodgkinson & Weitzman, 1996) but 89.9 in the California survey. As another example, Schervish and Havens (1998) reported that the average donation per household is much lower in Giving and Volunteering ($696) than in the General Social Survey ($1,099). It would be important to investigate the quality and comparability of these surveys under usual circumstances but differences like these make such an investigation indispensable.

This article investigates the quality and comparability of giving data in these six surveys and attempts to trace differences among them to underlying differences in survey methodology. The research strategy has three important features. First, it examines the prevalence of missing data, a potentially important problem because giving is not a highly salient event for many respondents. Second, instead of a narrow comparison of giving averages, the article uses relative distribution methods (Handcock & Morris, 1999) to provide a visually appealing comparison of entire giving distributions. Relative distribution methods also lead to a straightforward test for significant differences. Third, by including six major household-level surveys on giving, the results can help researchers decide which of the available data sets may be best suited for their projects.

The analysis produces four results. First, surveys fielded by organizations whose primary experience is in measuring public opinion produce much more missing data on amounts given to charity than do surveys fielded by organizations with extensive experience collecting other kinds of dollar-amount information. Second, surveys that cue respondent recall by asking about methods respondents may have used to transfer money into charities (input cues such as asking about giving in response to a telephone solicitation, through payroll deductions, at a shopping center, etc.) measure higher fractions who give but also a higher frequency of small gifts. Third, surveys that cue respondent recall

by asking about the output or purpose of the charities (such as asking about giving to religious organizations, the United Way, educational organizations, etc.) more frequently measure large amounts of giving, but only when fielded by organizations experienced in collecting dollar information. Finally, it is difficult to measure giving at the very top of the distribution without a high-income oversample. The Center Panel is the only survey without such an oversample that measures giving at the top of the distribution similar to that measured by the high-income oversample 1974 National Study. Even so, the National Study indicates larger giving at the very highest percentiles.

PREVIOUS LITERATURE

The consensus concern among survey researchers is that household surveys potentially underestimate giving because of recall bias—charitable donations often are not salient events and surveys typically ask about donations made during the previous calendar year, in many cases well over a year prior to the interview. Although social desirability may lead to some overreporting of donations, most survey researchers think that recall bias presents a more serious problem (e.g., Hall, 2001). Consequently, several methodological approaches to questionnaire design have arisen to address concerns about respondent recall. Rooney, Steinberg, and Schervish (2001) evaluated these approaches by fielding them simultaneously under the direction of a single public opinion research organization and found that output cue designs measured smaller average giving than did input cue designs (see also Rooney, Mesch, Chin, & Steinberg, 2005). This is surprising because it might be expected that input cue designs would facilitate recall of smaller gifts but not necessarily larger gifts.

None of the previous research has studied exactly where in the distribution of giving different surveys produce different results. Indeed, because giving is highly skewed, a few outliers could cause large differences in sample means even if the empirical distributions are otherwise similar. Neither has previous research considered the degree of missing data nor analyzed whether the oversampling of high-income respondents produces different empirical distributions of giving.

METHOD: MISSING DATA BOUNDS AND RELATIVE DISTRIBUTIONS

The effect of missing data is examined by determining the range of their potential effects on the estimate of the median gift. At one extreme, if all the missing data on amounts are set to zero, the resulting estimate of the median would be the lower bound of the potential effects of the missing data. At the other extreme, if all the missing data are set to a value large enough to move

the corresponding respondents above the median, the resulting estimate of the median would be the upper bound of the potential missing data effects. Although it is unlikely that respondents would fail to recall such large amounts, this is not the point. The point is that although reasonable analysts may disagree over their preferred imputation procedure they cannot disagree that any imputation procedure must produce an estimate of the median within the bounds just described (Manski, 1995).

The analysis then shifts to the use of relative distributions to describe differences in the lower bound distributions of giving—that is, the distributions that arise if missing amounts are taken as zeros (the lower bound to the estimate of the median is simply the median of the lower bound distribution). Two considerations lead to the decision to work with the lower bound distributions. First, the treatment of missing amounts as zeros is common in charitable giving research. Second, implementing a more sophisticated imputation procedure is complicated by the uneven quality of income data across the surveys. (Income would be the most important variable in any procedure to impute giving.) For those interested in alternative imputations, the present examination of the lower bound distributions can serve as a benchmark.

An intuitive description of the relative distribution of a comparison data set to a reference data set, for example giving in the General Social Survey to Giving and Volunteering in the United States, is to construct histogram bins defined by the deciles of Giving and Volunteering (the reference) and then place the General Social Survey (comparison) data into these bins. If the underlying distributions are the same, then aside from sampling variation the relative histogram would be uniform. If instead the General Social Survey measures higher giving than Giving and Volunteering does, the relative histogram would be skewed right. In this way the relative distribution gives an easily interpretable, visual description of the differences between two empirical distributions.

Figure 1 illustrates the construction of the relative distribution, $G(r)$, of a comparison random variable Y relative to a reference random variable Y_0, with respective cumulative distribution functions (cdfs) F and F_0. (See Handcock & Morris, 1999, Chapter 2, for additional details; this brief discussion follows their treatment.) In Figure 1B, y_r is the rth quantile of the reference distribution F_0. In Figure 1A, $G(r)$ is the probability that Y will take on values less than this quantile; that is, $G(r) = F(Y \leq y_r)$. Because the derivative $g(r) = dG(r)/dr$ is a valid probability density function—it is the density of the random variable defined to be the rank of Y relative to Y_0 (the realizations of these ranks are the relative data)—the relative distribution can be used in formal statistical analysis.

In particular, a test that two random variables have equal cdfs can be conducted by testing whether their relative distribution is uniform. Suppose we are interested in testing the equality of the cdfs at a $k \times 1$ vector of percentiles, γ; that is, H_0: $\boldsymbol{G}(\gamma) = \boldsymbol{\gamma}$, where $\boldsymbol{G}(\gamma)$ represents a $k \times 1$ vector whose i-th element is $G(\gamma_i)$. The test statistic $(\boldsymbol{G}(\gamma) - \gamma)'\Sigma^{-1}(\boldsymbol{G}(\gamma) - \gamma)$ is asymptotically chi-square
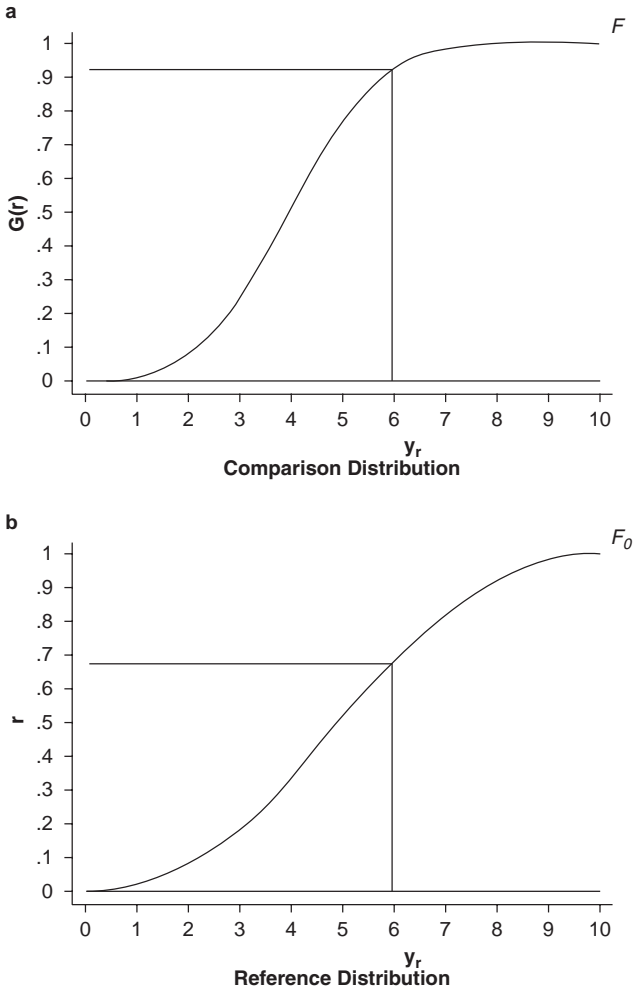
Figure 1a. Comparison Distribution
Figure 1b. Reference Distribution

with k degrees of freedom under the null. The i,j-th element of the covariance matrix Σ is ($i \geq j$):

$$\frac{G(\gamma_i)(1 - G(\gamma_j))}{m} + \frac{(\gamma_i(1 - \gamma_j)\, g(\gamma_i)\, g(\gamma_j)}{n} \tag{1}$$

where γ_i and γ_j are the i-th and j-th elements of $\boldsymbol{\gamma}$, n is the number of observations in the reference data set, and m is the number of observations in the comparison

data set (Handcock & Morris, 1999, Theorem 9.2.2.1). A kernel density estimate is used for $g(\cdot)$; because $g(\cdot)$ is uniform under the null, the relative data are reflected around zero and around one prior to the estimation. The test results in this article are not sensitive to typical bandwidth choices.

THE SIX DATA SETS

The data sets included in the analysis were purposefully selected. Giving and Volunteering is the present-day standard among U.S. household surveys of giving—it is frequently used by researchers, practitioners, and policy makers (e.g., Andreoni, Brown, & Rischall, 2003; Andreoni, Gale, & Scholtz, 1996; Clotfelter, 1997; Council of Economic Advisors, 2000; *Nonprofit Almanac*, 2002; U.S. Census Bureau, 2001, Table 560). As such, it is the obvious choice for the reference distribution with which to begin the relative distribution analysis. The General Social Survey is included because it adopted the Giving and Volunteering questions and thereby allows the comparison of results from identical output cue instruments fielded in the same year (1996) by different survey organizations.[1] Moreover, the General Social Survey is a heavily used data set among sociologists and economists. Giving and Volunteering in California is included because it is the only U.S. survey to use the input cue approach. Canada's Survey of Giving, Volunteering and Participating is included because it pioneered the input cue approach. The 1974 National Study of Philanthropy is included because it is the only survey of giving to have had the government use administrative records (income tax returns) to pinpoint a high-income oversample. Despite the National Study's age it is still used for research purposes (e.g., Schiff, 1990; Duncan, 1999). Finally, the Center Panel is included because it is the initial wave of the first panel survey of giving. Because the Center Panel is a part of the ongoing PSID, the future use of the Center Panel by sociologists and economists is anticipated to be high. Introductions and discussions of basic findings are available for Giving and Volunteering (Hodgkinson & Weitzman, 1996), the California study (O'Neill & Roberts, 2000), Canadian study (Hall et al., 1998), National study (Morgan, Dye, & Hybels, 1977), and the Center Panel (http://www .philanthropy.iupui.edu/COPPS.html).

Table 1 summarizes several important features of the surveys. The opinion research firms that fielded the Giving and Volunteering, General Social, and California surveys use interviewers not extensively trained to elicit responses to questions about dollar amounts.[2] In contrast, the organizations that fielded the Canadian, National, and Center Panel studies are widely recognized as expert in eliciting responses to questions about dollar amounts; their interviewers are better trained in this regard.

Each survey queries total giving through a series of component questions. As already discussed, the surveys differ in how these component questions cue respondent recall, and these cues are listed in Column 3. Although the

Table 1. Surveys With Data on Charitable Giving

Survey	Organization Fielding the Survey	Recall Cue	Response Rate	Special Characteristics	Scale Factor[a]
Giving and Volunteering in the United States (1996)	Gallup	Output	19[b]	—	1.22
General Social Survey (1996)	National Opinion Research Center	Output	76	—	1.22
Giving and Volunteering in California (2000)	Hebert	Input	35	Incomplete total for 40% of married couples. Response incentive—nonrandom	0.94
National Survey of Giving, Volunteering and Participating (1997)	Statistics Canada	Input	78	Incomplete total for 60% of married couples	1.20
National Study of Philanthropy (1974)	Michigan Survey Research Center and the U.S. Census	Output	78	High-income oversample. Total unknown if less than $100 (in 1973 dollars)	4.51
Center on Philanthropy Panel Study (2001)	Michigan Survey Research Center	Output	66[c]	$25 threshold	0.96

a. The scale factor is used to multiply the original survey data to account for growth in average household nominal income between the year covered by the survey and 1999.
b. The actual rate has not been published. This figure is taken from Kirsch, McCormack, and Saxon-Harrold's (2001) general discussion of response rates in the Giving and Volunteering series.
c. This is the cumulative response rate since 1989. The annual response rates vary between 95 and 98%.

71

survey questions cannot be presented in their entirety, it is helpful to see an example of each type of cue. (A more detailed description of all six questionnaires is available upon request.) Here is an illustration of an output cue taken from the Center Panel:

> Did you make any donations specifically for religious purposes or spiritual development, for example to a church, synagogue, mosque, TV or radio ministry? Please do not include donations to schools, hospitals, and other charities run by religious organizations. I will be asking you about those donations next.
>   [If yes] Altogether, what was the total dollar value of all donations [you/you and your family] made in 2000 towards religious purposes?

Subsequent components use a similar structure to ask about donations to educational organizations, health organizations, and so forth. The following illustrates an input cue taken from the Canadian survey:

> In the past 12 months, have you made a charitable donation by responding to a request through the mail?

If the answer is yes, the respondent is asked a series of questions to learn the number of such donations and their total amount. This is followed by components asking about donations in response to requests by co-workers, through payroll deduction, in memoriam, at public cash boxes, to door-to-door collectors, etc. Many of these specified methods of transferring contributions into charitable organizations are used for small donations, and by focusing on them input cue surveys might be expected to measure smaller levels of giving.

Column 4 lists survey response rates. Gallup and Hebert did not use multiple call backs and this likely explains the very low response rates in the Giving and Volunteering and California surveys. Response rates at these levels are cause for concern. The Center Panel's response rate is based on attrition from the PSID since 1989. Although attrition is always a matter of concern in panel studies, there is evidence suggesting that the PSID remains representative.[3]

Four other characteristics of the surveys are important to keep in mind and are listed in Column 5. First, in contrast to the other surveys, the California and Canadian surveys attempted to obtain an individual unit of account by subtracting one half of donations respondents made jointly with their spouses. This subtraction can be undone for some respondents but not for 40 and 60 percent of the married couples in the respective surveys, and for this reason alone giving measured by these surveys may be less than that measured by the other surveys. Unfortunately, the two affected surveys are also the two expected to produce lower measures of giving because of their use of input cues. Second, in an attempt to increase response rates the California survey offered to make charitable contributions on behalf of respondents, perhaps

producing a nonrandom sample selection toward generous respondents. Third, the National Study sought to determine the exact amount given to charitable organizations only if the respondent initially reported that their total contributions were greater than $100 (1973 dollars); thus, for many donors all we know is that their gifts were between $1 and $100. Fourth, the first question in the Center Panel screened out respondents who said they did not give more than $25. Hence, some panel respondents that I classify as *nondonors* may have, in fact, given amounts between $1 and $25 dollars. This classification will be further examined in the sensitivity analysis.

Several data handling decisions are important to mention. Attention is restricted to respondents who were heads of households or spouses of heads in the surveys that randomly selected respondents from a household listing. Unless otherwise indicated, survey weights are used for the Giving and Volunteering, California, Canadian, and National studies. Neither the General Social Survey nor the Center Panel require weighting, in the latter case because I use the PSID's nationally representative subsample. To account for inflation and real income growth, the data are scaled to 1999 levels using the growth in average household nominal income between the year covered by the survey and 1999. The scale factors are listed in the last column of Table 1.

RESULTS

Table 2 presents the prevalence of missing data in each of the surveys. For each survey, the first column lists the unweighted fraction of respondents having no missing data in any of the component questions and the last column contains the number of component questions about giving used in each survey. The remaining columns list the fractions having the indicated number of missing components. Column 1 shows that only 65% of the Giving and Volunteering respondents had no missing data. The 35% with some missing data consist of 7.6% missing only 1 of the 12 components, 2.2% missing 2 components, 3.1% missing between 3 and 5 components, and 22.3% missing 6 or more components. Clearly, missing data are extensive in Giving and Volunteering.

The percentages of respondents with no missing data in the General Social and California surveys are similar to that in Giving and Volunteering (71% and 67.3%); however, the fractions of respondents missing six or more components are much smaller in these surveys. Missing data are much less prevalent in the Canadian, National, and Center Panel surveys—the surveys fielded by organizations experienced in collecting dollar-amount information. In addition to interviewer training, the very infrequent missing data in the Center Panel—only 1.2% of the respondents have any missing data— may result from the experience PSID respondents have gained over the years answering questions about dollar amounts.

Table 2.   The Fraction of Respondents With Missing Data

Survey	Number of Component Questions With Missing Data					Number of Component Questions
	None	*1*	*2*	*3 to 5*	*6 or more*	
Giving and Volunteering in the United States	.649	.076	.022	.031	.223	12
General Social Survey	.710	.106	.031	.021	.133	12
Giving and Volunteering in California	.673	.168	.065	.072	.022	17
National Survey of Giving, Volunteering, and Participating (Canadian)	.836	.105	.027	.026	.006	17
National Study of Philanthropy	.917	.046	.018	.020	—	4
Center on Philanthropy Panel Study	.988	.008	.002	.002	.001	6

Note: For each survey, the table reports the fractions of respondents with various numbers of giving component questions containing missing data. The data are not weighted.

Figure 2 indicates how the missing data described in Table 2 can affect what each survey unambiguously reveals about the median gift. The figure presents lower and upper missing data bounds (the bottom and top of the vertical lines) on the median gift among the subsample of givers. The large distance between the lower and upper bounds in Giving and Volunteering, the General Social Survey, and the California survey reflects the difficulty these surveys had in getting respondents to answer the amount questions; in determining giving percentiles, the prevalence of missing data discussed in Table 2 has a big effect.

Compare the width of the missing data bounds to the sampling variation confidence intervals that would be reported in the typical giving study that ignores the missing data bounds. In Figure 2, the 95% confidence intervals (bootstrapped) for the lower bound medians are shown by the triangles; the intervals are very narrow compared to the wide missing data bounds in the Giving and Volunteering, General Social, and California surveys. Hence, the confidence intervals typically reported for these surveys could well be dominated by missing data problems.

The other three surveys with more "dollar experience" did much better: The distances between the lower and upper missing data bounds are in the $50 to $70 range in the Canadian, National, and Center Panel studies. These ranges are similar in magnitude to the respective confidence intervals for the lower bound medians.

As previously mentioned, the lower missing data bound on the median is the median of the lower bound distribution. More detailed statistics for the lower bound distributions are presented in Table 3. The first row contains the
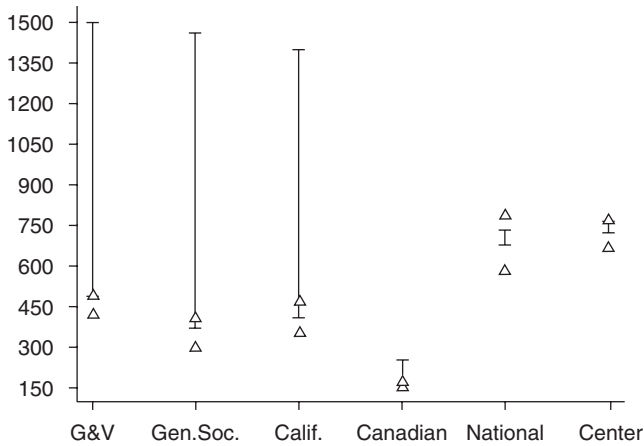
Figure 2. Median Gifts: Missing Data Bounds and Confidence Intervals
Note: G & V = Giving and Volunteering; Gen.Soc. = General Social Survey; Calif. = California survey. Ends of the intervals are the lower and upper bounds to the medians based on assigning missing data to zero or above the median. Triangles are 95% confidence intervals (bootstrapped) on the lower bounds. The Giving and Volunteering upper bound is indeterminate because just over one half of the head and spouse donors have some missing information about their donated amounts.

lower bound fractions that give. With the exception of the National Study, the input cue surveys (the California and Canadian studies) measure a higher fraction of the population who gives. Row 2 shows the average amounts given (including respondents who did not give in the average); the National Study with its good high-income oversample measures average giving higher than do the other surveys by a non-negligible amount. Row 3 shows that excluding the single largest gift affects the average in several surveys, especially the General Social, illustrating that the average is not necessarily a robust indicator of systematic differences between distributions. The same point applies to the average among the subsample of givers in Rows 4 and 5.

Systematic differences between distributions can be gleaned from patient study of the conditional distribution deciles in Table 3, Rows 6 through 14. Alternatively, the differences are quickly transparent in Figure 3's relative histograms. All the histograms are relative to Giving and Volunteering; hence the first histogram—Giving and Volunteering relative to itself—is uniform by definition. The next histogram shows that except at the tenth decile the General Social Survey measures more small gifts than does Giving and Volunteering.[4] The third histogram shows that the California survey also measures more small gifts than does Giving and Volunteering. The three remaining histograms show larger differences. The amounts in the Canadian survey are concentrated in the lowest three Giving and Volunteering deciles

Table 3. Summary Statistics for the Lower Bound Distributions of Giving

	Giving and Volunteering	General Social Survey	California	Canadian	National Study	Center Panel
Percentage donating	0.703	0.708	0.954	0.899	0.880	0.687[a]
Average amount given	901	1,218	1,132	507	2,109	1,341
Average without the maximum gift	798	835	1,042	504	2,087	1,244
Amount given, conditional on giving						
Average	1,399	1,951	1,256	492	2,435	1,979
Average without the maximum gift	1,239	1,337	1,155	489	2,409	1,837
Deciles 1	49	43	33	22	5	96
2	122	94	94	42	5	216
3	220	134	169	72	5	337
4	306	244	277	118	5	480
5	488	366	408	168	677	720
6	612	549	610	244	1,128	1,074
7	976	781	915	348	1,691	1,668
8	1,586	1,403	1,407	522	2,706	2,648
9	2,769	3,050	2,678	1,020	4,510	4,512
Maximum amount given	118,950	487,981	159,000	48,072	2,255,000	471,840
Next largest amount given	61,000	62,220	79,451	32,753	1,578,500	73,728
Number of observations	2,424	1,270	2,037	15,745	2,802	4,887

Note: Deciles, averages, and maxima are in 1999 dollars. The data are weighted (the numbers of observations are not).
a. The percentage donating in the Center Panel indicates whether the respondent gave more than $25; in the other surveys there is no such threshold.
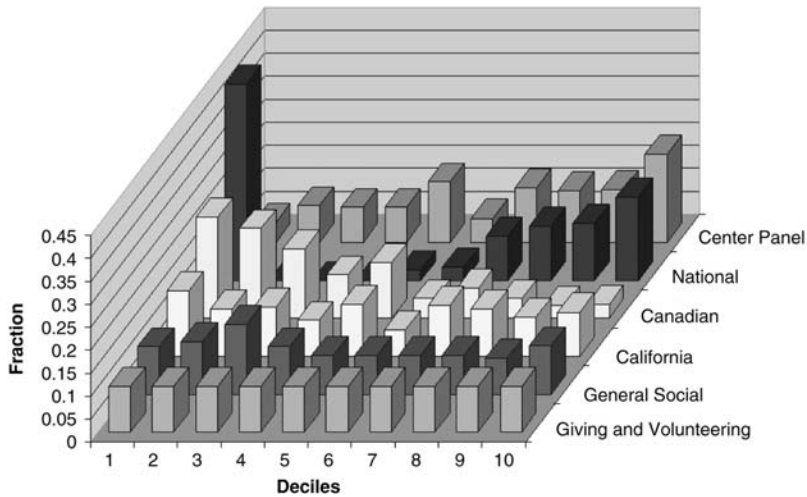
Figure 3.   Relative Distributions of Conditional Giving
Note: The figure presents the conditional distributions of giving (positive givers only) of the 1996 General Social Survey; the 2000 Giving and Volunteering in California; the 1997 Canadian Survey of Giving, Volunteering and Participating; the 1974 National Study of Philanthropy; and the 2001 Center Panel. Each of these five conditional distributions is presented relative to the conditional distribution of the 1996 Giving and Volunteering in the United States.
The conditional distribution of the 1996 Giving and Volunteering in the United States is presented relative to itself (the first histogram in the figure) and hence by definition is the uniform distribution; this makes it clear that the 1996 Giving and Volunteering in the United States is the reference distribution in the figure's relative distributions.

and the amounts from the National and Center Panel studies fall more frequently in the top deciles, especially the tenth.[5]

Table 4 examines the statistical significance of these differences using the test statistic in Equation 1. The differences between the General Social and Giving and Volunteering studies are statistically significant in the lower and middle deciles; the evidence is less strong for differences in the top three deciles ($p = .071$). The differences between the California and Giving and Volunteering studies follow a similar pattern: significant in the lower and middle deciles but not at the top. All the differences between the remaining surveys and Giving and Volunteering are statistically significant.

Because the National Study contains a good quality high-income oversample it makes more sense to use it rather than Giving and Volunteering as the reference distribution when analyzing differences at the top of the giving distributions. I focus on differences between the Center Panel and the National Study because it is clear from Table 3 (and implicitly from Figure 3) that the other four data sets measure much smaller amounts than does the National Study. (Statistical tests strongly confirm this.) The histogram of the Center Panel relative to the National Study in Figure 4 appears nearly uniform (aside

Table 4. Tests of Equality of Conditional Distributions of Giving With Giving and Volunteering in the United States

	Quantiles		
Data Sets	*.10, .20, .30*	*.40, .50, .60*	*.70, .80, .90*
General Social Survey	13.6 (.004)	12.6 (.006)	7.0 (.071)
Giving and Volunteering in California	18.3 (.000)	26.7 (.000)	2.4 (.494)
National Survey of Giving, Volunteering and Participating (Canadian)	309.8 (.000)	535.4 (.000)	635.0 (.000)
National Study of Philanthropy	2405.7 (.000)	369.6 (.000)	50.3 (.000)
Center on Philanthropy Panel Study	111.7 (.000)	138.3 (.000)	68.1 (.000)

Note: In Column 1, the null hypothesis is that the .10, .20, and .30 quantiles of the data set in a selected row are equal to those in Giving and Volunteering. In Column 2, the null hypothesis is based on the .40, .50, and .60 quantiles and in column 3 the .70, .80, and .90 quantiles. The results are the χ^2 test statistic. *p* values are in parentheses.

from the mass of Center Panel amounts in the fifth National Study decile)[6] indicating close agreement between the two surveys through the tenth decile. Testing that the two distributions are equal at the .70, .80, and .90 quantiles shows that they are statistically indistinguishable ($p = .888$).

The findings that Center Panel and the National Study are indistinguishable through the 90th percentile while the Center Panel conditional mean is 19% lower (see Table 3, Row 4) imply that differences must exist higher in the distribution. Figure 5 graphs the top 10 percentiles from each data set and shows that they match fairly closely through the 92nd percentile: The test of equality has a *p* value of .177. The divergence in the plots at the 93rd percentile is more noticeable ($5,472 in the Center Panel and $6,089 in the National Study), and the *p* value for this difference is .012.

SENSITIVITY ANALYSIS

The close match between the Center Panel and the National Study at the top of the distribution may be sensitive to classifying all those in the panel who were screened out by the $25 threshold question as if they were non-donors. Some of those screened likely gave small amounts and including them in the conditional distribution would shift the Center Panel distribution down and away from the National Study. To check this, I reclassified 6.6% of those screened (and previously treated as nondonors) to now be donors of small amounts and included them in the Center Panel's conditional distribution (6.6 is the percentage of respondents in the California study giving between 1 and 25 dollars—the highest such estimate from all the data sets). The top half of the reclassified conditional Center Panel distribution relative to the National Study is less uniform than in Figure 4, and the difference at the upper three deciles is on the margin of significance at 10% ($p = .108$). Nevertheless, the reclassified Center Panel distribution still matches the National Study more closely than the other four data sets.
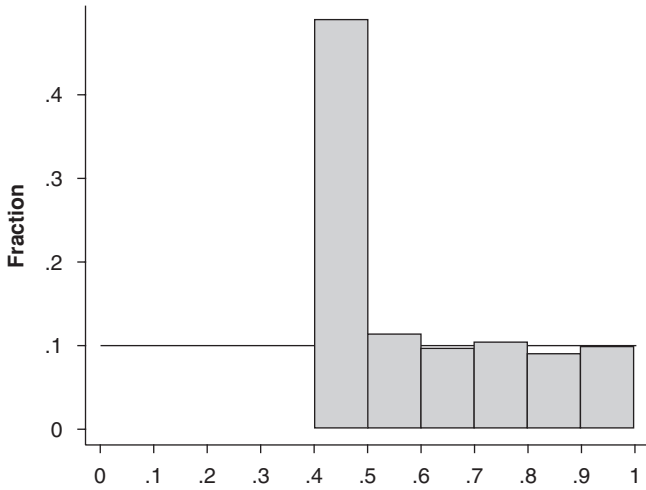
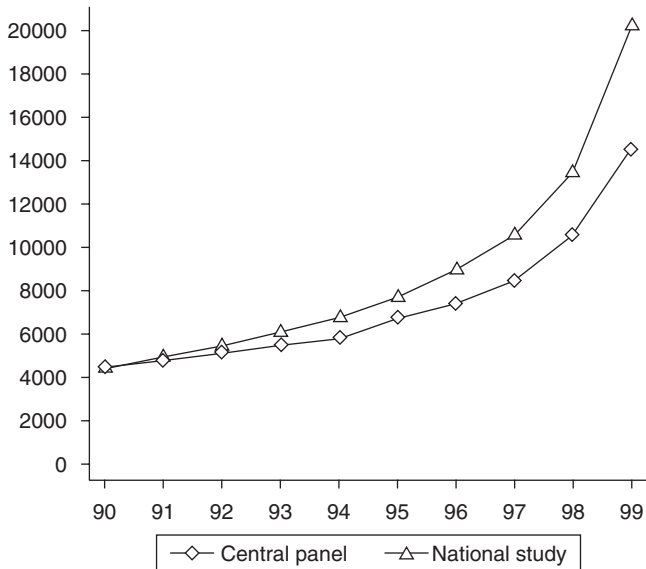**Figure 4. Center Panel Relative to the National Study**

Figure 5. Top Percentiles of the Center Panel and National Study

Finally, the result that among recent giving surveys the Center Panel provides better measurement of giving at the top of the distribution is robust to choosing an alternative reference distribution for comparisons at the top.

Wilhelm (2006) compared the Center Panel and Giving and Volunteering data to charitable deductions deciles from the Internal Revenue Service and found that the Center Panel matches the deduction deciles much higher into the distribution than does Giving and Volunteering.[7]

DISCUSSION AND CONCLUSIONS

This study is the first to analyze the quality and comparability of survey data on charitable giving by using several major household surveys of giving, describing missing data patterns, and using relative distribution methods. The analysis produces four main results.

First, survey organizations less experienced in getting answers to dollar-amount questions produce much more missing data when asking about giving; missing amount data in the Giving and Volunteering, General Social, and California surveys may have larger effects on giving estimates than does sampling variation. Second, although input cue surveys measure higher fractions of the population who give, Figure 3 suggests that the additional givers detected by input cues donated relatively small amounts.

Third, the design of the survey instrument appears to interact with the dollar-amount experience of the interviewers in producing estimates of the distribution of giving. The two surveys using output cues and dollar-experienced interviewers (the National Study and Center Panel) measured higher giving than either (a) surveys using output cues but without dollar-experienced interviewers (Giving and Volunteering and General Social) or (b) the survey using input cues and dollar-experienced interviewers (Canadian).[8] Additional evidence supporting the instrument-interacts-with-interviewers conclusion comes from Rooney et al. (2005); they fielded the Center Panel and Giving and Volunteering instruments using the same survey research organization (one without dollar-experienced interviewers) and obtained similar giving data from both instruments.

Fourth, among the surveys not containing a high-income oversample, only the Center Panel appears to be able to estimate giving high into the conditional distribution as judged by comparisons to the National Study and IRS data.[9] Even so, estimating the distribution of giving much above the 90th percentile does not seem to be feasible without a high-income oversample.

Several qualifications should be kept in mind when considering these conclusions. First, although I find the interviewers' dollar-amount experience a satisfying explanation of the pattern of results, there may be alternative explanations rooted in other differences between the surveys. It must be remembered that I am attempting to draw conclusions from what is, in effect, a small sample of six data sets.

Second, the conclusion that additional givers detected by input cue surveys donate small amounts might be altered had the Giving and Volunteering, General Social, and California surveys not had extensive problems with missing data. True, the conclusion is backed by the lower estimates from the Canadian

survey (which had less missing data) but those lower estimates may be because of differences in Canada's expenditure policy and social conditions or even the survey's treatment of married respondents; for these reasons, the conclusion that input cue surveys produce lower giving estimates even when administered by dollar-experienced interviewers is accordingly tentative.

Finally, it is not known whether choice of recall cue affects estimates of giving to specific charitable purposes such as religious giving or giving to help the needy. The advice a researcher should take from the present results on total giving is that the distribution of giving to a specific charitable purpose may well differ across surveys and that he or she should compare the distributions of (specific) giving from candidate surveys to consider whether any differences will have implications for his or her research. This article illustrates how this comparison can be carried out.[10] Also, I have not considered whether the National Study and the Center Panel, for example, produce similar joint distributions of giving and income. These questions are natural next steps in evaluating the quality of giving data.

Although these qualifications must be kept in mind, the results at present suggest that surveys using interviewers trained in obtaining information about dollar amounts and questionnaires that cue on charitable output facilitate the recall of giving. It seems that such a survey would be the appropriate choice for researchers interested in the behavior of people contributing at the median and higher quantiles. Among recent surveys of giving, only the Center Panel meets these requirements.

Finally, researchers should use caution when interpreting the results from mean regressions based on surveys without high-income oversamples because these surveys draw few observations from the top-giving decile and the top decile has a large effect on estimates of the mean gift. This creates two problems for mean regression analysis. First, assuming that regression relationships in the population are the same in the top decile as in the bottom 90th percent, the standard errors on predicted regression relationships for giving in the top decile will be very large. Second, the underlying assumption just evoked—that population regression relationships are the same in the top decile as in the bottom 90th percent—cannot be checked. Accordingly and in addition to using giving data that accurately reflect the median and higher quantiles, giving researchers should make use of quantile regressions to check the sensitivity of mean regression results.

Notes

1. Giving and Volunteering from 1996 is used so that the alignment with the General Social Survey is as close as possible. (The General Social Survey asked the giving questions only in 1996.) Giving and Volunteering was also fielded in 1988, 1990, 1992, 1994, 1999, and 2001.

2. This judgment is not entirely fair to the National Opinion Research Center, which administers the General Social Survey as well as surveys gathering high-quality dollar amount information such as the Survey of Consumer Finances (SCF). It is not known how many General Social Survey interviewers also worked on the SCF but even so the rarity and structure of other dollar-amount

questions in the 1996 General Social Survey suggest that little interviewer training would have been devoted to getting accurate responses to open-ended dollar-amount questions: Aside from the giving questions, only two questions dealt with dollar amounts and each of these were structured to have respondents select an income category from a list.

3. The 66% response rate is calculated by assuming the Panel Study of Income Dynamics (PSID) was representative as of 1989 and then cumulating response rates since then; 1989 was selected as the base year because of Fitzgerald, Gottschalk, and Moffitt's (1998) conclusion that the PSID had remained representative through that year. More recent PSID income data continue to align well with the Current Population Survey (Gouskova & Schoeni, 2002), suggesting that post-1989 attrition has not compromised the representativeness of the PSID.

4. This is different than the impression left by the Schervish and Havens (1998) finding discussed in the introduction. Furthermore, the averages of the two distributions are similar if the single largest gift in the General Social Survey is excluded (see Table 3).

5. The mass of National Study respondents in the lowest Giving and Volunteering decile occurs because the National Study did not query amounts if donors said their gifts totaled $100 or less. The lower bound for such donors is $1 (1973 dollars), which brought forward by the 4.51 scale factor generates the $5 decile boundaries in Table 3.

6. The mass consists of all Center Panel amounts greater than $5 but less than $677; once more this arises because the National Study does not quantify gifts less than $100 (1973 dollars). The very short bar in the first decile contains four Center Panel respondents who reported gifts totaling less than $5 (even though in the initial screening question they said that they gave more than $25).

7. Other than the general practice of conducting sensitivity checks whenever feasible, there are two possible problems motivating the consideration of an alternative reference distribution to the National Study scaled by income growth since 1974: Auten, Clotfelter, and Schmalbeck (2000) provided evidence suggesting that giving in the top deciles has not kept pace with income growth, and Jencks (1986) argued that the National Study may have overestimated giving at the top in the first place. The fact that the Center Panel also compares well to IRS data implies that these two possible problems are not affecting the results.

Because the SCF also obtains government assistance in pinpointing a high-income oversample it seems at first to offer another alternative reference distribution for comparisons at the top of the distribution. Unfortunately, the SCF measures charitable giving with only a single question, potentially creating a serious underreporting problem. Comparing gifts from the Center Panel above $500 (to match the SCF's $500 screen) to the SCF shows much higher giving in the Center Panel through the first eight deciles; the ninth deciles match.

8. Such interactive effects may be part of the explanation of why the present results differ from Rooney et al.'s (2001) finding of higher giving estimates with input rather than output cues—the interviewers collecting the Rooney et al. data did not have experience getting answers to questions about dollar amounts.

9. This is reminiscent of the finding that the PSID's estimate of the wealth distribution matches that obtained with the high-income oversample SCF, although in the case of wealth the match extends through the 98th percentile (Juster, Smith, & Stafford 1999).

10. Along these lines, the 2003 Center Panel has six component questions collecting data on giving to youth and family services, the arts, neighborhood improvement, the environment, international aid, and open-ended purposes. (The 2001 Center Panel has only one component question for all six of these purposes.) This presents an opportunity to study how more detailed questions affect reported amounts within a survey for which interviewers are well-trained in querying, and respondents are well-trained in providing, dollar-amount information.

REFERENCES

Andreoni, J., Brown, E., & Rischall, I. (2003). Charitable giving by married couples: Who decides and why does it matter? *Journal of Human Resources*, *38*(1), 111-133.

Andreoni, J., Gale, W. G., & Scholtz, J. K. (1996). *Charitable contributions of time and money*. Unpublished manuscript, University of Wisconsin–Madison.

Auten, G. E., Clotfelter, C. T., & Schmalbeck, R. L. (2000). Taxes and philanthropy among the wealthy. In J. B. Slemrod (Ed.), *Does Atlas shrug? The economic consequences of taxing the rich*. New York: Russell Sage Foundation and Harvard University Press.

Berking, H. (1999). *Sociology of giving*. Thousand Oaks, CA: Sage.

Clotfelter, C. T. (1997). The economics of giving. In J. W. Barry & B. V. Manno (Eds.), *Giving better, giving smarter: Working papers of the National Commission of Philanthropy and Civic Renewal*. Washington, DC: National Commission on Philanthropy and Civic Renewal.

Council of Economic Advisors. (2000). *Philanthropy in the American economy*. Washington, DC: Author.

Duncan, B. (1999). Modeling charitable contributions of time and money. *Journal of Public Economics*, *72*, 213-242.

Fitzgerald, J., Gottschalk, P., & Moffitt, R. (1998). An analysis of sample attrition in panel data: The Michigan Panel Study of Income Dynamics." *Journal of Human Resources*, *33*, 251-299.

Gouskova, E., & Schoeni, R. F. (2002). *Comparing estimates of family income in the Panel Study of Income Dynamics and the March current population surveys, 1968-1999*. Unpublished manuscript, University of Michigan, Institute for Social Research.

Hall, M. (2001). Measurement issues in surveys of Giving and Volunteering and strategies applied in the design of Canada's National Survey of Giving, Volunteering and Participating. *Nonprofit and Voluntary Sector Quarterly*, *30*, 515-526.

Hall, M., Knighton, T., Reed, P., Bussiere, P., McRae , D., & Bowen, P. (1998). *Caring Canadians, involved Canadians: Highlights from the 1997 National Survey of Giving, Volunteering and Participating*. Ottawa: Statistics Canada.

Handcock, M. S., & Morris, M. (1999). *Relative distribution methods in the social sciences*. New York: Springer-Verlag.

Hodgkinson, V. A., & Weitzman, M. S. (1996). *Giving and volunteering in the United States: Findings from a national survey*. Washington, DC: Independent Sector.

Jencks, C. (1986). Who gives to what? In W. W. Powell (Ed.), *The nonprofit sector: A research handbook*. New Haven, CT: Yale University Press.

Juster, F. T., Smith, J. P., & Stafford, F. (1999). The measurement and structure of household wealth. *Labour Economics*, *6*, 253-275.

Kirsch, A. D., McCormack, M. T., & Saxon-Harrold, S. K. E. (2001). Evaluation of differences in giving and volunteering data collected by in-home and telephone interviewing. *Nonprofit and Voluntary Sector Quarterly*, *30*, 495-504.

Komter, A. E. (Ed.). (1996). *The gift: An interdisciplinary perspective*. Amsterdam, the Netherlands: Amsterdam University Press.

Lee, L., Piliavin, J. A., & Call, V. R. A. (1999). Giving time, money, and blood: Similarities and differences. *Social Psychology Quarterly*, *62*, 276-290.

Manski, C. F. (1995). *Identification problems in the social sciences*. Cambridge, MA: Harvard University Press.

Morgan, J. N., Dye, R. F., & Hybels, J. H. (1977). Results from two national surveys of philanthropic activity. In *Research papers sponsored by the Commission on Private Philanthropy and Public Needs, Volume I: History, trends, and current magnitudes*. Washington, DC: U.S. Department of the Treasury.

Nonprofit almanac: The new nonprofit almanac and desk reference. (2002). San Francisco: Jossey-Bass.

O'Neill, M., & Roberts, W. L. (2000). *Giving and volunteering in California*. San Francisco: University of San Francisco.

Rooney, P. M., Mesch, D. J., Chin, W., & Steinberg, K. S. (2005). The effects of race, gender, and survey methodologies on giving in the US. *Economics Letters*, *86*, 173-180.

Rooney, P. M., Steinberg, K. S., & Schervish, P. G. (2001). A methodological comparison of giving surveys: Indiana as a test case. *Nonprofit and Voluntary Sector Quarterly*, *30*, 551-568.

Schervish, P. G., & Havens, J. J. (1998). Embarking on a republic of benevolence? New survey findings on charitable giving. *Nonprofit and Voluntary Sector Quarterly*, *27*, 237-242.

Schiff, J. (1990). *Charitable giving and government policy: An economic analysis*. New York: Greenwood.

U.S. Census Bureau. (2001). *Statistical abstract of the United States, 2001*. Washington, DC: Government Printing Office.

Vesterlund, L. (2006). Why do people give? In R. Steinberg and W.W. Powell (Eds.), *The nonprofit sector: A research handbook*. New Haven, CT: Yale University Press.

Wilhelm, M. O. (2006, July 1). New data on charitable giving in the PSID. *Economics Letters*, *92*, 26-31.

Mark O. Wilhelm is an associate professor of economics and philanthropic studies at Indiana University Purdue University–Indianapolis. His most recent research is on the transmission of philanthropic behavior to children, generational changes in giving, the roles of empathy and principles of care in helping behavior, and the effect of low income and family structure disruption during childhood on subsequent giving and volunteering in young adulthood. He is the founding director of the Center on Philanthropy's Panel Study—data on the giving and volunteering of more than 7,000 families that is a module within the Panel Study of Income Dynamics. Waves 2001 and 2003 of the Center on Philanthropy's Panel Study are available at present. Wave 2005 will be available at the end of 2006. The data are freely available via the PSID Web site (http://www.philanthropy.iupui.edu/studies.htm#panel) and easier to use abstracts are available at https://oncourse.iu.edu/access/content/user/mowilhel/Web_page/data.htm. Other information about the Center Panel can be found at http://www.philanthropy.iupui.edu/COPPS.html.