Submissions to Curriculum and Case Notes should be sent to David Reingold, Indiana University, School of Public & Environmental Affairs, 1315 E. 10th Street, Bloomington, IN 47405-1701.

INCREASING THE TRANSPARENCY OF STATED CHOICE STUDIES FOR POLICY ANALYSIS: DESIGNING EXPERIMENTS TO PRODUCE RAW RESPONSE GRAPHS

Dipika Sur, Joseph Cook, Susmita Chatterjee, Jacqueline Deen, and Dale Whittington

Abstract

We believe a lack of transparency undermines both the credibility of, and interest in, stated choice studies among policy makers. Unlike articles reporting the results of contingent valuation studies, papers in the stated choice literature rarely present simple tabulations of raw response data (that is, a table or graph showing the percentage of respondents agreeing to purchase a good or service, or vote for a proposed management plan as a function of price). We describe an approach for adding "policy" tasks to a standard orthogonal-in-attribute-levels research design that allows the researcher to plot raw responses as a function of changes in only one characteristic of the offered good or service. We demonstrate this approach using data from a stated choice study of private demand for cholera and typhoid vaccines in Kolkata, India, carried out in the summer of 2004. © 2006 by the Association for Public Policy Analysis and Management

INTRODUCTION

The use of stated choice experiments (sometimes called choice model experiments) to generate information for policy analysis is growing rapidly in both industrialized and developing countries. A recent review by Adamowicz (2004) found that publication of papers on choice experiments for environmental valuation has grown from 1 in 1991 to 18 in 2003 and that the trend for such studies is upward. His tally

Journal of Policy Analysis and Management, Vol. 26, No. 1, 189–199 (2006)

Published by Wiley Periodicals, Inc. Published online in Wiley InterScience (www.interscience.wiley.com)
DOI: 10.1002/pam.20237

excludes choice experiments published in other (nonenvironmental) fields such as health and marketing.

Proponents of stated choice (SC) experiments list several advantages over the related contingent valuation (CV) technique.[1] First, SC experiments are more efficient in that they collect more valuation information from each respondent. These individual-level data allow the analyst better means for modeling the heterogeneity of preferences. Second, where the attributes of a given policy or project still remain to be designed, SC experiments allow the analyst to predict willingness to pay (WTP) for many different versions of the program without resorting to large split-sample CV studies. Third, SC experiments allow the analyst to observe more directly how each respondent trades off different program attributes, again without relying on a split-sample CV experiment and comparisons *between* respondents.

The contingent valuation technique was subjected to rigorous examination both inside and outside academic economics throughout the early 1990s, spurred by interest in the Exxon *Valdez* oil spill damage assessment (Diamond & Hausman, 1994; Hanemann, 1994; Portney, 1994). An expert panel convened by the National Oceanic and Atmospheric Administration (NOAA) in 1993 (Arrow et al., 1993) proposed a set of minimum standards that should characterize high-quality CV studies.

Stated choice studies have not been subjected to the same degree of criticism or held to the same standards. One might reasonably see SC experiments as an extension of the CV approach and conclude that most of the recommendations from the CV debate (that is, constructing credible scenarios and conducting high-quality surveys) should be applied equally to both SC and CV studies. That view overlooks the fact that the questions posed to respondents in a stated choice exercise are often cognitively more difficult to answer than a single referendum-format CV question. Respondents in SC surveys must process more information, make trade-offs that are (by the researcher's design) nonobvious, and then repeat the choice exercise anywhere from 3 to 20 times. Few SC studies report the number of stated choice respondents who answer in ways that violate economic theory; that is, transitivity, monotonicity, etc. (for exceptions, see Alpizar & Carlsson, 2003; Carlsson & Martinsson, 2001; Cook, Whittington, Canh, Johnson, & Nyamete, 2006; and Johnson, Mathews, & Bingham, 2000). Interest is now growing in understanding how choice complexity qualitatively and quantitatively affects respondents' answers (DeShazo & Fermo, 2002; Swait & Adamowicz, 2001).

CV researchers often attempt to demonstrate that the data underlying their analyses are of high quality by reporting raw results from the discrete choice (referendum) valuation question, generally in the form of a table or graph showing how the percent of respondents agreeing to pay varies by price. Indeed, most consumers of

[1] Contingent valuation studies elicit a respondent's willingness-to-pay for only one hypothetical good or service. In contrast, stated choice (choice model) studies typically ask respondents to choose between several different configurations of the good or service. To use an example from our research, suppose the researcher would like to estimate an individual's willingness to pay for a cholera vaccine where the attributes of the vaccine are its price, effectiveness in protecting against the disease, and the duration of protection. A CV approach would present each respondent with only one vaccine configuration (for example, a cholera vaccine that is 50% effective for 3 years and costs US$2) and ask the respondent if she would purchase the vaccine if it were available. A different subsample of CV respondents would be asked about a different configuration (that is, different cost, different effectiveness), so that the necessary comparisons are made *across* subsamples. A stated choice approach would ask each respondent to choose between two or more different alternatives—a cholera vaccine that is 70% effective for 3 years with a price of US$10, or a cholera vaccine that is 50% effective for 3 years with a price of US$2. Respondents typically complete several of these choice tasks, so that the researcher observes how *each individual* trades off the different attributes of the good.

this literature would probably have doubts about any study that did not report such information. CV researchers then proceed to analyze responses using multivariate statistical models that better explain the determinants of demand; but readers typically regard that original graph or table of raw responses as a first quality check.

Stated choice studies, on the other hand, rarely report (in either table or graph form) raw responses to choice tasks. Most papers list the attributes of the program or management plan used in the design, including the price levels, and discuss the procedure for assembling these attributes into choice tasks. Most SC studies offer only one example from the array of choice tasks that each respondent was asked to perform. Readers of stated choice papers thus have very little opportunity to gain an intuition about the quality of the underlying choice data without relying on an understanding of multinomial or mixed logit estimation techniques. We report here a straightforward approach for adding tasks to a stated choice experimental design that will yield graphs that give the reader a transparent, simple understanding of the underlying raw data in the same way that the familiar CV graph (percent of the sample that said (voted) "yes" versus price) does.

Although quite useful for non-technical consumers of these studies, the advantage of these policy (or raw response) graphs is not simply in communicating results. Similar graphs can be constructed with predictions based on the results of the statistical models. The principal advantage is that the graphs provide an intuition on the underlying *raw* data. This is important for several reasons. First, raw response data are critically important in the early stages of research design (for example, pretesting) when the final attribute levels to be included in the research design (namely, price) are still uncertain. Second, like in CV studies, results based on raw data can be used to validate more sophisticated econometric results. If the raw data show little response to price, for example, one would be quite skeptical of statistical results showing a price effect. Third, the transparency of raw response data may be especially important because health and environmental valuation papers are often read by non-economists and members of the policy community, who are often less familiar with econometric models. These readers would likely have more confidence in studies that present results based on both raw data and predicted choices.

DESIGNING CHOICE TASKS FOR POLICY GRAPHS

To maximize efficiency and minimize the number of choice tasks that each respondent must complete, stated choice researchers typically use orthogonal choice task designs. These designs require that the attributes of a good or program have different levels for each alternative in a choice task. For example, suppose a choice task asks a respondent to choose between three cholera vaccines, and that the attributes of each vaccine are its effectiveness, duration, and price. An orthogonal design requires that all three cholera vaccines offered in one choice task have different prices and different levels of effectiveness and duration. Though efficient, this design makes it impossible to construct simple graphs to display the raw data because too many variables change at once.

Our basic approach, by contrast, is to add several non-orthogonal tasks that are carefully designed to allow a comparison where only one variable is changing at a time. Adding these non-orthogonal tasks does come at a cost to the research effort. Many SC researchers may feel that a study's resources are better spent in adding additional attributes or levels (or in making tasks simpler). From an efficiency standpoint, the non-orthogonal tasks we propose do a poor job of collecting additional preference information from respondents. We believe, however, that the additional transparency achieved by a view of raw response data is often worth this opportunity cost.

To illustrate the advantage of this additional transparency with a concrete example, we use data from a stated choice study of private demand for cholera and typhoid vaccines in Kolkata, India, carried out in the summer of 2004. As part of a larger stated preference study (including both CV and SC surveys) in Kolkata, we administered 200 in-person, stated choice interviews in the Beliaghata neighborhood of Kolkata. The neighborhood is a mix of lower- and middle-income households; mean per capita income was on the order of US$1 per day. Respondents were selected randomly from the most recent voter rolls.

The interviewer gave each respondent information on the two diseases, explained our hypothetical cholera and typhoid vaccines, and had the respondent complete one practice choice task. To provide time to think, the interviewer then gave the respondent an informational card showing all of the possible vaccine attributes and levels (see Table 1) to review overnight. The interviewer returned the next day to complete the actual choice tasks and to collect socioeconomic information about the respondent.[2]

Figure 1 presents a sample choice task that was shown to respondents. Respondents were asked to choose between three alternatives—a cholera vaccine, a typhoid vaccine, or neither. They were not initially given the opportunity to purchase both vaccines. After completing all of the choice tasks, however, the interviewer revisited any task where the respondent said they would purchase a vaccine. For each of these tasks, the interviewer asked the respondent if they would want to purchase *both* vaccines if this were possible.

Using an orthogonal design with the number of attributes and levels shown in Table 1 would have required each respondent to complete 12 choice tasks (we will call these the "orthogonal" tasks). Because this number of tasks seemed too demanding for respondents, we split our sample of 200 into three groups (n = 66, 67, 67 respondents). Each respondent then had to complete 4 orthogonal choice tasks from the total set of 12. The order in which they completed these orthogonal tasks was determined randomly before the interview. In addition, each respondent completed two non-orthogonal "policy" tasks (Table 2). To avoid contaminating the orthogonal design, the policy tasks were always completed after the orthogonal tasks. This raises the possibility, however, that responses may differ between orthogonal and policy tasks because of fatigue. To avoid this, in future applications

Table 1. Attributes and levels used in the vaccine choice tasks.

Attribute	Attribute Level
Vaccine type	Cholera, Typhoid
Effectiveness	50%, 70%, 80%
Duration	3 yrs, 20 yrs
Price[a]	US$0.22, 1.11, 11.11

[a] Prices were given to respondents in Indian rupees (Rs. 10, Rs. 50, Rs. 500).

[2] For a more in-depth discussion of the use of this time-to-think treatment in SC research, see Cook et al., 2006).

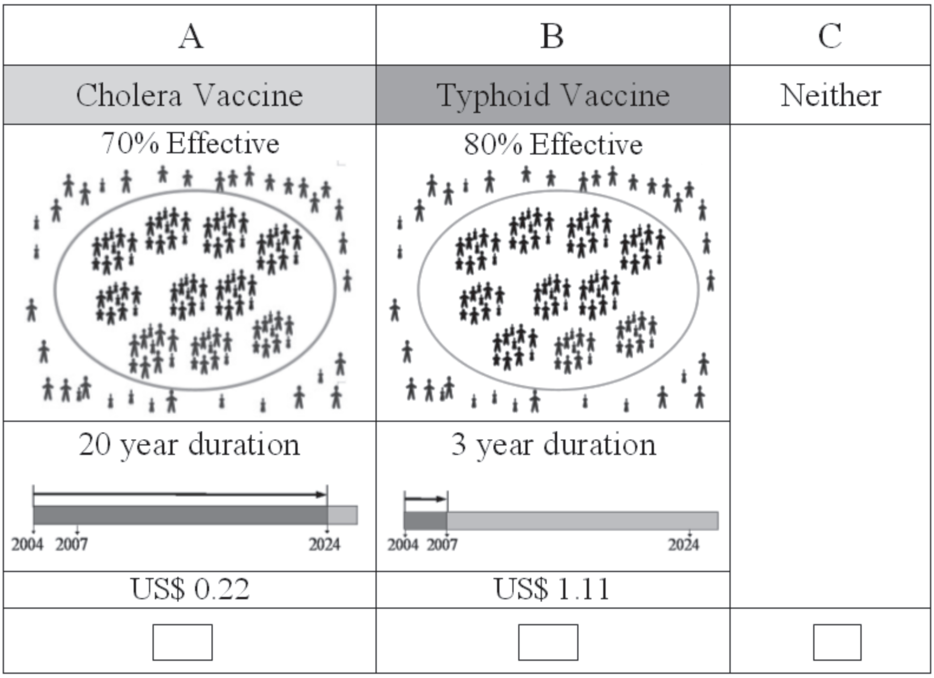


Figure 1. An Example of a Choice Task Used in the Survey.

analysts might explicitly build these non-orthogonal tasks into the overall experimental design.[3]

By comparing the raw responses (that is, what percent of the 66 respondents answering each task said they would purchase neither vaccine, the cholera vaccine only, the typhoid vaccine only, or both vaccines), we can construct graphs that illustrate how different attribute levels affect the choice of vaccine. Comparing tasks A5, B5, and C5 (see Table 2) allows us to compare raw responses when all of the attributes remained constant except the price of the vaccine. Comparing tasks B5 and B6 allows us to compare raw responses when only the vaccine's duration changed. Tasks A6, B6, and C6 allow a comparison of responses by vaccine effectiveness. Because the cholera and typhoid vaccines have the exact same attributes on all of the policy tasks (again, *not* the case in the orthogonal tasks), the raw responses to each of these choices also present a direct, easily interpretable comparison of demand for cholera and typhoid vaccines.

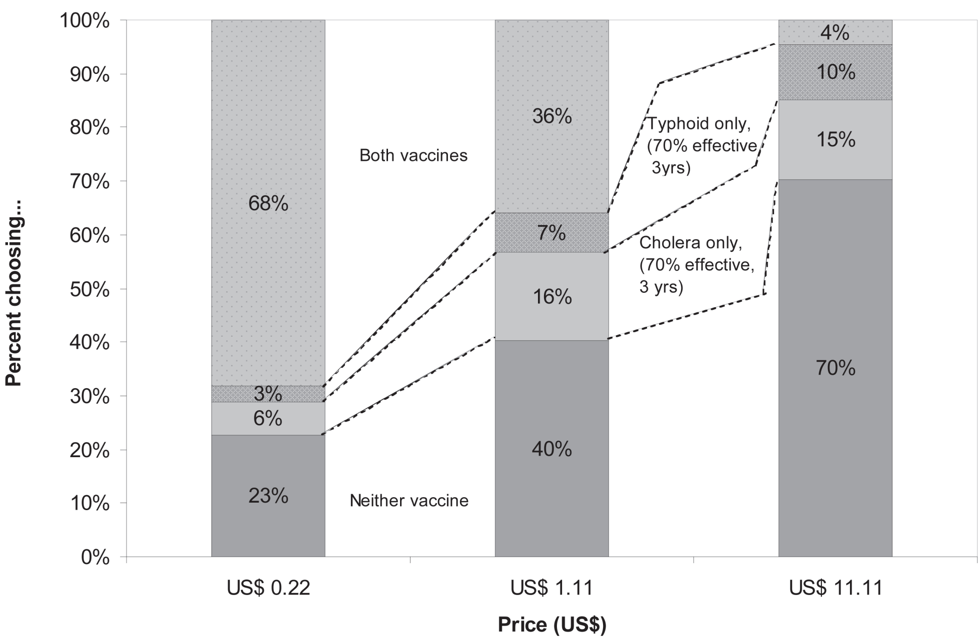
RESULTS

Figure 2 is an analog to the common "percent yes" graph in the CV literature. It shows that as the price of the vaccine increased and all other attributes remained

[3] Building the tasks explicitly into the experimental design has another advantage. Since the policy tasks are likely to be cognitively easier for respondents to answer, the results from these simpler tasks can be used to validate the results from the more efficient, but likely more confusing, orthogonal tasks. (We thank a reviewer for pointing this out.)

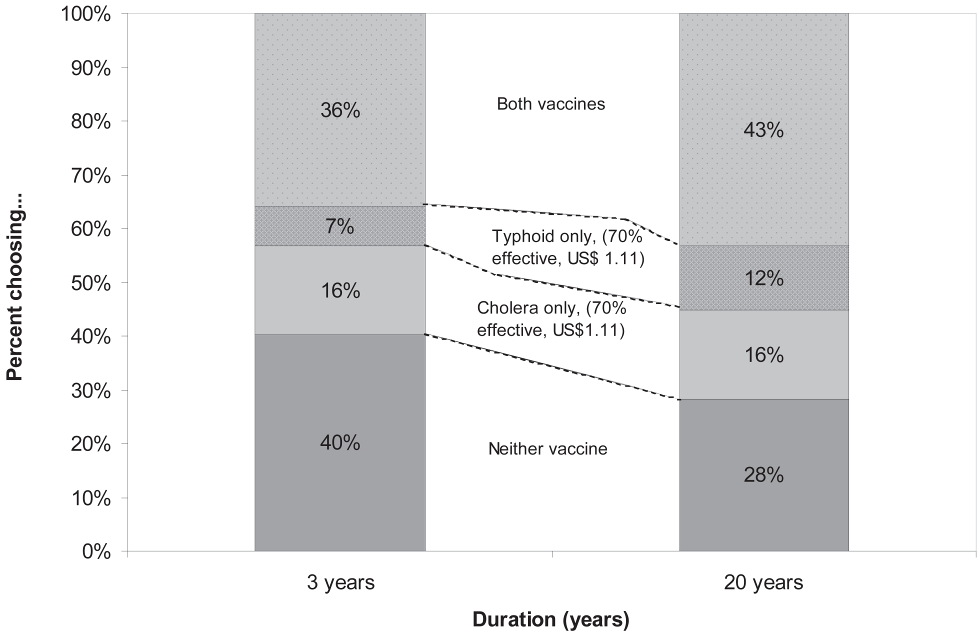
Table 2. Designing the choice tasks for "policy," or raw response, graphs.

Task #	First Group A of Respondents (N = 66)			Second Group B of Respondents (N = 67)			Third Group C of Respondents (N = 67)		
1, 2, 3, 4	*4 Tasks, Orthogonal Design, Order Determined Randomly*			*4 Tasks, Orthogonal Design, Order Determined Randomly*			*4 Tasks, Orthogonal Design, Order Determined Randomly*		
5	Task A5			Task B5			Task C5		
	Cholera	Typhoid	Neither	Cholera	Typhoid	Neither	Cholera	Typhoid	Neither
	70%	70%		70%	70%		70%	70%	
	3yr	3yr		3yr	3yr		3yr	3yr	
	US$0.22	US$0.22		US$1.11	US$1.11		US$11.1	US$11.1	
6	Task A6			Task B6			Task C6		
	Cholera	Typhoid	Neither	Cholera	Typhoid	Neither	Cholera	Typhoid	Neither
	50%	50%		70%	70%		80%	80%	
	20yr	20yr		20yr	20yr		20yr	20yr	
	US$1.11	US$1.11		US$1.11	US$1.11		US$1.11	US$1.11	



Each bar represents the answers to one choice task (approximately 66 respondents per task).

Figure 2. Percent of Respondents Choosing Either of the Two Vaccines, Both Vaccines, or Neither, When Both Vaccines Are 70% Effective for 3 Years, *by Price.*



Each bar represents the answers to one choice task (approximately 66 respondents per task).

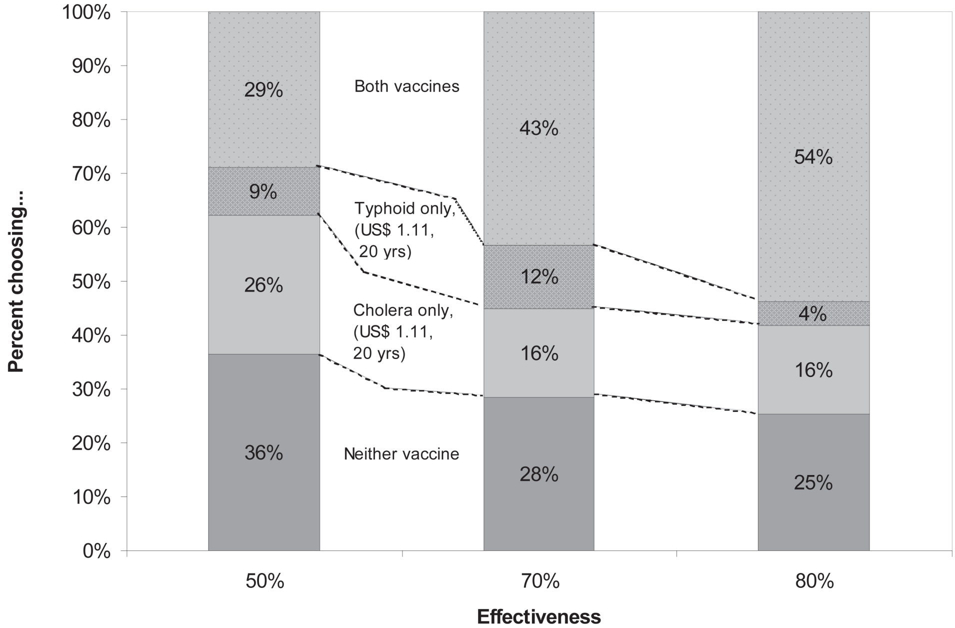
Figure 3. Percent of Respondents Choosing Either of the Two Vaccines, Both Vaccines, or Neither, When Both Vaccines Are 70% Effective and Cost USD 1.11, *by Duration.*

constant, a higher percentage of respondents chose neither vaccine, and a lower percentage of respondents chose both vaccines. A significant number of respondents were interested in buying both vaccines when the price was low (US$0.22), and 77% bought at least one vaccine at this price. Policy makers can immediately see that modest user charges could recover some immunization program costs without significantly compromising vaccination coverage rates.

Similarly, when the duration of the vaccine increased from 3 years to 20 years and all other attributes remained the same, more respondents decided to buy both vaccines and fewer chose neither, though the change in demand is much smaller than with the response to price (Figure 3). Finally, as the vaccine's effectiveness increased from 50% to 70% to 80%, more respondents chose both vaccines, and fewer chose neither vaccine (Figure 4). Demand increased more when effectiveness improved from 50% to 70% than when it improved from 70% to 80%.

Demand for cholera and typhoid vaccines does not appear to be very different, though the percentage of people choosing the cholera vaccine *only* is greater than the percentage choosing the typhoid vaccine *only* on all six policy tasks (A5, A6, B5, B6, C5, C6). It would thus be surprising if the results of an econometric analysis showed higher demand for typhoid vaccines.

Finally, we include for comparison the results of a mixed logit / hierarcharical Bayes (ML/HB) analysis of responses to all tasks, not only the policy tasks. Cook et al. (2006) use a ML/HB estimating procedure developed by Train and Sonnier (2005) that allows the attribute partworths to be correlated and modeled as transformations of the normal. Table 3 presents the results of a model where all effects-



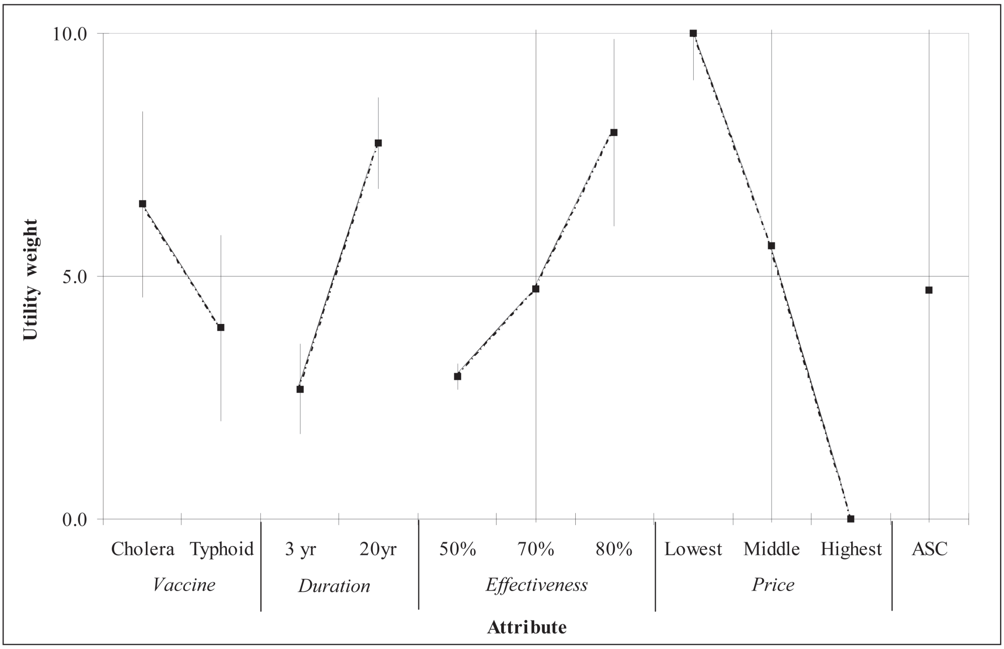
Each bar represents the answers to one choice task (approximately 66 respondents per task).

Figure 4. Percent of Respondents Choosing Either of the Two Vaccines, Both Vaccines, or Neither, When Both Vaccines Have a 3-Year Duration and Cost USD 1.11, *by Effectiveness.*

Table 3. Results from a mixed logit / hierarchical Bayes model which includes data from all cards.[a]

Attribute		Parameter (Standard Error)
Cholera vaccine	mean	0.20 (0.04)**
	standard deviation	0.05 (0.00)**
20-year duration	mean	0.39 (0.05)**
	standard deviation	0.06 (0.00)**
70% effectiveness	mean	–0.07 (0.07)
	standard deviation	0.06 (0.00)**
80% effectiveness	mean	0.42 (0.07)**
	standard deviation	0.06 (0.00)**
Lowest price	mean	0.73 (0.07)**
	standard deviation	0.06 (0.00)**
Middle price	mean	0.06 (0.07)
	standard deviation	0.06 (0.00)**
Alternative-specific coefficient	mean	–0.08 (0.08)
	standard deviation	0.53 (0.00)**

[a] All variables are effects-coded so that coefficients represent marginal effects away from a mean utility which is scaled to zero. Positive coefficients increase the probability of a choice. Variables for typhoid vaccine, 3-year duration, 50% effective, and highest price are excluded from the table but can be calculated as the negative sum of the included variables. * indicates significance at the 5% level, ** at the 1% level.



Error bars indicate one standard deviation away from the mean

Figure 5. Utility Weights of Vaccine Attributes, Derived from Mixed Logit / Hierarchical Bayes Coefficient Estimates.

coded partworths have normal distributions except the lowest price, which is modeled as a normal truncated at zero from below (the lowest cost should never decrease the probability of choosing that alternative). These results are consistent with the results from the policy graphs. They indicate that respondents a) prefer cholera to typhoid vaccines, b) are more likely to choose a vaccine with a 20-year duration than one with a 3-year duration, c) prefer vaccines with higher effectiveness, and d) are more likely to choose an alternative with the lowest price, ceteris paribus.

Figure 5 demonstrates similar findings by the plotting utility weights of each attribute. These weights are derived by recoding the mean coefficients to a common scale of 0 to 10, with 5 indicating the mean effect; coefficients above 5 increase the probability of choosing that alternative. The cost of the vaccine is the most important attribute: The lowest (highest) price has the most positive (negative) effect on utility. Changes in effectiveness and duration have approximately similar effects on utility. The graphical results from the policy tasks yield the same intuition.

DISCUSSION

Stated choice studies are growing in popularity and are increasingly used to formulate advice to policy makers. The policy graphs we present are an easy and transparent approach for presenting the results of stated choice studies that can complement the results obtained from econometric analyses. Although it is possible to construct similar graphs using simulations based on logit parameter estimates, we believe it is important to present comparisons of *raw data* to decision makers. These comparisons can also help the researcher set final design prices in pretesting, as well as help readers evaluate the underlying quality of the data. Many SC researchers will undoubtedly object that adding non-orthogonal tasks "wastes" precious sample space, and that adding more attributes or attribute levels would be a better use of a study's limited resources. While we agree that this trade-off is real and important, we believe that providing a transparent look at the underlying stated choice data will both increase interest in such studies and give policy makers a greater degree of confidence in the information collected. These benefits will often be well worth paying for.

ACKNOWLEDGMENTS

This research is part of the Diseases of the Most Impoverished Program (DOMI), administered by the International Vaccine Institute with support from the Bill and Melinda Gates Foundation. The DOMI program works to accelerate the development and introduction of new-generation vaccines against cholera, typhoid fever, and shigellosis. The program involves a number of parallel activities including epidemiological studies, social science studies, and vaccine technology transfer. The results will support public decision-making regarding immunization programs for cholera and typhoid fever. We would like to thank Brian Maskery, Malay Lahiri, and Margaret Browne for assistance with fieldwork, and Vic Adamowicz and F. Reed Johnson for insightful comments on an earlier draft. We also thank Dr. Sujit Kumar Bhattacharya, Director of the National Institute for Cholera and Enteric Diseases in Kolkata, for his help making this research possible.

DIPIKA SUR is Deputy Director, Epidemiology, National Institute of Cholera and Enteric Diseases, Kolkata, India.

JOSEPH COOK is a PhD Candidate in the Department of Environmental Sciences and Engineering, School of Public Health, University of North Carolina at Chapel Hill.

SUSMITA CHATTERJEE is Fellow cum Faculty at the Indian Institute of Bio-social Research and Development (IBRAD), Kolkata, India.

JACQUELINE DEEN is a Research Scientist at the International Vaccine Institute, Seoul, South Korea.

DALE WHITTINGTON is Professor in the Department of Environmental Sciences and Engineering, School of Public Health, University of North Carolina at Chapel Hill.

REFERENCES

Adamowicz, W. L. (2004). What's it worth? An examination of historical trends and future directions in environmental valuations. Australian Journal of Agricultural and Resource Economics, 48(3), 419–443.

Alpizar, F., & Carlsson, F. (2003). Policy implications and analysis of the determinants of travel mode choice: An application of choice experiments to metropolitan Costa Rica. Environment and Development Economics, 8, 603–619.

Arrow, K., Solow, R., Portney, P. R., Leamer, E. E., Radner, R., & Schuman, H. (1993). Report of the NOAA Panel on contingent valuation. Federal Register, 58, 4601–4614.

Carlsson, F., & Martinsson, P. (2001). Do hypothetical and actual marginal willingness to pay differ in choice experiments? Journal of Environmental Economics and Management, 41, 179–192.

Cook, J., Whittington, D., Canh, D. G., Johnson, F. R., & Nyamete, A. (forthcoming). Reliability of stated preferences for cholera and typhoid vaccines with time to think in Hue, Vietnam. Economic Inquiry.

DeShazo, J. R., & Fermo, G. (2002). Designing choice sets for stated preference methods: The effects of complexity on choice consistency. Journal of Environmental Economics and Management, 44, 123–143.

Diamond, P., & Hausman, J. (1994). Contingent valuation: Is some number better than no number? Journal of Economic Perspectives, 8(4), 45–64.

Hanemann, W. M. (1994). Valuing the environment through contingent valuation. The Journal of Economic Perspectives, 8(4), 19–43.

Johnson, F. R., Mathews, K. E., & Bingham, M. F. (2000). Evaluating welfare-theoretic consistency in multiple-response, stated-preference surveys. TER Technical Working Paper No. T-0003, Triangle Economic Research, Durham, NC.

Portney, P. (1994). The contingent valuation debate: Why economists should care. Journal of Economic Perspectives, 8(4), 3–17.

Swait, J., & Adamowicz, W. (2001). The influence of task complexity on consumer choice: A latent class model of decision strategy switching. Journal of Consumer Research, 28, 135–148.

Train, K., & Sonnier, G. (2005). Mixed logit with bounded distributions of partworths. In R. Scarpa & A. Alberini (Eds.), Applications of simulation methods in environmental resource economics (pp. 117–134). Dordrecht, Netherlands: Springer.