Environmental and Resource Economics **22:** 323–367, 2002.

Improving the Performance of Contingent Valuation Studies in Developing Countries

DALE WHITTINGTON
Departments of Environmental Sciences & Engineering, Public Policy, and City & Regional Planning, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA (e-mail: dale_whittington@unc.edu)

Abstract. This paper discusses three main reasons why so many of the contingent valuation studies conducted in developing countries are so bad. First, the contingent valuation surveys themselves are often poorly administered and executed. Second, contingent valuation scenarios are often very poorly crafted. Third, few CV studies conducted in developing countries are designed to test whether some of the key assumptions that the researcher made were the right ones, and whether the results are robust with respect to simple variations in research design and survey method. The paper concludes that research on stated preference methods in developing countries is critically important to the successful implementation of these methods because (1) there is no empirical evidence to suggest that rapid, "streamlined" CV surveys yield reliable, accurate results, and (2) there is a significant risk that the current push for cheaper, simpler CV studies could discredit the methodology itself. Moreover, the policy debates to which CV researchers are asked to contribute are often of tremendous importance to the well-being of households in developing countries. Because the costs of policy mistakes can prove tragic, it is critical that VC researchers push for excellence in this research enterprise and that funding agencies think more carefully about the value of policy-relevant information in the fields in which the contingent valuation method is being used to study household preferences and behavior (e.g., water and sanitation services, urban air pollution, soil erosion, deforestation, biodiversity, watershed management, ecosystem valuation, vaccines for the poor).

Key words: contingent valuation method, demand assessment, developing countries, household surveys, stated preferences

Introduction

Over the last ten years there has been a surge of activity in the use of the contingent valuation method (and increasingly in other stated preference techniques) to measure the value of environmental and health-related outcomes from projects, policies, and regulations in developing countries (Carson et al. 1995). I believe that the findings of many of these contingent valuation (CV) studies are inaccurate and unreliable, and that there is a pressing need to improve the quality of CV studies being conducted in developing countries. This conclusion will come as no surprise to the many critics of the contingent valuation method, who argue that stated preference methods are inherently incapable of accurately measuring indi-

viduals' preferences – in either industrialized or developing countries (Diamond and Hausman 1994; Diamond et al. 1993).

My own view is much less pessimistic than the majority of the critics of stated preference approaches. I believe that it is possible to obtain high quality, policy relevant information from well designed and soundly executed stated preference studies. However, over the past several years I have become increasingly discouraged about the quality of many of the contingent valuation studies that are being conducted in developing countries.[1] In this paper I will not offer any evidence to support my assertion that many of the contingent valuation studies conducted in developing countries are of poor quality. Moreover, I need to be clear that this assertion is not based on a systematic study of a representative sample of the contingent valuation studies that have been conducted in developing countries. Rather my conclusion is based on personal experiences in five related activities.

First, I regularly review developing country applications of the contingent valuation method (CVM) for numerous academic journals that publish such work. Second, I review proposals from young scholars in developing countries who apply to the Economy and Environment Program for Southeast Asia (EEPSEA) and the South Asia Development and Environmental Economics Program (SANDEE) for research funding to conduct such studies.[2] Third, I supervise doctoral students at the University of North Carolina at Chapel Hill and other research universities who are learning to carry out contingent valuation studies in developing countries. Fourth, I receive many unsolicited requests from researchers throughout the world who are working in this field and who want help with the design of their research or the interpretation of their findings. Fifth, I work on a more or less continual basis with the World Bank on a variety of projects and as a result see many of this organization's internal efforts to use the contingent valuation method in operational work. My conclusion that all is not well in this field is thus based on my first-hand experiences from these activities.

Based on such personal experiences, I have concluded that there are three main reasons why so many of the contingent valuation studies conducted in developing countries are so bad. The first is that the contingent valuation surveys themselves are often poorly administered and executed. Second, contingent valuation scenarios are often very poorly crafted. When one reads them, one cannot help but wonder what respondents must have thought during the interviews.

The third is that few CV studies conducted in developing countries are designed to test whether some of the key assumptions that the researcher made were the right ones, and whether the results are robust with respect to simple variations in research design and survey method. The absence of such "split-sample" tests does not necessarily mean that the findings are inaccurate or unreliable, but it does make it difficult to place much confidence in their results, or to learn much from the study about how to do better CV surveys in the future. I elaborate on each of these three problems below.

Problem 1: Poor Survey Implementation

The principal problem with many CV surveys conducted in developing countries is that the surveys themselves are poorly executed. Many of the economists directing CV studies are not themselves well trained in household survey methods or the practical issues involved in sampling in developing countries, and they thus tend to underestimate their importance. It is typically difficult for readers of journal articles and technical reports to discern the scope of survey implementation problems because researchers naturally enough do not want to highlight any difficulties they have had. Similarly editors of most journals tend to assume that CV surveys have been professionally implemented, and reviewers typically have almost no way to determine the quality of survey execution. Moreover, editors do not want authors to fill up the pages of their journal with mundane details about survey implementation.

In my opinion many of the puzzling, inconsistent results that one often finds in CV results are due to poorly trained enumerators and the resulting enumerator bias.[3] The enumerator is the critical link in the implementation of a CV survey in terms of effectively communicating a well-crafted CV scenario to respondents. Even the best CV scenario may make little sense to a respondent if a well-trained enumerator does not deliver it smoothly and sensitively. Economists will recognize this as a principal-agent problem, in which the CV researcher (the principal) typically does not know the enumerators (the agents) before the survey, and has no long-term relationship with them. Training and incentivizing enumerators is not a trivial task. It is quite difficult for researchers new to the CV field to learn more about managing and training enumerators in developing countries. I have thus included an appendix (A) to this paper that discusses some of the practical issues involved.

Problem 2: Poorly Crafted Contingent Valuation Scenarios

The Roman poet Ovid tells the story of king Midas, whose loyalty to the gods earns him one wish. Anything he asks for will be his. However, because he does not think carefully, he ends up wasting his wish. He chooses to ask that everything he touches be turned to gold. When he subsequently turns his own daughter to gold, he must ask the gods to undo the wish to bring her back to life. CV researchers in developing countries often find themselves in the place of king Midas: asking for things that they really didn't mean.

Crafting good CV scenarios turns out to be far harder than is commonly recognized. In my experience, it is the single largest hurtle to the achievement of a high-quality CV study. It is unfair to call attention to a particular researcher or study because the problems in crafting CV scenarios are so ubiquitous. It is also difficult for policy analysts to appreciate the widespread scope of this problem because many published papers do not report the precise CV scenario(s) or the valuation question(s) that were presented to respondents.[4]

Crafting a good CV scenario amounts to writing a short story about the problem or situation that is the focus of the survey, and then posing an interesting choice (or decision) for the respondent. A useful way to think about a CV scenario is that the researcher must describe a hypothetical "deal" to the respondent, and then the respondent can either accept or reject this deal. Putting together a sound deal structure is difficult in most fields; nonmarket valuation is no different.

If the researcher is to learn something from the respondent's acceptance or rejection of this deal, the respondent must be (1) sufficiently intrigued by the story to listen closely to the aspects of the deal being described, and (2) able to understand the characteristics of the deal *as the CV researcher intends*. This requires that a balance must be struck between not insulting the respondent's intelligence and not overestimating what a respondent knows about a subject. Because exactly the same CV scenario must be delivered to each respondent in a survey, finding the right balance in a CV scenario between intelligibility and sufficient detail is often difficult.[5] It requires that the CV researcher utilize focus groups and carefully administered pretests to learn how different groups in a population are responding to the CV scenario, i.e., whether they are engaged by the story, understand the deal proposed, and provide answers that are influenced by the price of the hypothetical good or service.

Crafting a good CV scenario, however, requires more that simply judicious use of focus groups and pretests, indispensable as these tools are. The CV researcher must be able to put herself in the place of a respondent and understand how a respondent would consider the different facets of the hypothetical deal that she is being asked to evaluate. In my experience there is a curious irony involved in this task.

Critics of the contingent valuation method often argue that there is no economic content to respondents' answers to questions posed in CV surveys because respondents do not face real economic choices (or an actual budget constraint). The problem with many CV scenarios is not, however, that respondents have trouble thinking in economic terms about hypothetical choices. Rather it is that the CV researcher pays too little attention to the economy theory of household decision making when crafting the CV scenario, i.e., the CV researchers themselves cannot construct hypothetical choices that make economic sense to respondents.

Some CV researchers working in developing countries seem to have a particularly difficult problem understanding the linkage between the CV scenario and the selection of an elicitation procedure. It is not widely understood that the choice of elicitation procedure can depend on local culture and customs. Nor in my experience is it widely understood among CV researchers working in developing countries how the choice of elicitation procedure depends on the extent to which the hypothetical good or service described in the CV scenario can be characterized as a public good (i.e., consumption by one person does not reduce the amount of the hypothetical good or service available for consumption by others). There are several common mistakes.

Table I. The linkage between the CV scenario and the choice of elicitation procedure

Type of elicitation procedure	CV scenario describes a hypothetical **private** good or service	CV scenario describes a hypothetical **public** good or service
Open-ended maximum WTP valuation question	Case I	Case II
Closed end, Yes/No valuation question	Case III	Case IV

Table I shows the relationship between two types of elicitation procedures (an open-ended maximum WTP question and a closed-end, discrete Yes/No valuation question) and two types of hypothetical goods and services (private and public). For each of the four situations (Cases I–IV in Table I) I will discuss where CV researchers often go wrong and a recommended approach. To make the discussion more concrete, I will use the example of a CV scenario and elicitation procedure for a hypothetical HIV/AIDS vaccine. To keep the example simple, assume that the HIV/AIDS vaccine is 100% safe and effective and lasts for an individual's entire lifetime.

One could craft a CV scenario that offered a HIV/AIDS vaccine to an individual and/or his family members (Whittington et al. 2000). In this case the hypothetical product would be a private good. Although consumption of this private good would have some positive externalities that could conceivably enter into the individual's decision as to whether or not to get vaccinated, most respondents would probably base their decision on whether to purchase a vaccine largely on self interest and concern for their own family members.

A different CV scenario involving a hypothetical HIV/AIDS vaccine could describe a government-financed mass vaccination campaign – or a campaign to target high-risk individuals and/or anyone who wanted a vaccine. Such a mass vaccination campaign would have many of the characteristics of a public good because it would not only eliminate any risks to the respondent and his family of being infected, but also would eliminate new infections in the entire population. A respondent might be willing to pay more for such a program than for a vaccine for himself due to altruistic reasons or the macroeconomic benefits of eliminating the disease.

Case I in Table I combines a private good and an open-ended maximum WTP valuation question. In this example the valuation question would entail asking the respondent the most he would be willing to pay for a hypothetical HIV/AIDS vaccine for himself. This valuation question would make sense for our hypothetical HIV/AIDS vaccine because a person would only need one vaccine. In general, a maximum WTP question for a private good or service requires that the quantity of the good be fixed (in this case at one vaccine). If we changed the CV scenario

to ask what is the respondent's maximum WTP for vaccines for his household members, we would need to know how many vaccines he would buy. In Case I a common mistake CV researchers make is to ask for the respondent's maximum WTP without clearly specifying in the CV scenario the quantity of the good or service that the respondent is to assume.

In most situations involving private goods and services, what the researcher would like to know is the respondent's demand curve, so there is no obvious *single* quantity to propose. One approach for estimating an individual's (or household's) demand curve is to ask the respondent the maximum he would be willing to pay for a series of different quantities. However, this is typically a difficult line of questioning because individuals may never have carefully considered how they would actually use different amounts of a private good or service that they may have never used at all before.

Case II combines a public good and an open-ended WTP valuation question. This combination is typically a mistake. The problem is that a respondent needs to know what others are going to pay for the public good before he can determine what he would be willing to pay. In most cultures the issue is one of perceived fairness: a respondent may logically reason that all individuals in similar circumstances should share the costs of the public good and pay the same. It would not be fair for the respondent to have to pay his maximum if others did not have to pay the same amount (since they would also receive the full benefit from the provision of the public good).

For example, if the CV scenario described a government-sponsored mass vaccination campaign against HIV/AIDS, it would not be advisable to ask an individual the maximum he would be willing to pay for such a program because others would also need to contribute to finance the program.[6] The respondent might understandably be reluctant to contribute his maximum WTP if he were not assured that others in like circumstances would do the same. A related problem is that there is no politically realistic payment vehicle for such a scenario that could identify each respondent's real preferences for such a program and extract each respondent's maximum willingness to pay. The Case II CV scenario and open-ended valuation question will thus seem *very* hypothetical to the respondent – and not a sensible economic or political choice.

Case III combines a closed-end, Yes/No, discrete choice valuation question with a private good. In this case the CV scenario describes the hypothetical good or service, and the respondent is then asked whether or not he would buy it at a specified price. For most private goods and services, the elicitation procedure should allow the respondent to specify the quantity of the good or service he would buy at the specified price. For example, a respondent might be asked how many HIV/AIDS vaccines he would buy for his family if the price was US$30 per vaccine (and perhaps which family members would receive them). In this case, the respondent's answer yields one point on the household's demand curve for HIV/AIDS vaccines. As in Case I, a common mistake in Case III is for the CV

researcher to forget the basic economic point that the research objective is to obtain a demand curve and thus fail to ask the respondent to clearly specify the quantity to be purchased at the stated price.

Case IV combines a CV scenario for a hypothetical public good with a closed-end, discrete choice valuation question. The closed-end, Yes/No valuation question is typically a desirable approach to valuing a hypothetical public good (Mitchell and Carson 1989). The CV scenario often includes a description of a government program or management plan that will provide the public good or service if a majority of respondents votes for the plan and agrees to pay the specified amount. For example, the enumerator would present the CV scenario that described a government-sponsored mass vaccination program to the respondent, and then ask whether he would vote for such a program if it would cost his household US$50 per year in increased taxes. A common mistake here is to use an open-ended maximum WTP question to determine the range of prices to use for the closed-end, discrete choice valuation question(s). The pretest should be done with the CV scenario and the exact valuation questions used in the final survey.[7] The use of an open-ended maximum WTP question for pretesting will result in the same incentive-compatibility problems described in Case II above.

In reality the political factors involved in the provision of many government-supplied goods and services make crafting a CV scenario and valuation question(s) more complicated than implied by the simple typology presented in Table I. The economic and political issues involved in the provision of government programs are almost always intertwined. To engage the respondent and to capture the political reality of a local situation, CV researchers often need to have a carefully nuanced sense of the political forces at play before trying to draft a CV scenario. It is not uncommon for individual respondents to be concerned with both a public choice at a policy level and a private, household decision as to how to respond personally should a macro policy change occur. Typically policy makers are interested in households' preferences about both choices. In the first, public choice an individual may have both self interested and altruistic reasons for favoring one policy direction over another. In the second, personal decision considerations of self-interest are usually assumed to predominate.

To illustrate the complexity that this kind of situation can create for the task of crafting a CV scenario, I will use an example from Kathmandu, Nepal. In the spring of 2001, Subhrendu Pattanayak, Bal Kumar, and myself designed and carried out a CV study in Kathmandu to determine households' willingness to pay for improved water services. We designed the CV survey to gauge households' reactions to a possible plan to engage a private sector operator in order to improve several attributes of the service provided by the existing piped distribution system. Because households in the Kathmandu Valley are currently obtaining their water from a variety of different sources and have different housing arrangements, it did not make sense to ask all households in the sample precisely the same questions about whether they would support a plan to improve water services, and what they would

do if a new, improved water service were available. We thus designed different versions of the CV questionnaire for various groups of households.

There are two main groups of households in the Kathmandu Valley: (1) households with connections to the existing National Water Supply Corporation (NWSC) piped distribution system; and (2) households without connections to the existing NWSC distribution system. Each group received a different version of the survey instrument. Respondents were told about a plan that would engage the private sector to improve the water supply system. They were asked to suppose that the improved system would provide 24-hour service, that water would be safe to drink from the tap, and that the private operator would provide accurate billing of the water they received. Households with connections to the existing distribution system were then told . . .

> I want you to suppose that the improved water service for households in the Kathmandu Valley with a private NWSC connection would result in a total monthly water bill for a typical household like yours of (200/400/600/800/1000/1300/1600/2000 Nepalese rupees).[8] Let's assume that a water bill of this size would entitle a typical household to about 500 liters of water per day.[9]

Respondents were then asked whether they would vote for the water supply improvement plan. The eight different amounts of monthly water bills were randomly assigned to subsamples of respondents with private connections, i.e., some sample respondents received one monthly water bill, and other subsamples received different monthly water bills.

Next, respondents were told . . .

> Now, I want you to suppose that in fact most people did vote for the plan to improve the water supply system. Assume that the typical household's monthly water bill for 500 liters of water per day increased to [200/400/600/800/1000/1300/1600/2000 NPR]. What do you think your household would do?
>
> (1) Stay connected and pay the higher water bill
> (2) Disconnect and find water elsewhere
> (3) Don't Know

Households currently without connections were also asked how they would vote on the plan, but they were told that they could choose to have either a private or a shared connection with improved water service. Respondents were told to assume that the monthly cost of a shared connection would be half the cost of a private connection.

These two valuation questions actually pose complex choices and require careful consideration on the part of the respondent. The first question asks about the respondent's willingness to support (vote for) a public action (in this case the involvement of a private sector operator in the implementation of service improvements); the second question asks about private behavior (what the respondent's

Table II. The decision problems created for the respondent by the two CV questions

Current situation (status quo): Choice set 1	*First CV Question*: Would you vote to have Choice set 1 or Choice set 2?	*Second CV Question*: Suppose you had to choose between the options in Choice Set 2 (i.e., the majority voted for the plan). What would you do?
Connection to NWSC connection with poor service and low water bill	Choice set 1 (status quo):	Connect to improved private water system
vs.	*vs.*	*vs.*
Not connect; rely on public taps, private wells, and other sources	Choice set 2: Connect to improved private water system at specified price vs. not connect and rely on public taps, private wells, and other sources	Rely on public taps, private wells, and other sources
Comment: We know the decision the respondent's household made regarding this choice.	*Comment*: Respondent tells the enumerator which choice set he/she prefers by indicating whether he/she would vote for the plan	*Comment*: Respondent tells the enumerator whether his/her household would connect to the improved water system if having a connection with a low water bill and low quality was no longer an option

household would actually do if confronted with a choice to connect to the new system). We felt that it was necessary to ask both questions because they mirror the political economy of water supply improvements in Kathmandu.

The *status quo* decision problem for most households in the Kathmandu Valley is a choice between (1) a connection to the distribution system that provides poor service at a low cost, and (2) relying on water sources other than a private connection. The two CV questions effectively create two new decision problems for the respondent. As depicted in Table II, the first CV question asks the respondent to choose between *two choice sets*:

(1) Status quo [no plan]: (a connection to the distribution system that provides poor service at a low cost) vs. (relying on water sources other than a private connection)

(2) Change [plan implemented]: (a connection to an improved distribution system that provides good service at higher cost) vs. (relying on water sources other than a private connection). The first CV question does not ask the respondent to indicate what he would actually do, only from which of the two choice sets he would prefer to choose.

For a household without a private connection, the main reason not to vote for the plan (i.e., answering YES to this first CV question) is that the option of connecting to the existing system in the future is removed.[10] Presumably this is not likely to be perceived as a large cost because the respondent's household has the option of connecting to the existing system now and decided not to do so.[11] On the other hand, a respondent with a private connection has a strong incentive to vote NO if he prefers the existing system to the improved system because the implementation of the plan would remove the option of having a connection to the existing system.

The second CV question restricts the respondent's options by asking him to imagine that the plan to improve water services was indeed approved and implemented, and that the first choice set above was removed from consideration, i.e., his household no longer had the option of having a connection that provided poor service at a low price. This restriction of the respondent's options to the second choice set could reduce his household's welfare, even if he indicated that his household would stayed connected and pay the higher price for the improved service.

Table III shows the six alternative ways a respondent might rank three possible "states of the world": (A) connected to a new private system with high quality and a higher monthly bill; (B) connected to existing system with low quality and a low monthly bill, and (C) disconnected from piped distribution system, rely on other sources. What can a respondent's answers to the two CV questions and his choice of water sources in the current situation tell us about his household's preferences?

Consider first a household that already has a private connection. We can rule out rankings 2, 5, and 6 because in these three rankings being disconnected from the existing system (C) is preferred to having a private connection to the existing system (B), but in fact we know the household chose B over C. What can the answers to the two CV questions tell us about the respondent's preferences among the remaining rankings (1, 3, and 4)?

The two CV questions yield four possible patterns of responses (Table IV). For example, consider a household that currently has a connection to the piped distribution system. Two of the possibilities are straightforward. First, such a respondent could vote for the plan and then agree to stay connected to the distribution system at the new, higher tariff (Pattern 1 in Table IV). We interpreted a "yes/yes" response to the two CV questions as indicating that the respondent holds ranking 1, not ranking 3 or 4. That is, if the respondent voted for the plan and said his household would connect to the new system at the offered monthly bill, we assume he prefers a connection to the new system at the specified monthly bill to his connection to the old system, and experiences a welfare gain from the implementation of the plan and paying the higher tariff.

Second, a respondent could vote against the plan, and, if the plan were implemented, decide to disconnect from the piped distribution system (Pattern 4). We interpreted such a "No/No" response to the two CV questions as indicating that the respondent holds ranking 4, not ranking 1 or 3. That is, if the respondent (whose

Table III. Possible rankings of household water supply "states of the world" (most preferred to less preferred)

Ranking 1	Ranking 2	Ranking 3	Ranking 4	Ranking 5	Ranking 6
A. Connected to new private, system at specified price (high quality, high prices)	A. Connected to new private, system at specified price (high quality, high prices)	B. Connected to existing system (low quality, low prices)	B. Connected to existing system (low quality, low prices)	C. Disconnected; rely on other sources	C. Disconnected; rely on other sources
B. Connected to existing system (low quality, low prices)	C. Disconnected; rely on other sources	A. Connected to new private, system at specified price (high quality, high prices)	C. Disconnected; rely on other sources	A. Connected to new private, system at specified price (high quality, high prices)	B. Connected to existing system (low quality, low prices)
C. Disconnected; rely on other sources	B. Connected to existing system (low quality, low prices)	C. Disconnected; rely on other sources	A. Connected to new private, system at specified price (high quality, high prices)	B. Connected to existing system (low quality, low prices)	A. Connected to new private, system at specified price (high quality, high prices)

Table IV. Possible responses to the two CV questions

Answer to 1st CV Question:	Answer to 2nd CV Question (Connect to improved system?): YES	Answer to 2nd CV Question (Connect to improved system?): NO
Vote for the plan? YES	Response Pattern 1: Yes/Yes (Vote for the plan and connect to the new system)	Response Pattern 2: Yes/No (Vote for the plan, but not connect to the new system)
Vote for the plan? NO	Response Pattern 3: No/Yes (Vote against the plan, but connect to the new system)	Response Pattern 4: No/No (Vote against the plan and not connect to the new system)

household is already connected to the existing system) voted against the plan and said that his household would disconnect from the existing system if the plan were implemented, we assume that he prefers a connection to the existing system to the new system, and would experience a welfare lose if the plan were implemented and his household choose to disconnect from the system.

Third, a respondent could vote against the plan, but then indicate that if the plan were implemented, her household would decide to stay connected to the new system (Pattern 3). In other words, she preferred the status quo situation, but if this option were removed from her choice set, her household would choose the new system and higher tariff rather than disconnect from the piped distribution system. We interpret such a "No/Yes" response to the two CV questions as indicating that the respondent holds ranking 3, not ranking 1 or 4. That is, if the respondent (whose household is already connected to the existing system) voted against the plan, but said that her household would stay connected to the distribution system if the plan were implemented, we assume that she prefers a connection to the existing system to a connection to the new system, but prefers a connection to the new system to being disconnected. In this instance the household would be made worse off as a result of the implementation of the plan for improved water services, even though it would decide to remain connected and pay the higher price.

Fourth, a respondent could vote "Yes" for the plan and then decide not to stay connected to the piped distribution system (Pattern 2). A respondent might value the option of connecting to an improved distribution system in the future, but know that his household could not afford it now. Conceivably a respondent could vote for the plan for altruistic reasons (e.g., because he believed that it would improve the health of his neighbors and school-age children), knowing that his household could not afford the new water service. We interpreted a "Yes/No" response to the two CV questions as indicating that the respondent holds ranking 1, not ranking 3 or 4.

For households that are not currently connected to the distribution system, we can rule out rankings 1, 3, and 4 because in these three rankings B is preferred to C, but in fact we know the household chose C over B. We again attempt to interpret the four possible patterns of responses to the two CV questions, but in this case a household that votes for the plan may prefer either a private connection or a shared connection to the new system. If a respondent without a private connection votes for the plan and then says that his household would connect to the improved system (a "Yes/Yes" response), we interpret this as an indication of ranking 2. We interpret a "No/No" response as an indication of ranking 5 or 6 (we cannot determine how the unconnected household ranks a connection to the new system versus a connection to the old system).

A "No/Yes" response for an unconnected household would seem to be an inconsistent; it is, however, conceivable that the respondent is worried about the implications of the plan for others (perhaps the poor) and votes against it. But once the plan is passed, the respondent is free to indicate his personal preference to

connect his household to the new system (either a private or shared connection. Nevertheless we interpret a "No/Yes" response pattern as an inconsistent answer.

We interpreted a "Yes/No" response for an unconnected household as an indication that the respondent holds ranking 5. We interpret a "YES" vote to the first CV question as an indication that the respondent prefers a connection to the new system to a connection to the old system, ruling out ranking 6. A "No" vote to the second CV question indicates that the respondent prefers being disconnected to having a private connection to the new system. In this case being connected to the new system is preferred to being connected to the existing system, and the "Yes" vote to the first CV question might indicate a willing on the part of the respondent to pay for having this option in the future.

Elsewhere we have presented the results of the analysis of these data from Kathmandu (Pattanayak et al. 2001). The point of presenting the framework of this CV scenario here in some detail is to illustrate that crafting CV scenarios that more accurately capture the linkages between a respondent's public and private decisions can create conceptually complex choices for the respondent. Too often CV researchers give insufficient thought to local political realities, and, as a result, their CV scenarios come across to respondents as contrived or out of touch.

Problem 3: Failure to Test for the Effects of Variations in Survey Design on the Results of Contingent Valuation Studies

The scholarly literature on the contingent valuation method stresses the importance of carrying out a variety of split-sample experiments in order to better understand how respondents may be reacting to the CV scenario and the elicitation procedure (e.g., Carson et al. 1997, 1998; Lauria et al. 1999; Whittington et al. 1992, 1998). There are two main reasons for conducting split-sample experiments in CV research designs. The first is that, as described above, the CV researcher almost always faces some difficult choices in the study design with respect to the crafting of the CV scenario and elicitation procedure. The CV researcher can utilize split-sample experiments to gain insight into the question of whether she made the "right" design choice with respect to a few critical issues. For example, the CV researcher can learn whether respondents' answers can be manipulated by changes in survey design or the information presented in the CV scenario in ways that a "reasonable observer" might anticipate. A failure of such experiments to elicit the response that is anticipated raises doubts about the confidence one can place in the CV results and the research design choices that were made. In other words, a different research choice (e.g., elicitation procedure) should elicit different answers from respondents, and the failure to induce different responses may suggest that people did not see the answers to the CV questions as having real consequences.

Second, conducting a CV survey provides the funding agency an opportunity to learn more about what works and does not work in terms of alternative research design choices in a particular culture. Funding agencies need to know whether

one approach to conducting CV surveys is more likely to yield accurate, reliable answers and the cost implications of different research designs. The use of split-sample experiments provides the funding agency the opportunity to gather information that can benefit its future operational work.

These kind of split-sample experiments have not been widely used in CV studies in developing countries. It is thus difficult for readers of such studies to assess the robustness of the results. The failure to conduct split-sample experiments also represents a lost opportunity to learn more about how to improve CV studies in developing countries. Given the inherent difficulties of understanding respondent behavior in different cultures, spilt-sample experiments should be standard practice in this field. I will illustrate the importance of carrying out such tests with two examples, one from a CV study conducted in Tanzania, and the other from a CV study in Odessa, Ukraine.

WILLINGNESS TO PAY FOR IMPROVED SOCIAL SERVICES IN TANZANIA

In 1993 the World Bank funded a large, ambitious household survey in Tanzania (Human Resource Development Survey – Tanzania) designed to better understand households' expenditures on social services and households' preferences and willingness to pay for a variety of social and health programs. Individuals in over 5000 randomly selected households were interviewed throughout Tanzania. The survey instrument was 85 pages long, and each interview took over 2 hours to complete. This survey work was organized and directed by Dr. Charles Griffin, the Project Director and a Senior Health Economist at the World Bank.

There was actually more one than one round of this Human Resource Development Survey, and Dr. Griffin and his colleagues experimented with several slightly different contingent valuation scenarios. In all the rounds each survey instrument included several CV scenarios. For example, one CV scenario dealt with the respondent's willingness to pay for a health insurance card that would provide limited coverage of heath care costs; another elicited willingness to pay information on a girls' secondary education program (see Appendix B). Others elicited willingness to pay information on a community health clinic, child spacing, and a hypothetical malaria vaccine.

A split-sample experiment was incorporated in the research design; it entailed the use of three different bidding games to elicit the respondents' willingness to pay for each proposed government program. Figures 1A–C present the sequence of questions in each of the three bidding games used for one of the CV scenarios (health insurance card). One third of the respondents in the sample received a bidding game with a sequence of questions that started with a low price of 1000 Tsh; another third of the sample received a second bidding game that started with a medium price of 25,000 Tsh, and the final third of the sample received a third bidding game that started with a price of 50,000 Tsh.[12] These three bidding games were randomly assigned to respondents in the study.

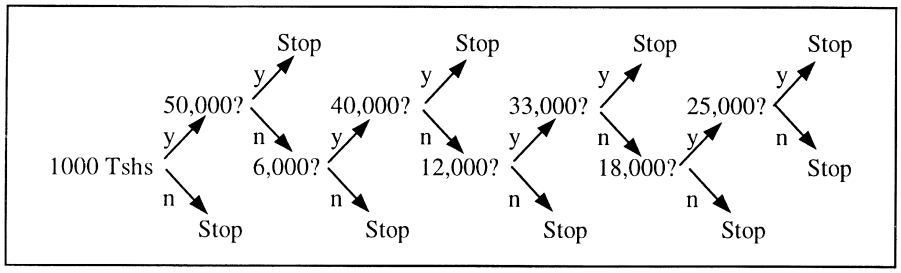

Figure 1A. Bidding game with low starting point – Human Resources Development Survey (Tanzania) – offered prices in Tsh (US$1 = 550 Tsh).

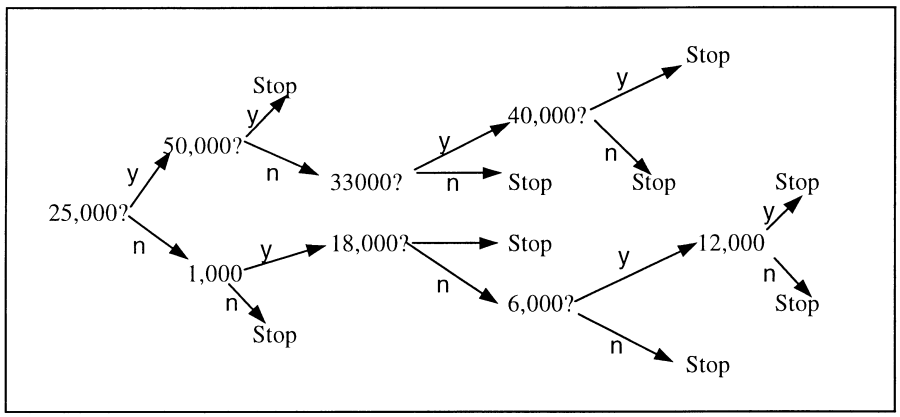

Figure 1B. Bidding game with medium starting point – Human Resources Development Survey (Tanzania).

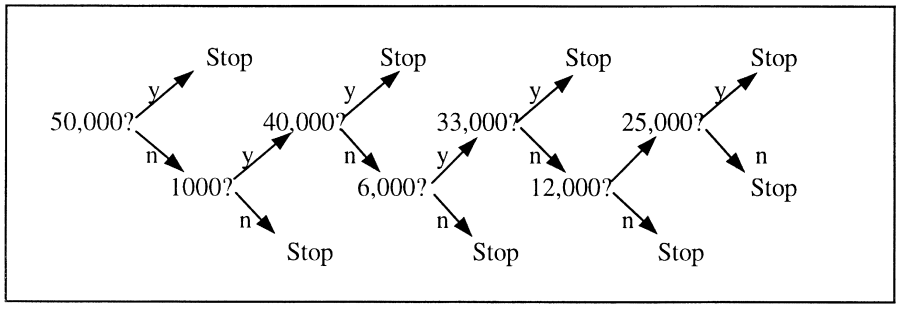

Figure 1C. Bidding game with high starting point – Human Resources Development Survey (Tanzania).

Table V. Number of questions respondent receives in three bidding games versus respondent's willingness to pay (Human Resources Development Survey – Tanzania)

Number of Yes/No questions that the respondent receives in . . .			
Respondent's willingness to pay:	Version A	Version B	Version C
< 100 Tshs per visit	1	2	3
500 Tshs per visit	3	4	4
3500 Tshs per visit	7	4	7
5000 Tshs per visit	8	3	7
9000 Tshs per visit	6	4	5
15000 Tshs per visit	4	4	3
> 16000 Tshs per visit	2	2	1

The purpose of this split-sample experiment was to test whether respondents' willingness to pay would be influenced by the magnitude of the first price that they received and the sequence of follow-up questions. There are two different perspectives on such a "starting point" test. One is that the different starting points convey information about the cost of the good or service provided. From this perspective different starting points *should* induce different answers from respondents. In other words, if the split-sample test elicited different answers, one would conclude that respondents were in fact taking the CV scenario seriously.

A second perspective is that a respondent holds a precise willingness-to-pay amount in his mind that he is trying to accurately communicate to the interviewer. Respondents would have to answer very different numbers of questions depending on which of the three bidding games he received (Table V). In effect, a honest respondent who wanted to "tell the truth" about his willingness to pay was sent through a maze of questions that differed greatly depending on which of the three bidding games he was randomly assigned. There are few CV researchers or social psychologists who would predict that poor, often illiterate respondents in Tanzania could find their way through such a maze during the course of such a long interview. In this interpretation, if respondents who received different starting points provided essentially the same WTP answers, one would have greater confidence that they were revealing their "true" WTP.

When this Tanzania Human Resource Development Survey was in the planning stage, Dr. Griffin asked me for help on the design of the CV components of the survey instrument. I felt strongly that the survey instrument was too long, and urged him to shorten it. I also pressed him not to use such a stringent split-sample test for starting point bias because I was sure that respondents would be too heavily influenced by the huge differences in the three starting points. Fortunately, Dr. Griffin ignored my advice on both points.

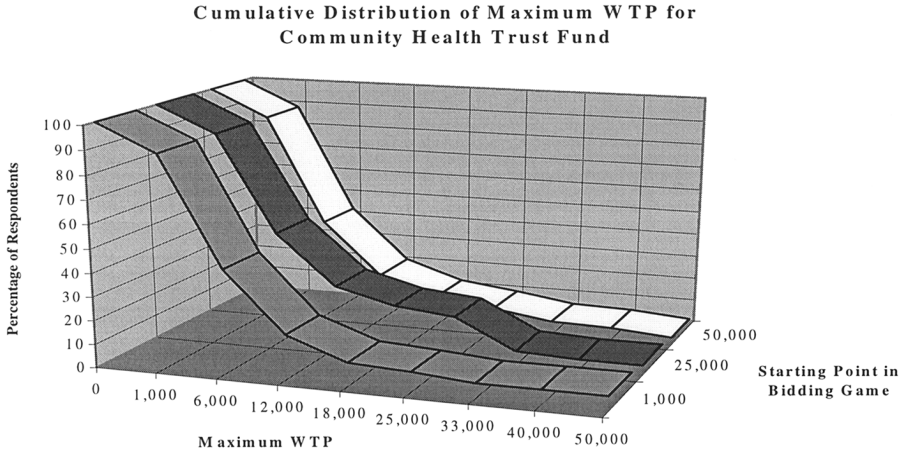

Figure 2A. Frequency distribution of WTP responses for three Bidding games – community health trust scenario (Human Resource Development Survey – Tanzania).

Figures 2A–B present the frequency distributions of willingness-to-pay responses for the three bidding games for two of the CV scenarios included in the survey instrument, based on respondents' final answer.[13] As shown, the results for the three bidding games look quite similar. There is, however, a "starting point" effect in respondents' answers to the first question in the bidding game with the middle starting price (e.g., 25,000 Tsh for the health insurance card; Figure 2A). More respondents appear to have answered "yes" to the starting price in this middle bidding game (i.e. more respondents said that they were willing to pay this price than the other two bidding games would suggest). Is this, however, a "problem"?

This "kink" in the frequency distribution for the middle starting point occurs in both CV scenarios, but it is larger in the WTP distribution for the Girls Scholarship Fund. This could be interpreted as lending support to the hypothesis that the initial starting price of 25,000 Tsh is conveying information about the perceived cost of sending a girl to secondary school. Respondents are probably less likely to perceive what the cost of the community health trust fund would be, and thus are less sensitive to the starting point used.

If one concludes that the starting point effect with 25,000 Tsh is due to "yea saying," then the correct approach for the policy analyst would be to use the results from the 1,000 Tsh or 50,000 Tsh starting points. Alternatively, if the responses contingent on the 25,000 Tsh starting point are true, then using the results from the 1,000 Tsh or 50,000 Tsh starting points would be conservative. If the policy decision is the same in either case, then the choice of starting point does not influence the choice of policy.

In fact, Dr. Griffin concluded that for policy purposes the results were in fact sufficiently close to treat them as indistinguishable from one another. He also felt that in general the respondents in the survey must have made an extraordinary effort

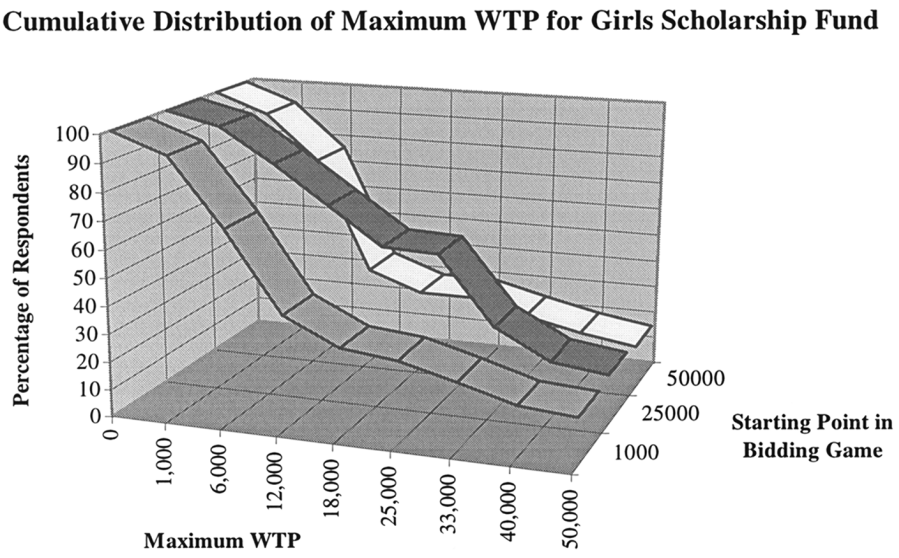

Figure 2B. Frequency distribution of WTP responses for three bidding games – girls' secondary education program (Human Resource Development Survey – Tanzania).

to pay attention to the questions they were being asked and to try to the truth about their willingness to pay. In effect, the respondents found their way through the three mazes that Dr. Griffin had created and all three groups of respondents found themselves in essentially the same place at the end. An important outcome of the split-sample experiment was that Dr. Griffin had more confidence in the accuracy and reliability of the CV results.

WILLINGNESS TO PAY FOR IMPROVED MUNICIPAL WATER SUPPLY SERVICES IN ODESSA, UKRAINE

In 1996 Jennifer Davis, Kristin Komives, Dafina Gercheva, and I designed and supervised the implementation of a contingent valuation study in Odessa, Ukraine for the World Bank (Davis et al. 1996). The objective of this CV study was to measure households' willingness to pay for improved water services in Odessa. This CV survey included a split-sample experiment to check the robustness of respondents' answers to valuation questions to changes in the type (or mode) of interviews administered to respondents.

It is generally accepted that some modes of survey administration are preferable to others if a lot of information needs to be conveyed to respondents or if the information is itself complex. However, CV researchers are often under considerable pressure to reduce the costs of survey administration. Since there is quite limited information about the magnitude of such mode effects in different cultures, we chose to use this project as an opportunity to learn more about the advisability of such cost-cutting measures.

By way of background, most households in Odessa live in multifamily housing and have in-house piped water services. A typical apartment is served by multiple water lines, and none are metered. However, the quality and reliability of their piped water service is currently poor. Few households have water supplied to their households or flats 24 hours a day. Most households are quite concerned about the contamination of the water from piped connections and the vast majority of the population treats its drinking water. On the other hand, this poor water service is very cheap. A small fixed charge for water service is included in most households' combined monthly bill for rent and infrastructure services.

In the CV scenario we told respondents about the characteristics of a modern piped water supply system, i.e., the water was available 24-hours a day, 365 days a year on demand, that the water from the tap was potable without further treatment, and that connections were metered and customers were billed depending on the quantity of water they used. The enumerator then told the respondent:

> Suppose your family would be able to have the type of water service I have just described. The cold water[14] would be available 24 hours a day, 365 days a year, with sufficient pressure. It would be safe to drink directly from the tap without further treatment. In your apartment/house, water meters would be installed. You would only pay for the water your family actually used.
>
> Suppose that your family could have this type of modern water service. Of course, this higher level of service would be more expensive than your current service. Because you would pay for the amount of water your household actually used, your water bill might vary each month. But, suppose that in an average month the cost of cold water for each member of your household would be _____ Krb. That means that your family, which has _____ members, would have to pay a total of about _____ Krb per month for cold water.
>
> What would you choose? To keep the cold water supply services you are using now and pay the same amount you are paying now, or pay approximately _____ each month for your whole family for improved water service?[15]

Five different average per-person monthly costs were used in this scenario: 200,000 Krb. (US$1.08), 300,000 Krb. (US$1.62), 500,000 Krb. (US$2.70), 700,000 Krb (US$3.78), and 1,000,000 Krb. (US$5.41). One of these five per-person monthly costs was randomly assigned to each respondent.[16]

What was unusual about the research design was that we used four different types of interviews. The total sample size was 2259 households. Of these 767 individuals were interviewed by telephone; 737 were interviewed in-person in their homes; 504 were interviewed in-person outside their homes (intercept survey); and 251 were interviewed during the course of small participatory meetings in their neighborhoods. In the group meetings, it was not possible to randomly assign different prices to different participants. One price was announced to all participants in a meeting (and we used the same price in all the meetings). Participants then discussed the pros and cons of the proposed change in water service.

The comparison of individuals' WTP responses across these four groups is not straightforward because of differences in the socioeconomic characteristics of the resulting samples, despite that fact that considerable effort was made to obtain a representative sample in each of the four groups of respondents (Davis 1998; Davis and Whittington 1998).[17] Nevertheless, the frequency distributions of WTP responses for the four groups are quite striking (Figure 3). For the three groups that received the five different randomly assigned per-person monthly costs (i.e., respondents in the household, intercept, and telephone surveys), the percentage of respondents who agreed to pay for the improved, modern water service declined as the monthly cost increased. From this one can infer that respondents were in fact listening to the CV scenario and that their answers depend on the price they were told in a manner consistent with expectations. However, the frequency distribution for respondents in the telephone survey is quite different from the distributions for the household and intercepts survey groups.[18] In my opinion, this result lends support to the conventional wisdom in the CV research community that one should avoid using telephone surveys even though they are less expensive to implement and can be completed more quickly than in-person interviews.

Of perhaps even more interest are the results for the participatory meetings because such data collection methods have widespread support in the development planning business (Whittington and Davis 1994). Figure 3 shows two data points for the percentage of individuals who agreed to pay the per-person monthly charge of US$2.70: (1) Focus Group – Public Vote; and (2) Focus Group Revised (Private Response). The first depicts the percentage of respondents who publicly stated, after an extensive group discussion, that they would agree to pay the offered price. After this public vote, respondents were offered an opportunity to revise their public decision and report it confidentially (see Davis 1998, for details of this experiment). The second data point (Focus Group Revised) shows the percentage of respondents who agreed to pay the offered price in this private, confidential answer. As shown in Figure 3, substantially fewer respondents stated that they would accept the new improved water service when they were offered an opportunity to report a confidential, private answer. Forty one percent of participants in the participatory meetings actually changed their answers when offered the opportunity of confidentially.

Not only were many respondents' revised, private answers different from their public answers, but also participants revised their answers in all different directions. As shown in Figure 4, 15 percent of the participants changed their public answer of accepting the improved source to refusing to pay for the new source, and 6 percent who accepted the improved source publicly revised their answer to "Don't know" in their private answer. On the other hand, 9 percent in public said that they would not accept the improved water service (or did not know if they would), but then revised their answers in private and said they would accept the service. The percentage of respondents in the group meetings who indicated that they would pay the price of the improved water system in their revised answers

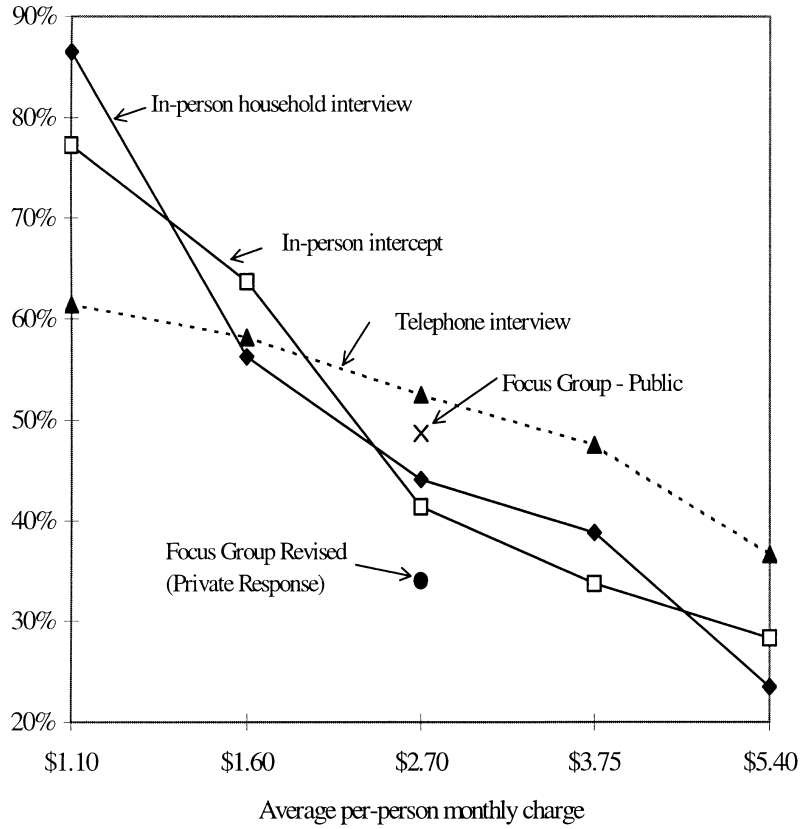

Figure 3. Percent of respondents willing to pay for improved water service vs. per-person monthly cost for different types of interviews [Source: Davis 1998].

is only about half the percentage of respondents in the telephone interview who agreed to pay. The policy conclusions one would reach from this study would be quite different depending upon which type of interview (mode) one assumed yielded the most accurate results.

The findings of these experiments in the Odessa CV study lead one to quite difficult conclusions than the Tanzania survey about the ease of implementation of CV surveys and the accuracy and reliability of results. Based on the Tanzanian results, one might infer that it was a relatively simple, straightforward matter to conclude a CV survey among respondents in a poor developing country. The results from Odessa suggest that the potential problems of cross-cultural communications and/or survey implementation may be severe, and that there are real risks involved in the use of cost-cutting measures such as telephone interviews and participatory meetings with public votes. Without these split-sample experiments, policy analysts could easily have false confidence in what they think they know about the demand for improved water services in Odessa, and perhaps the next time they

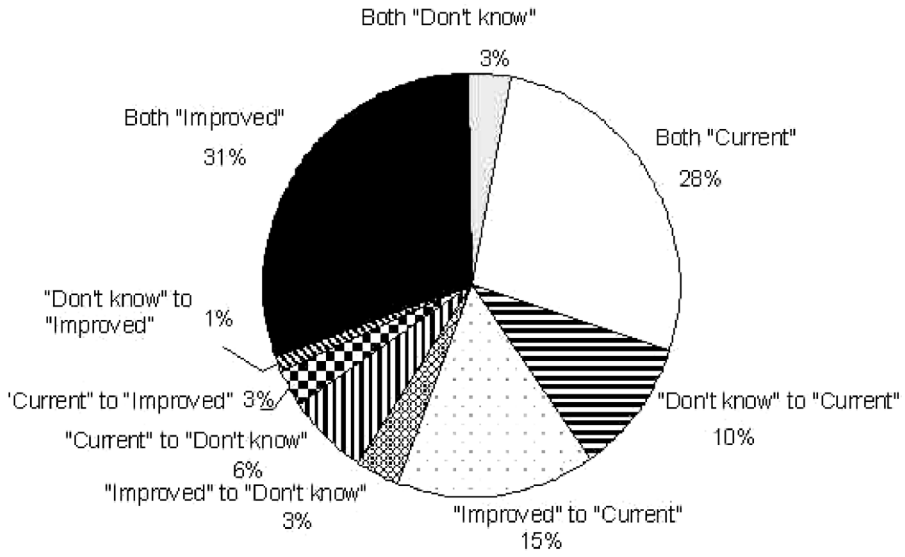

Figure 4. Responses of participants in groups – public and private responses [Source: Davis 1998].

might have been too easily persuaded to rely on less expensive modes of data collection, such as telephone interviews or group meetings with public votes.

Conclusions

Almost thirty years ago Nobel Lauriat Amartya Sen (1973) offered economists the following encouragement to purse research using stated preference methods:

> We have been too prone, on the one hand, to overstate the difficulties of introspection and communication and, on the other, to underestimate the problems of studying preferences revealed by observed behavior.

In some sense Prof. Sen perhaps foretold the current popularity of the contingent valuation method in developing countries, but he certainly never implied that the research agenda would be easy. In this paper I have discussed three problems with existing CV practice in developing countries.[19] Solutions to these three problems will require that CV researchers spend more time and money on their work. Yet the push from mission-oriented agencies is for faster, less expensive "streamlined" CV studies with a practical, operational focus.

Research on stated preference methods has never been popular with donor agencies even though such research is critically important to the successful implementation of these methods. This is unfortunate for two main reasons.[20] The first is that there is no empirical evidence to suggest that rapid, "streamlined" CV surveys yield reliable, accurate results. There is thus a significant risk that this push for cheaper, simpler CV studies could discredit the methodology itself.

Second, and much more important, the policy debates to which CV researchers are asked to contribute are often of tremendous importance to the well-being of households in developing countries. Water and sanitation services, urban air pollution, soil erosion, deforestation, biodiversity, watershed management, ecosystem valuation, vaccines for the poor – these are all policy areas in developing countries in which CV researchers are working on policy questions that directly affect the lives of millions of people. Information on household preferences on such matters is central to both improved policy design and a sound appraisal of policy options. Because the costs of policy mistakes can prove tragic, it is critical that CV researchers push for excellence in this research enterprise and that funding agencies think more carefully about the value of policy-relevant information in these fields. Undertaking more research on stated preference methods will eventually lead to higher quality, and, in some cases, quicker and cheaper studies. However, we are still a long way from the point where it is possible to do high-quality CV surveys with minimal effort or expense.

Acknowledgement

I am indebted to Richard Carson for many excellent comments on an early draft of this paper.

Appendix A: Guidance Notes on Managing and Training Enumerators for Contingent Valuation Surveys in Developing Countries

1. Introduction

The goal of a contingent valuation (CV) researcher is to gauge the attitudes and perceptions of a study population and to listen carefully to respondents' voices. The messages from the respondents to the researcher are filtered through the enumerators. The team of enumerators can thus be viewed as an instrument for listening to people. Unless this team of enumerators is well trained and committed, it is next to impossible for the CV researcher to accomplish her objectives. There will inevitably be some static or "noise" in the people's messages as they are transferred from respondents to the researcher, but good enumerators can help minimize distortions. It is not an exaggeration to say that the primary job of the CV researcher, after designing the questionnaire itself, is to train and manage the team of enumerators.

2. Who Makes a Good Enumerator?

One might wish to find a team of a well-trained, committed enumerators already in place working for a survey research institute or market research firm, and indeed in many developing countries this expertise is increasingly available. The more usual case however is that relatively inexperienced individuals will have to be trained to serve as enumerators. Moreover, even the availability of experienced local enumerators does not obviate the need

for careful field training and supervision. CV surveys pose unique and complex challenges even for experienced enumerators.

For example, I have never found a team of local enumerators familiar with the kinds of split-sample experiments now used in CV surveys. Nor have many experienced enumerators attempted to use photographs and diagrams to convey the kinds of technical or risk information presented in some CV scenarios. Employing experienced enumerators is almost always a good idea, but it is not a substitute for the active involvement of the CV researcher in the training of enumerators and the administration and management of a survey.

If experienced enumerators are not available, a wide variety of individuals can be trained to serve effectively as enumerators. Good enumerators need to enjoy people and feel at ease in the community in which they are interviewing. University students are often obvious candidates for the job of enumerator, in part because of their flexible schedules. I often use them, but in some countries they are poor choices because of class differences between university students and respondents. They also often believe that they can formulate the questions better than the research team, and are tempted to deviate from the explicit instruction that they stick to the exact wording in the questionnaire. I tend to prefer social workers, community organizers, sociologists, health professionals (e.g. nurses), planners, and teachers, but disciplinary or employment backgrounds are not a particularly good guide for selecting enumerators. Interviewers from these groups may create undesirable disparities in status between the interviewer and the respondent. It is not good for the respondent to feel that the interviewer is "smarter" than he/she is, or to get the impression that the interviewer is judging him or her. One does not want respondents trying to impress the interviewers.

One might imagine that extroverted individuals would make the best enumerators because they typically enjoy meeting people and making conversation. This is not always true, however; extroverts can tend to dominate a shy, introverted respondent. In my experience quiet, seemingly shy women often make the best enumerators. They are not threatening to respondents, and almost everyone will agree to be interviewed by such a person.

I have a running joke in our surveys that the quality of an enumerator can be judged by how much weight they gain during fieldwork. Good enumerators make respondents feel comfortable and at ease. Great enumerators make respondents feel so comfortable that the respondent often invites the enumerator to stay for a meal!

My least favorite candidates for enumerators are government officials themselves. They are often aloof from the problems of people, and overly confident that they already know everything about local conditions. This is equally true of the medical profession and professional economists, and another reason not to use them. Furthermore, they are often likely to talk down to respondents. In fact, it can be difficult for respondents to believe that sensitive information (like income) will be held confidential from tax authorities if the interviewer is a government official.

It is often highly beneficial for government officials to listen to the views of ordinary citizens. The job of being an enumerator will force government officials to see things in a different way, and they will be more effective in their job as a result of listening to views and opinions that they might not otherwise hear. However, the CV researcher should be first and foremost concerned with obtaining accurate and unbiased information that

is not influenced by the views of the interviewer. Thus, government officials and other professionals (like doctors) are not ideal candidates for interviewers.

There is one important exception to this generalization about government employees making poor enumerators. Some national statistics or census agencies employ enumerators full-time, and it may be possible to hire such individuals for some short-term CV surveys. Such government employees will obviously be experienced enumerators and understand the many problems interviewers face in the country. On the other hand, because they are permanent public sector employees, they may prove difficult to motivate and train.

3. Training Enumerators

The job of training enumerators in the administration of contingent valuation surveys involves two related but distinct tasks. The first is to ensure that the enumerators understand the objectives of the survey and the subject matter contained in the CV scenario. The second is to provide enumerators with the skills to conduct a lengthy, high-quality in-person interview. If the CV researcher is fortunate enough to employ professional enumerators, then the second task will be less involved than if one is forced to use individuals with little or no survey experience. However, even experienced, professional enumerators will need careful instruction in the objectives of the survey and the material in the CV scenario.

The first step in training enumerators is thus to explain what the study is about. The concepts of economic value and "maximum willingness to pay" are often difficult for the researcher to translate and for some noneconomists to grasp. Open-ended willingness-to-pay questions require that the CV researcher convey the notion of the *maximum* amount an individual is willing to pay; this can be particularly difficult to translate. For example, in a CV study I conducted in Haiti, in response to an early version of an open-ended CV question, a respondent asked one of our interviewers,

> What do you mean the maximum I would be willing to pay? You mean when someone has a gun to my head?

In fact, the CV researcher is trying to determine the maximum amount the respondent would be willing to pay for the hypothetical good or service in the context of the *existing institutional regime* within which individuals are free to allocate their personal or household's financial resources. This does not mean that the respondent is free to choose the price, or pay what she would like (which logically would be zero). The CV researcher would like to measure the amount of income the household could give up after obtaining the goods and services from the project, product, or policy that would make her just as well-off as she would have been if the project had not been built (or the good or service provided).

When a referendum elicitation procedure is used, respondents themselves will not need to be asked an open-ended question about the maximum they would be willing to pay for a hypothetical good or service. Instead, spilt-sample techniques will be utilized, and different prices will be assigned to randomly selected respondents. In the United States there is a fairly large cadre of professional interviewers who work on survey after survey. Split-sample questions are routinely used in U.S. in CV surveys, and such professional interviewers rarely question their use.

This is not the case in most developing countries, and the interviewers themselves will want to understand the reason for the split-sample experiment. One particularly common source of confusion relates to the distinction many people want to make between willingness and ability to pay. It is important for the CV researcher to clearly communicate to interviewers that the purpose of the valuation question is to determine what the respondent would do if he had to make a *real economic commitment* (i.e., faced with an actual budget constraint). In other words, the objective of the CV study is to determine how much respondents are willing *and* able to pay.

The next step in introducing the CV survey to enumerators is to carefully explain the subject matter of the CV scenario. I have conducted numerous CV surveys in which the majority of the enumeration team did not understand the basic scientific and technical aspects of the hypothetical good or service being offered to respondents. Because respondents will often ask questions or request clarifications of matters raised in the CV scenario, it is essential that enumerators be able to respond in a knowledgeable, informed manner.

For example, in several surveys I have managed about demand for improved sanitation services, the vast majority of both respondents and enumerators had never seen or heard of a sewer system with a wastewater treatment plant. The administration of the CV interview required that enumerators use drawings, figures, and photographs to explain to respondents how a sewer system works. As another example, I managed a CV survey in Mexico designed to assess the demand of non-HIV positive individuals for a hypothetical AIDS vaccine. We wanted respondents to understand that triple drug therapy with antiretrovirals was an alternative (but far from perfect) substitute commodity that would extend the lives of most individuals who contracted the virus. However, most uninfected people did not know about triple drug therapy or the costs of such a drug regime. This meant that we had to spend time in the interview providing some information about this treatment.

During training the enumerators should be encouraged to ask questions about the technical material in the CV scenario, and the project director should provide answers for the whole enumeration team to hear. Enumerators need to understand that they must closely follow the script when presenting the CV scenario to the respondent. However, the script of a CV scenario typically asks respondents whether they have any questions about the materials presented. At this point in the interview, it is essential that all the enumerators on the team have a solid, common understanding of the subject matter of the survey so that they will all answer respondents' questions in a consistent way.

During pretesting, it is advisable to ask the enumerators to write down questions that respondents ask during the presentation of the CV scenario and discuss the appropriate response to these questions in a group setting. One technique I sometimes use is to ask one enumerator for a question that a respondent asked them about the CV scenario, and then ask another enumerator how they would have answered this question. I then invite discussion from the group of enumerators before offering my proposed response. If there is a large group of interviewers or the survey is being done nationally, the researcher should develop a formal set of the questions respondents are most likely to raise and the appropriate answers that interviewers should give.

If the enumeration team consists of individuals with little or no experience, then training must include instruction in how to conduct in-person interviews. Table A.1 lists some "do's" and "don't" that should be discussed with enumerators during training.

Table A.1. Rules of good interview practice (do's and don'ts for enumerators)

No.	Advice	Comment
1	*Read every question exactly as written in the questionnaire – do not improvise*	Research on the art of asking questions shows that the precise wording of questions may significantly affect a respondent's answers. If each enumerator develops her own way of asking questions, one can never be sure that the same question is being asked. We need to make sure that each respondent is answering the same question. Reading the question exactly also makes the interview shorter.
2	*Read the question slowly enough so that the respondent can understand.*	An enumerator has seen each question hundreds of times before. It's natural for the enumerator to want to go quickly over a question that he knows so well, but it's the first time for the respondent. The enumerator thus needs to speak slowly.
3	*Wait for the respondent to answer*	Some enumerators will read the question once, then look up and repeat the question, and sometimes even start a lengthy explanation, before letting the respondent answer! Ask once very clearly, and let the respondent think.
4	*If the respondent can't answer, repeat the question.*	The respondent may not have been paying attention the first time. If, after the second reading the respondent still can't answer, go to the next question.
5.	*Remain absolutely neutral about the respondent's answers*	Never express surprise, approval, disapproval, judgment, or doubt about a response. Don't let your facial expression change. Just record the answer. For example, if a respondent says that they would be willing to pay a very large amount for a good or service, the enumerator should not say, "wow!" If a respondent gives an answer that is factually wrong, the enumerator should not reveal that he knows the answer is incorrect.
6	*Do not act embarrassed about a respondent's answers to sensitive questions.*	This will increase the embarrassment of the respondent, not reduce it. Be very matter of fact.
7	*Never suggest an answer unless the instructions say to read the answers to the respondent.*	For example, if the respondent is having difficulty estimating the most he will pay for a good or service, *do not* prompt him with suggestions like ... "would you pay more than US$xx? More than US$yy? Less than zz?"
8	*Don't repeat the respondent's answers.*	This is repetitive and wastes time.

Table A.1. Continued

No.	Advice	Comment
9.	*Conduct the interview in private.*	That means that the interview should not be in earshot of other people in the household. If someone doesn't want to leave, the enumerator should offer to interview him or her separately. If they still won't leave, then the enumerator should explain to the respondent that he will have to return later.
10	*Do not give advice to respondents on personal matters.*	Enumerators should refer respondents to the appropriate authorities for answers to questions that may arise that are outside the scope of the interview.
11.	*Answer directly any questions the respondent may have about the purpose of the survey.*	Respondents are entitled to know the purpose of the survey and how they have been selected to be interviewed. The enumerator should not be reluctant to take time to provide clear, detailed answers to such questions.
12	*Listen carefully to the respondent's answer*	It is very off putting to the respondent if the enumerator is inattentive. Moreover, the respondent may be offering an answer that is in fact different than it first appears to be. In such cases the enumerator needs to be listening carefully to hear what is actually being said.

Sources:
(1) Martha Ainsworth, The World Bank. Personal communication.
(2) Amon, Joseph, Tim Brown, Jan Hogle, Joan MacNeil, Robert Magnani, Stephen Mills, Elizabeth Pisanti, Thomas Reble, Tobi Saidel, Christine Kolars Sow. (2000). *Behavioral Surveillance Surveys (BBS): Guidelines for Repeated Behavioral Surveys in Populations at Risk of HIV*. Family Health International URL: www.fhi.org. Funded by United State Agency for International Development, and the United Kingdom Department for International Development.

One of the most common difficulties I have encountered in training inexperienced enumerators is their initial inability to appreciate the importance of each respondent receiving exactly the same information in the CV scenario. Many enumerators will feel they can "improve" the CV scenario if they tell it "their way," or adapt it to fit the characteristics of a particular respondent. For example, enumerators may assume that they can use more sophisticated language when explaining the scenario to well-educated respondents, or assume that high-income respondents will already know some information that poor respondents do not. They may feel that they should be trusted to present the CV story in the manner best suited to a particular situation, and that the CV researcher should give them the opportunity to express their individual, creative talents with the material in the scenario.

Because CV scenarios can often take several minutes for an enumerator to present, there is a natural temptation for enumerators to tell the story in their own way, i.e., not follow the script exactly. If this tendency is not nipped in the bud, over the course of a

CV survey, each enumerator's version of the CV scenario can evolve in quite different and sometimes startling ways.

To make the point that enumerators must stick to the exact wording of the CV scenario, and this discipline does not squelch all opportunities for the enumerator to express their individual talents, I like to draw the analogy between a CV scenario and a Shakespeare play. A Shakespearean actor must deliver the exact lines of the play. This does not mean, however, that all performances of Hamlet are the same, or that there is no room for individual actors to demonstrate their individual talents. Similarly, the fact that enumerators should stick with the exact wording of a CV scenario does not mean that they cannot give "great performances."

To show enumerators how easy it is for wording to be transformed if they are not diligent; I sometimes play the following child's game with them. The enumerators sit in a large circle and one person whispers a few sentences on some subject to the person next to him on his right. This person then passes on the "message" to the person on her right, and the process continues until the message goes around the circle and returns to the individual who started it. The person who was the source of the message then announces to the group both the initial message and what arrived back. The point of the exercise is to illustrate how difficult it is to deliver the exact wording in a CV scenario to every respondent in a survey.

Perhaps the biggest single problem I have had in training enumerators in good interview practice is that some inexperienced enumerators see their role as an educator who is to convince respondents that they should be willing to pay for the hypothetical good or service offered. I've often seen inexperienced enumerators pleased when a respondent accepts the CV scenario and disappointed when a respondent rejects the proposed offer. This is most likely to be a threat to the validity of the study when educated individuals living in urban areas are taken into rural areas to interview respondents with lower socioeconomic status. Similarly if government officials are used as enumerators, they are likely to see themselves as agents of change or proponents for modernization, and thus try to convince respondents that they need the service described in the CV scenario. The Project Director thus needs to stress that it is legitimate for the respondent to indicate an unwillingness to pay or a low willingness to pay for the hypothetical good or service offered. It is not the role of the enumerator to change the respondent's mind in this regard.

I also find that it is important to stress during training that it is the enumerator's responsibility to make the respondent feel comfortable and at ease during the interview. This seems like simple common sense, but enumerators need to recognize that what makes one person relaxed may make another tense or suspicious. It is thus the enumerator's responsibility to assess the personality of the respondent and to adapt their behavior and interview style to the particular circumstances. For example, a quiet, introspective respondent may be put off by an overly extroverted, lively presentation, while another respondent might thoroughly enjoy it.

I like to use two analogies during enumerator training to make this point. The first is that a good enumerator is like a mirror: the respondent feels comfortable during the interview because he sees the enumerator as someone like himself, sympathetic to his circumstances and preferences. The second is that a good enumerator is like a chameleon: she is able to adjust her presentation of the questions in the survey instrument to the characteristics of the particular respondent. Both analogies seem to help make the point that a good enumerator is not free to simply "be herself," but must pay close attention

to how a respondent is reacting to her and adjust her behavior accordingly to make the respondent feel comfortable.

Another common problem I have encountered with inexperienced enumerators is their failure to appreciate the importance of obtaining high-quality responses to *all* the questions in the survey instrument. For individuals without experience in multivariate statistical analysis, it is ease to assume that if one of their interviews is 90 percent complete, that they have collected a lot of useful information. The CV researcher must stress that if an interview is missing just a single variable (e.g., age or gender), then the entire interview will likely be thrown out of the data set.

For questions such as a respondent's age, education, and income, enumerators need to be instructed on how to probe for answers and not to accept a "don't know" answer from the respondent. If some or all of the enumerators' compensation is paid on the basis of the number of interviews completed, the CV researcher can drive home the importance of answering all the questions by refusing to pay for interviews with missing responses.[21] Enumerators may perceive this practice to be an unreasonable drive for perfection unless the CV researcher carefully explains during training why it is necessary. On the other hand, "don't know" responses are often acceptable answers to the contingent valuation questions. During training it should be made clear to the enumerators when probing of respondents is called for and when it is not.

Enumerators definitely need practice in administering the questionnaire in order to avoid these and other pitfalls. For example, most CV questionnaires contain rather complicated skip patterns, and enumerators need to be thoroughly familiar with these before the actual survey is launched. I routinely break the enumeration team into pairs and let them practice the questionnaire on each other, one being the enumerator and the other the respondent. I also like to have a pair of enumerators role-play the interview in front of the entire team of enumerators, and then let the group critique the interview.

Such role-playing is no substitute, however, for a significant number of supervised practice interviews in the field. Real respondents will raise questions and pose challenges that enumerators will not create for each other in a role-playing exercise. It is thus important to schedule time for supervised in-field practice interviews, and have a debriefing session in which enumerators have plenty of time to discuss the problems that arose and then for the project researcher(s) to discuss these problems among themselves and explain how they should be addressed.

It can be very detrimental to the quality of the training to let interviewers do multiple practice interviews on their own (i.e., without supervision). This can develop and reinforce bad habits. Having a second interviewer observe the first one committing multiple errors can make the situation even worse. The best procedure is for a trainer to accompany each two-interviewer team into the field and observe and critique one interview of each person before letting the trainees loose to conduct interviews on their own.

Launching a survey with a poorly trained enumeration team is about the most serious mistake a CV researcher can make. Training enumerators takes time, and it is a judgment call as to when enumerators are ready to conduct the actual survey. In most of the CV surveys I have designed and managed, training occurs as an integral part of a process of questionnaire design and pretesting. It can easily take 4–6 weeks to develop a questionnaire, train enumerators, and then pretest and revise the questionnaire. The training of the enumerators is occurring throughout this period, at the end of which I hope to have both a polished survey instrument and a skilled enumeration team. This model of questionnaire

development and enumerator training works best with relatively small surveys, or with large surveys in one geographic area that can be closely supervised. It is generally not practicable for large national surveys.

Another model is to conduct the pretest with a handful of experienced enumerators. At the end of the pretest, the research team finalizes the questionnaires, and the experienced enumerators help train new interviewers and perhaps become field supervisors. In this model, more questionnaire development does not occur during the training program.

Typically I initially hire more enumerators for training than I need and make it clear that only a fixed proportion of the group will be retained after training for the full survey. Sometimes I manage this winnowing process with a written test. Other times I ask enumerators to perform a mock interview, and then make a subjective judgment as to whether the individual can do the job.

Table A.2 presents an example of a schedule for a two-week enumerator-training program. Day 1 would be spent on a general introduction to . . .

(1) the objectives of the research project;
(2) introductions of the key project personnel and the trainees;
(3) the scope of the fieldwork being planned;
(4) the responsibilities of each team member;
(5) the organization of the fieldwork; and
(6) an explanation of what is meant by a "split-sample," and the concept that different households will get different questionnaires.

The latter is particularly important; enumerators have to understand that they cannot give any questionnaire to any household. They must use the questionnaire that has been assigned to a specific household.

On Day 2 the CV research team makes a presentation on the technical and scientific subject matter that is the focus of the research. For example, this might include a discussion of various technologies and/or scientific processes unfamiliar to enumerators. The CV research team should invite questions from trainees and allow plenty of time for discussion of this technical material. During the remainder of Day 2 the CV research team presents the CV questionnaire and explains the structure and flow of the questionnaire. This will include the major parts or sections of the questionnaire, and the skip patterns.

On Day 3 two members of the CV research team role-play the entire CV questionnaire in front of the group of trainees. Before the role-play, the enumerators should have had a careful demonstration of how to explain the CV scenario. This includes the use of visual aids, which must be used in exactly the same manner by each interviewer. This will require considerable practice in front of the group, and the trainees should be allowed to critique each other's performance. This will reinforce the notion that everyone must read from the same script. The person demonstrating the CV scenario should emphasize this by making sure that he/she is shown to be glancing at the script periodically, to make sure that they are not missing anything.

Enumerators then need to be instructed in how to conduct good interviews (see Table A.1). Trainees are divided into groups of two and asked to practice the questionnaire with each other. The CV researchers should circulate among the pairs of trainees, listen to how they are doing, and offer helpful suggestions for improvement. This is a good point in the training program to discuss the issue of the confidentiality of responses and the importance of maintaining confidentiality during the course of the survey. Finally, trainees should be encouraged to study the questionnaire in the evening.

Table A.2. Sample schedule for a two-week training program for enumerators

	Monday	Tuesday	Wednesday	Thursday	Friday
Week 1	*Day 1*: 1. Introduction of project staff; 2. Explanation of the objectives of the research project; 3. Introduction to the CVM; 4. CV elicitation procedures; 5. Exercise: Enumerators work in small groups to write a CV scenario.	*Day 2*: 1. Presentation on the technical/scientific subject matter of the research; 2. Overview of the technical subject matter in the CV scenario; 3. Trainers answer questions about technical issues involving subject matter; 4. Discussion of the structure of the draft CV questionnaire.	*Day 3*: 1. Researchers role play the CV questionnaire in front of the group; 2. Discussion of good interview practice (do's and don'ts); 3. Trainees are separated into two groups and practice the interview with each other; 4. CV researchers circulate, listen to trainees practice the interview, and offer constructive comments.	*Day 4*: 1. Pairs of trainees practice administering the questionnaire in front of the group; 2. Group discussion of the pros and cons of each trainee's interview style; 3. CV researchers answer questions about the structure and content of the questionnaire; invite trainees' suggestions for changes; 4. Discussion of the logistical arrangements for practice interviews for Day 5	*Day 5*: 1. Trainees work in pairs; each trainee conducts interviews with two respondents and listens to two interviews by his or her partner (4 total interviews); 2. Discussion of the day's interviews and any problems that arose in the practice interviews; 3. Trainees suggest improvements to the questionnaire; 4. Trainees study and memorize questionnaire over the weekend.
Week 2	*Day 6*: 1. Videotaping of selected trainees' practice interviews; 2. Showing of videos to entire group; 3. Discussion and critique of trainees' performances in videotaped interviews; 4. Suggestions to trainees on ways suggest improvements to the questionnaire.	*Day 7*: (same as day 6) 1. Videotaping of selected trainees' practice interviews; 2. Showing of videos to entire group; 3. Discussion and critique of trainees' performances in videotaped interviews; 4. Suggestions to trainees on ways suggest improvements to the questionnaire.	*Day 8*: 1. Trainees conduct practice interviews alone; 2. CV researchers and field supervisors sit in on selected interviews; 3. Discussion of the day's interviews and any problems that arose in the practice interviews; 4. Trainees suggest improvements to the questionnaire.	*Day 9* (same as day 8): 1. Trainees conduct practice interviews alone; 2. CV researchers and field supervisors sit in on selected interviews; 3. Discussion of the day's interviews and any problems that arose in the practice interviews; 4. Trainees suggest improvements to the questionnaire.	*Day 10*: 1. CV researchers conduct oral and written tests of the trainees; 2. CV researchers announce final selection of enumerators for the project.

On Day 4 several pairs of trainees are asked to practice administering the questionnaire in front of the group. The entire group then discusses the pros and cons of these practice interviews. The trainees should be invited to ask questions about the structure and content of the questionnaire. Trainees may make suggestions for things for the CV researchers to include in the training manual. Sometimes the researchers will overlook instructions that trainees need in terms of "what to do if . . .". The CV research team should quickly incorporate any suggestions from trainees that it feels will improve the implementation of the fieldwork. If errors are discovered in the printed questionnaires during training, the researchers will have to develop a plan to quickly modify them.[22] The remainder of Day 4 is spent planning the supervised practice interviews in the field test on Day 5.

On Day 5 of the training program, enumerators go to the field in pairs to practice the questionnaire with actual households. Each member of the pair of trainees conducts two interviews and listens while his partner conducts two interviews. Members of the CV research team supervise these practice interviews. After the supervised pairs of trainees finish their four interviews, the entire group meets back at project headquarters to debrief and discuss the interviews. Trainees are invited to discuss any problems that arose and ask for guidance or assistance in how to handle particular problems. Trainees are again invited to suggest things that should be included in the training manual.

Days 5 and 6 are spent videotaping the trainees and then playing the tapes of the mock interviews in front of the group.[23] For many trainees, this will be the first time that they have watched themselves on videotape, and this can often be a powerful learning experience and an important way to improve enumerators' performance. The logistics of videotaping a large number of trainees can be complicated, and the schedule may need to be adjusted to reflect the fact that more time is required for this stage of training, or the number of trainees taped may need to be reduced.

On Days 8 and 9, trainees again go to the field to conduct supervised practice interviews. However, this time they do not go in pairs; they conduct the interviews alone. Members of the CV research team observe as many of these interviews as they can. At the end of the day, the group of trainees meets to debrief and discuss the day's work. Trainees are again invited to identify errors in the questionnaires and items to be added to the training manual.

Day 10 of the training program is devoted to both written and oral tests of the trainees so that the CV research team can make a final selection of which individuals to use as enumerators in the actual survey. There are many different ways of devising this final evaluation. If the group of trainees is relatively small, the CV research team will have had ample time over the two-week training program to form judgments of the capabilities of different enumerators.

I have a word of caution, however, about forming impressions of trainees' talents too quickly. Some individuals may start poorly, but just get better and better as enumerators. Others may have some strengths, but have a difficult time breaking bad habits. It is important for the members of the CV team to keep an open mind about individuals' capabilities, and consult with each other about what they see as the strengths and weaknesses of different trainees before reaching a consensus on who should be hired for the actual survey. But regardless of the poise and dynamism of the interviewer, a candidate who is not willing to stick to the exact script of the CV scenario and the questionnaire will not be a good interviewer and should not be hired.

4. Rewarding Enumerators

The CV researcher typically has three kinds of rewards to offer an enumerator: (1) money, (2) references and connections for future work, and (3) fun. All of these rewards should be used to motivate an enumerator to do good work. I discuss each in turn.

(i) MONEY

The basic dilemma facing the CV researcher is how to provide incentives for enumerators to both conduct high-quality interviews and at the same time use their time efficiently. There are two principal ways of providing enumerators with monetary compensation. The first is to pay a fixed daily rate irrespective of the number of interviews completed. The second is to pay enumerators a fixed amount for each questionnaire finished. Both have advantages and disadvantages. In many countries one of these approaches is common, and switching to the other may cause confusion or be culturally inappropriate.

The advantage of using a fixed daily rate is that the enumerator does not have an incentive to rush the interview. The disadvantage is that the enumerator has no monetary incentive to use her time efficiently, either during the course of the interview or in finding respondents. I often prefer this compensation method because it enables me to stress that my main concern is the quality of the interview, and that I do not want enumerators to rush an interview in order to make more money.

The advantages and disadvantages of paying for each questionnaire completed are just the converse of paying a fixed daily wage. Some enumerators will not be able to resist rushing an interview in order to make more money. Others may even be tempted to fill out a questionnaire without interviewing anyone. On the other hand, paying enumerators on a per-interview basis does give a strong incentive not to waste time.

It is possible to combine these two compensation schemes in an attempt to get the best of each. One way to do this is to pay for each questionnaire completed, but guarantee each enumerator a daily minimum just in case an enumerator has a series of long interviews (for example with elderly respondents). Such variations in the length of interviews should even out over a long survey, but they may not in a shorter survey. This approach does not, however, alleviate the incentive to rush interviews.

Another way of combining these compensation schemes is to pay a fixed daily wage to each enumerator, but to structure a bonus incentive to the group of enumerators. For example, each enumerator could be paid an additional lump sum amount if the team accomplishes a specified goal by a certain date. Promising a specified bonus at the end of the fieldwork to all enumerators who complete the survey also provides a strong incentive to finish the job.

However a compensation scheme for enumerators is structured, the issue still arises of how much money enumerators should be paid in absolute terms. In establishing compensation rates, I recommend keeping in mind two guidelines. First, enumerators should feel very well paid and understand that keeping their job depends on good performance. Enumerators should thus be paid a sufficient amount that being fired represents a significant loss. Second, enumerators should not be paid so much that they feel the CV researcher has no knowledge of local wage conditions or wage rates; the CV researcher does not want to pay enumerators so much that they consider the researcher foolish.

Within these two guidelines, there is often a broad range for negotiation. If the CV researcher does not have knowledge of local wages, one way to get a quick sense of the local economy is to determine the wage rate for an unskilled laborer. As a crude rule of thumb, I rarely pay enumerators more than 5–10 times this local wage. Another benchmark is the starting wage for a college graduate in a government job. I would not typically pay an enumerator more than two times this wage.

But these general guidelines are no substitute for a high level of political sensitivity in settling on and/or negotiating monetary compensation for enumerators. Particularly in situations where applicants for positions as enumerators have no experience working with foreign donors or international agencies, their expectations about wages can be very high. Potential enumerators may, for example, expect to be paid international wage rates.

I have had two experiences in which I got off on the wrong foot because of initial negotiations over wages. Both were in rural situations in which local wage rates for unskilled labor were exceedingly low (less than US$1 per day). Once in Haiti I literally had a group of potential enumerators react violently to a proposal I made for wage rates that I considered very generous. I managed to leave, but not before a crowd of potential enumerators attacked my car.

In a survey conducted in rural Ethiopia, a confrontation with enumerators over wages was not physically threatening, but there was a similar huge gap between expectations of what an international donor should pay and local wage rates. In this case enumerators initially expected to be paid an order-of-magnitude higher wage working for an international donor-funded research project than on local research projects. Much of the misunderstanding arises because potential enumerators typically have no idea how a CV researcher's budget is established. In some instances it may help to explain this process.

(ii) REFERENCES

A related but different reward for enumerators is help in obtaining future work. There are several easy things a CV researcher can do to help enumerators obtain future employment. I prepare a letter of reference for all of our enumerators when they have completed a survey. In some cases I have prepared a more formal certificate that acknowledges an enumerator's contribution and recognizes what they have learned. Both are usually greatly appreciated. Another simple thing I like to do is to include a list of enumerators' names, address, and contact numbers as an appendix in the project report so that future researchers who need enumerators will be able to find our enumerators. I promise enumerators both these things when I begin a survey as an additional motivation incentive.

(iii) FUN AND OWNERSHIP

Finally, as with any job, intangible benefits can be among the most important rewards. For some young people working on such a project will be their first exposure to a complicated project management schedule, and they often relish the challenges and excitement this poses. Some of our enumerators always immerse themselves in this experience more than others, but for many people the long hours and intensity of the task can prove exhilarating.

I try to foster this sense of excitement and camaraderie in any way I can. It is important for the CV researcher and the team of enumerators to share housing and meals whenever

possible. I find opportunities for informal time together, including recreational activities. At the end of our surveys I like to throw farewell parties for enumerators. All of these activities are designed to make the experience for enumerators more than simply a job and develop a sense of ownership in the research product.

5. Field Supervision

No matter how carefully trained a team of enumerators is, or how well they are compensated, a CV researcher needs to establish a set of systematic supervision procedures to enhance the quality of the fieldwork. There are three essential components of good field supervision. The first is to review enumerators' completed questionnaires for errors soon after they are finished. The second is to assess the quality of enumerators' performance. The third is to ensure that enumerators are actually conducting the interviews that they say they are.

When field supervision reveals serious problems with an enumerator's work, her employment may have to be terminated. I discuss some of the steps that can be taken to prepare enumerators for such an eventuality. Finally, good field supervision includes not only ferreting out problems with enumerators' work, but also looking after the safety of both enumerators and field supervisors.

(i) DETAILED REVIEW OF COMPLETED QUESTIONNAIRES

Social scientists who have not actually designed and managed survey work often fail to appreciate the importance of sitting down and reviewing enumerators' questionnaires very soon after they are completed. This kind of immediate review of questionnaires accomplishes several things. First, it conveys to each enumerator the message that you are keenly interested in the information they are collecting, and that you care about the quality of the work they are doing. Second, it enables the research team to quickly identify problems with the questionnaire and correct them. Third, it enables the field supervisors to get to know each enumerator.

Fourth, the field supervisors can provide immediate positive feedback to enumerators that are doing well and correct those that are not. Fifth, it enables the field supervisors and CV researcher to emphasize to the enumerators that they need to be engaged in the interview process and detect inconsistent answers. Sixth, the field supervisors can standardize judgment calls as to how to interpret respondents' answers to particular questions, and thus prepare questionnaires for data entry. In a large survey, immediately reviewing all of the completed questionnaires can be quite time consuming, and may not always be feasible. However, it should be done if at all possible, and sufficient time and resources need to be allocated for this purpose.

(ii) ASSESSING THE QUALITY OF ENUMERATORS' PERFORMANCE

In addition to reviewing each enumerator's completed questionnaires, it is also important to accompany enumerators and watch how they handle actual interviews. This task logically falls to the field supervisors, but the CV researcher should also spend time observing enumerators conducting interviews. One might imagine that this is not possible unless

the CV researcher understands the local language. Without meaning to de-emphasize the importance of local language skills, much can be learned by simply observing the body language of an enumerator and respondent, and following the tempo of an interview.

One technique that I have found useful in fieldwork supervision is to post a large chart in the main office or headquarters of the field operations. The rows of this chart are the names of the enumerators; the columns are the days of the week(s) that the actual interviewing of respondents is underway. Every day I ask the enumerators to come into the field operations center to return their completed interviews and pick up blank questionnaires. While they are in the office, each enumerator is asked to write on the chart the number of questionnaires they completed on the designated day (sometimes we also keep track of the number of each version of questionnaire completed as well as the totals). In this way the enumerators can easily see both the overall progress of the survey, and how the number of questionnaires that they have conducted compares with other enumerators.

In my experience this technique serves to benchmark standards for a day's work interviewing. Most enumerators will not want to be a laggard in terms of the number of questionnaires completed. This technique thus provides enumerators with an incentive to at least keep up with the group average of the number of interviews completed per day.

In a large national survey, it is typically not feasible for enumerators to meet regularly at a central project headquarters, and thus posting a chart of the overall progress of the survey is not much help in providing feedback to individual enumerators. In this case it may prove useful for field supervisors to share this kind of information with enumerators in the field.

(iii) RANDOM RE-CHECKS OF ENUMERATOR'S INTERVIEWS

In any large survey the CV researcher must expect and plan for the fact that some enumerators will be tempted to either falsify questionnaires, or interview respondents that are easier to reach than those in the sample frame. Field supervisors must arrange to return to randomly selected respondents and verify that they were indeed interviewed, and that they gave the answers recorded on the questionnaire.

It can prove particularly challenging for field supervisors to find individuals who have already been interviewed when sampling procedures are used that do not identify particular individuals in advance. In such cases the CV researcher may need for the questionnaire to include information on a respondent's name and address (as well as detailed instructions on how to locate his or her dwelling). Such information may seem to contradict any promises enumerators make about the confidentiality of the interview.

An alternative procedure to the use of random spot checks is to have the field supervisor randomly select a few of an enumerator's questionnaires, and ask the enumerator to take the field supervisor back to these specific households. The field supervisor would then confirm with the respondent that he was interviewed, and that these were the answers he gave. This requires that the enumerator make detailed notes on the location of all interviews conducted.

To be effective the field supervisor needs to revisit a respondent soon after the interview. In most cases it is unrealistic to expect that an enumerator can locate a specific respondent a long time after an interview is finished. This approach requires that the field supervisor be physically located in the vicinity of the enumerators. If the field supervisor is spread

across several towns or communities, the likelihood of a field supervisor arriving without warning is reduced.

(iv) FIRING ENUMERATORS

Good field supervision usually means that poor and/or dishonest enumerators will be found; such individuals' employment must be terminated. I expect to fire some enumerators in every large survey, and prepare field supervisors and enumerators during training for this. I tell the group of enumerators at the beginning of the study that not everyone makes a good enumerator or enjoys this job, and that their continued employment depends upon how well they perform. There are numerous reasons why an enumerator might need to be fired. The most egregious is the falsification of data. Another reason would be if he/she revealed confidential information about a respondent to others, or was observed talking to others about the responses of a specific individual.

On occasions I have discussed with the entire group of enumerators problems that I was having with a particular enumerator, and let the group decide how this individual's problem should be handled. Such a process makes it clear to all enumerators the high standards of performance the CV researcher expects.

One may find that an entire team or group of enumerators have colluded to deceive the CV researcher. To my knowledge, this has only happened to me once (the culprits were a group of economics graduate students). Firing a group of individuals can be even more sensitive and complicated than a firing an individual, in part because the group may try harder to defend the integrity of its members, and definitive proof of collusion may be lacking.

I have no general guidance to offer on this subject except to make two points. First, if either a group of people or a single individual must be fired, the entire management team of the research project must proceed in a careful, deliberative fashion, and reach a consensus on how to respond. But action must be taken quickly and decisively. Second, the management team must explain to the enumerators who remain what happened, why a particular course of action was chosen, and answer any questions they may have about the events.

(v) SECURITY RISKS

It is always important to be cognizant of any personal security risks that enumerators may be taking and discuss this issue openly with the team. The survey researcher should be aware of the concerns that enumerators may have about working in certain neighborhoods or after dark, and take such fears seriously. In some of our surveys such concerns have in fact dictated how we have organized the fieldwork. For example, in Odessa, Ukraine, we sent teams of enumerators into neighborhoods together so that they would be near each other, and some women preferred to work in pairs rather than go into an apartment alone. Similarly, in a survey in Sofia, Bulgaria, in certain neighborhoods some female enumerators preferred to be accompanied by a male enumerator when they went into a respondent's apartment. In Timor in Indonesia I adjusted the sampling protocol itself to ensure that enumerators would be picked up from assigned drop-off points at precisely the time (at dusk) when the vehicle drivers promised to be there.

6. Concluding Remarks: How Long Should an Enumerator Work on a Survey?

Professional enumerators employed by survey or market research firms obviously work on one survey after the other; such people often work for years in this capacity. However, for many university students or others who are hired for work on a CV survey, they do not see themselves as long-term enumerators. At some point they will likely tire of the work as the survey proceeds. Based on my experience there is an optimal amount of time that most enumeration teams should work. This varies depending on a host of factors, such as the difficulty and schedule of the survey (e.g., the number of evenings and weekends worked); the length of time away from home and family; alternative employment opportunities open to enumerators; and housing accommodations when enumerators are away from home.

There is no general rule about how long a team of enumerators should continue. Of course, when a project begins enumerators should be informed as to how long the work is expected to last. But schedules often change, and the CV researcher should not simply assume that a team of enumerators can continue working past agreed upon deadlines. Even if enumerators agree in advance to work for three or four months on a survey, they may simply not know how they will feel after two months of interviewing.

Whatever schedule is agreed upon in advance, the CV researcher should be alert to the onset of "enumerator fatigue," and adjust the work schedule when it becomes a problem. The solution may be as simple as allowing enumerators to take some time off, but practically this may compromise other aspects of the fieldwork logistics. The most important point to emphasize is that the CV researcher needs to have open lines of communication with enumerators and to continually assess how they are feeling as the survey work unfolds.

A related issue concerns the need that may arise to rotate the membership of people on a team. For a survey that runs for a long time, interviewers and field supervisors may become too close for the field supervisor to be able to objectively assess the work of the interviewers under his or her direction. Thus, for a long survey it is a good practice to rotate interviewers and/or supervisors across teams.

Appendix B: Two Contingent Valuation Scenarios – Human Resources Development Survey (Tanzania)

COMMUNITY HEALTH TRUST FUND

Today the government tries to provide free medical care to everyone. But often people do not use government medical services because drugs are out of stock, or because other doctors are easier to get to, or just because they prefer to use other medical services. In any case, you often have to pay money for medical care when you become ill, even when you get help from government facilities.

The government is considering a new way of helping Tanzanians pay for medical care. I will now describe the idea.

(1) Your household could purchase a health card for your family.

(2) The government would pay for half the cost of the health card.

(3) This card would allow you to choose a medical facility for your family, which would take care of you at no additional cost for 6 months. You could choose a govern-

ment dispensary, a health center, or a hospital. Or you could also choose a mission dispensary or hospital or you could choose a private doctor.

(4) Your doctor would be required to keep good records for you and your family and to help you stay healthy, not just treat you when you are ill.

(5) At the end of 6 months, you could buy another health card or drop out of the program. The government would continue to pay half the cost of the card. You could also choose a new medical facility, or you could continue to use the same facility for another 6 months.

(6) The health card would cover all costs of getting regular medical care, such as consulting with the nurse or the doctor, any drugs you might need, and injections.

However, if you have to be admitted to the hospital to stay overnight, the health card would not pay. Instead, the government would offer low-cost hospital care in case you need it. The government would no longer attempt to provide free medical services, and government facilities would operate the same as other facilities. Therefore, if your family did not buy a health card, you would have to pay the full cost of health care anywhere you go, whether you use a private facility, a mission facility, or a government facility.

This is a simple idea, but it changes the way you would pay for health care. Instead of paying when someone in your family gets sick or is injured, you would pay for the card when you are healthy. But there is always the chance you would not need to use a doctor during the 6 months. In that case you would pay for the card but not get any services.

Now I will repeat the main features of this proposal. First, you would buy a health card for your household. The government would pay about half the cost. So if the card cost Tsh. 20,000, you would pay Tsh. 10,000 and the government would pay the remaining Tsh. 10,000. Second, you would have to choose one medical facility to take care of your family for 6 months. During that 6 months, you would get services from that facility by using your health card, without paying more money. But if you use any other facility, you would have to pay the full cost without any help from the government. Third, at the end of 6 months, you could buy another card or quit the program. You could also switch to a new facility for the next 6 months or continue with the same facility. The main feature is that the government would help you pay as long as the facility you choose is part of the program. Today, to get help from the government, you must to a government facility.

Do you have any questions about the health card? If you do, I can try to answer them now.

Now I am going to describe the health card program again and ask how much you would be willing to pay for the card. You would buy a health card for your family. It would pay for all medical care your family needs for 6 months except if someone has to stay in a hospital. You would choose a single health facility to take care of your family for those 6 months. You would not have to pay any more money to use that facility. But if you use any other facility, you would have to pay. At the end of six months, you would be able to choose whether to participate again, and you could choose a different facility. Please keep in mind how much you pay for medical care today when someone gets sick and whether it would be worth it to you to buy a health card to avoid these payments. Sometimes the health card might cost you more than you pay now; sometimes it might cost you less. It depends upon how often your family needs to use a doctor. If you offer to pay more for the health card than you pay for medical care today, you will have to give up something else to pay the additional amount. If the health card saves you money, though, you could purchase other things with the money you save. I would also like to remind you that this

program would only work if your payment and the government's matching contribution are combined; otherwise there would not be enough to pay for medical services for 6 months. Therefore, please think carefully about how much you would really pay for such a card when I ask you the questions. Are you ready for the questions about what you would pay, or would you like more information first?

Would you pay Tsh (1,000/25,000/50,000) for a health card that would entitle your family to medical care for 6 months, from a facility that you would choose? The government would match the amount you would pay.

GIRLS SECONDARY EDUCATION SUPPORT

Many children are unable to go to secondary school because there are very few spots open in government schools and many families are too poor to be able to send their children to a private school. For this reason, the government is planning to provide funds for girls to attend secondary school if they have done well in primary school, but their families are too poor to afford to send them to a private secondary school. I will ask you later how much you would be willing to contribute to be part of such a program if it were available to you.

Under the program, government and private secondary schools would be given funds for bursaries to support an increase in the school's enrollment. The bursaries would support girls who are identified by local communities as disadvantaged, and those who score well on the Primary School Leaving examination (Standard 7 exam). The bursaries would pay for some of the girls' expenses, including fees, uniforms, and books. The girls' families would also be asked to contribute in some ways. For example, families would assist with the girls' transportation to and from school and might have to pay some other costs.

How would the girls be selected for the bursary? First, the School Committee and the Village Council would develop rules to determine whether a household needs help. These rules should be well known in the village. Second, the School Committee and the Village Council would prepare a list of girls who meet these conditions. They would come from poor families, but would be doing well in primary school and are likely to do well on the Primary School Leaving Examination. That list would be publicized in the village, to make sure that everyone knows who has been selected. The DEO would make a final list of eligible girls available to secondary schools. The secondary schools would contact the girls and offer them support for four years of secondary education. Girls and their parents could choose the secondary school they like the most, and the girls would enroll. The government would pay for at least part of the cost of fees, uniforms and books for each girl selected; parents would have to contribute to cover transportation and other expenses not covered by the bursary.

Now I am going to briefly describe the program and ask you whether you would be willing to participate in it at various levels of parental contribution. Remember, this program is designed to increase the number of girls successfully completing secondary school. The School Committee and Village Council will select girls who are likely to do well in secondary school, but whose families are too poor to afford the school fees. While the government will be providing funds to secondary schools for the girls' education, parents will also be asked to contribute to their child's education. For example, if your daughter were selected, you would have to pay for transportation and for any costs not covered by the bursary. Remember, if you choose to contribute, if your daughter is selected,

you will have to give up something else to have the money to contribute. The more you would be willing to contribute, the less you would have to purchase other things with the money.

Would you be willing to pay Tsh. (1,000/25,000/50,000) for the transportation and incidental expenses for one daughter for one year of secondary school under this program?

Notes

1. One should not infer from this that I believe that stated preference studies conducted in industrialized countries are on average superior to those in developing countries. In fact, I am agnostic on this point. I have not read a sufficient number of the contingent valuation studies conducted in the United States and Europe to draw a firm conclusion on this question.
2. The interested reader can learn more about the EEPSEA and SANDEE programs from their respective Websites: www.eepsea.org and www.sandeeonline.org.
3. A simple and rather obvious check for this particular problem is for CV analysts to include enumerator dummies as explanatory variables for respondents' WTP answers to see if interviews conducted by a particular enumerator are systematically different from others (see Cropper el al. 1999, for an example of this type of analysis).
4. Journal editors should take three steps to improve the evaluation and practice of contingent valuation surveys. First, they should require that the main CV valuation question be included in the journal article. Second, they should require authors to submit their survey instrument (in both the language in which it was administered and an English translation) with their article to help reviewers evaluate the paper. Third, the entire CV survey instrument should be made available as a *.pdf file on the journal's and/or the author's website so that other researchers can study the entire questionnaire (including any color graphics). If the survey instrument is posted on the author's website, the journal editor should require that the author's html address be included in the journal article.
5. In some general population surveys, it is in fact not possible to ask some groups of respondents exactly the same CV questions as other groups. For example, when estimating the demand for improved infrastructure services, the CV researcher will typically need a different CV scenario for owners and renters. In effect, the research involves two related but distinct CV surveys (one for renters and one for owners). But for each group, all respondents must receive the same CV scenario.
6. The public goods character of a mass vaccination campaign would also apply to a well-designed program to vaccinate high-risk groups.
7. Admittedly there is something of a chicken and egg problem here. While it is clear that the open-ended pretest question will not yield an accurate estimate of the willingness-to-pay distribution, it may help the researcher get in the right range of prices when working at an early stage on the development of the survey instrument.
8. At the time of this survey, US$1 = 72 Nepalese rupees.
9. We could not ask renters without connections precisely the same contingent valuation questions as owners because renters without connections would have less incentive to pay the connection charges associated with a shared or private connection. We thus asked renters to suppose that the landlord would pay the connections charges, and the renter would face an increase in his monthly rent in exchange for the improved water service provided by the private operator.
10. Households without a private connection that relied on water from neighbors with piped connections might be concerned that the price of water that their neighbors charged them would increase if the plan for the new improved system were implemented.

11. We assume here that unconnected households have the option to connect to the piped distribution system. In most neighborhoods in Kathmandu, this is a reasonable assumption.
12. At the time of the survey (1995) US$1 = 550 Tsh.
13. These frequency distributions are based on data from a follow-up survey round that included 476 households in four districts in the central part of the country: Igunga, Manjoni, Kilosa, and Kibaha.
14. In Odessa, as in many places in Eastern Europe and the former Soviet Union, hot water is supplied through a separate piped network. Our CV survey did not propose any changes in the hot water service.
15. Note the difficulty we (and other CV researchers working in the water supply sector) faced in constructing a CV scenario that was understandable to respondents and accurately communicated to respondents the relationship between the monthly bill and the quantity of water used. Ideally one would ask how much water a household would use at a stated price, but only rarely can a respondent answer such a question. Here we asked the respondent to imagine an average month's water bill. Implicitly we are assuming that the household uses as much water is it wishes (or a "reasonable" amount of water, even though the connection is metered. This is probably not a bad assumption in Odessa where the marginal cost of additional water use was essentially free at the time of the survey, and, since most households lived in multi-family housing units, lawn and garden watering was not an issue.
16. At the time of this survey, there was a system of discounts in effect in Odessa for different types of "deserving" households, e.g., military veterans. Respondents were told to assume that discounts on cold water service would no longer be honored, i.e., that they would have to pay the full amount quoted.
17. People who came to the participatory community meetings were on average considerably older and poorer than respondents in the other three groups. They were also more likely to be female and have less education.
18. Even though the results for the in-person household and intercept surveys appear similar, multivariate analysis reveals that they are statistically different from each other (Davis 1998).
19. Notably absent from my list is poor econometric analysis. This is not because I feel that econometrics is unimportant, but in my opinion it is not yet the primary concern. Good econometrics cannot make up for the three types of problems I have described. But good econometric analysis is of course crucial if the CV researcher succeeds in collecting high-quality data. See Hanemann and Kanninen (1999) for an excellent discussion of the issues involved in the analysis of discrete-response CV data.
20. A third reason is that such research is cheap in developing countries (Whittington 1998), so, other things equal, one should think that funding agencies would want more of it.
21. Of course, this can simply lead to fabrication of data. It is thus important during field supervision to have the supervisors not only confirm that the interview took place, but also that key data that could be easily fabricated were indeed given to the enumerator by the respondent.
22. An example might be an error in a skip instruction.
23. Note that videotaping could easily occur before the first practice interviews (rather than after, as is suggested here).

References

Bateman, I., I. H. Langford and J. Rasbash (1999), 'Elicitation Effects in Contingent Valuation Studies'. Chapter 15 in Ian J. Bateman and Kenneth Willis, eds., *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries*. Oxford University Press, pp. 511–539.

Carson R. T., J. L. Wright, N. J. Carson, A. Alberini and N. E. Flores (1995), *A Bibliography of Contingent Valuation Studies and Papers*. San Diego: Natural Resources Damages Assessment Inc.

Carson, Richard T. (1997), 'Contingent Valuation Surveys and Test of Insensitivity to Scope', in R. J. Koop, W. Pommerhene and N. Schwartz, eds., *Determining the Value of Non-Markets Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods*. Boston: Kluwer, pp. 127–163.

Carson, R. T., N. E. Flores and N. F. Meade (2001), 'Contingent Valuation: Controversies and Evidence', *Environmental and Resources Economics* **19**, 173–210.

Carson, R. T., N. E. Flores and W.M. Hanemann (1998), 'Sequencing and Valuing Public Goods', *Journal of Environmental Economics and Management* **36**, 314–323.

Cropper, M., M. Haile, J. Lampietti, C. Poulos and D. Whittington (1999), *The Value of Preventing Malaria in Tembien, Ethiopia*. World Bank Policy Research Paper, No. 2273. December.

Davis, J., D. Gercheva, K. Komives and D. Whittington (1996), *Household Water Supply Conditions and Willingness to Pay for Improved Services in Odessa, Ukraine (with Implications for Serving the Poor)*. A Report to the World Bank. October, 74 pp.

Davis, J. (1998), *Assessing Community Preferences for Development Projects: Does Mode Matter?* Ph.D. Dissertation. Department of Environmental Sciences and Engineering, University of North Carolina at Chapel Hill.

Davis, J. and D. Whittington (1998), 'Participatory Research for Development Projects: A Comparison of the Community Meeting and Household Survey Techniques', *Economic Development and Cultural Change* **41**(1), 73–94.

Diamond, P. A. and J. A. Hausman (1994), 'Contingent Valuation: Is Some Number Better than No Number?' *Journal of Economic Perspectives* **8**(4), 45–64.

Diamond, P. A., J. A. Hausman, G. K. Leonard and M. A. Denning (1993), 'Does Contingent Valuation Measure Preferences? Experimental Evidence', in J. Hausman, ed., *Contingent Valuation: A Critical Assessment*. Amsterdam, Elsevier Science Publishers.

Hanemann, W. M. and B. Kanninen (1999), 'The Statistical Analysis of Discrete-Response CV Data', in Ken Willis and Ian Bateman, eds., *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries*. Oxford University Press, pp. 302–443.

Lauria, D. T., D. Whittington and K. Choe (1999), 'Household Demand for Improved Sanitation Services: A Case Study of Calamba, Philippines', in Ken Willis and Ian Bateman, eds., *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries*. Oxford University Press, pp. 540–581.

Mitchell, R. C. and R. T. Carson (1989), *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Resources for the Future. Washington D.C.

Pattanayak, S., D. Whittington, J. Yang and B. Kumar (2002), *Willingness to Pay for Improved Water Supply in Kathmandu Valley, Nepal*. Research Triangle Institute. Final Report to the Water and Sanitation Program for South Asia, New Delhi: World Bank. August, 2001.

Sen, A. K. (1973), 'Behavior and the Concept of Preference', *Economica* **40**.

Whittington, D., V. K. Smith, A. Okorafor, A. Okore, J. L. Liu and A. McPhail (1992), 'Giving Respondents Time to Think in Contingent Valuation Studies: A Developing Country Application', *Journal of Environmental Economics and Management* **22**, 205–225. Reprinted in Kenneth

G. Willis, Kenneth Button, and Peter Nijkamp, eds., *Environmental Valuation*. Cheltenham, UK: Edward Elgar Publishing, 1999. Volume 1, Chapter 11.

Whittington, D. and J. Davis (1994), 'A Review and Assessment of Techniques for Systematic Client Consultation', A Report to the World Bank, Washington DC, December 16, 66 pp.

Whittington, D. (1998), 'Administering Contingent Valuation Surveys in Developing Countries', *World Development* **26**(1), 21–30.

Whittington, D., O. Matsui-Santana, J. Freiberger, G. van Houtven and S. Pattanayak (2001), 'Individuals' Willingness to Pay for a HIV/Aids Vaccine: Evidence for Guadalajara, Mexico', Proceedings of the XIII International AIDS Conference. Durban, South Africa.

Whittington, D., D. T. Lauria and K. Choe (1997), 'The Effect of Giving Respondents 'Time to Think" on Tests of Scope: An Experiment in Calamba, Philippines', in Ray Kopp, Werner Pommerehne and Norbert Schwarz, eds., *Determining the Value of Non Market Goods: Economic, Psychological and Policy Relevant Aspects of Contingent Valuation Methods*. Kluwer Publishers, pp. 219–234.