ABSTRACT

Marc Jeuland: Estimating the private benefits of vaccination against cholera in Beira, Mozambique: a travel cost approach

The paper reports the results of a study that estimates households' private demand for cholera vaccines based on their travel behavior. We take advantage of an unusual natural experiment. In January 2004, more than 41,000 residents in Beira, Mozambique received two doses of the new-generation oral recombinant toxin B subunit killed whole-cell rBS-WC cholera vaccine during the first vaccination trial to test its effectiveness in an endemic cholera zone of Africa. The trial was designed to target about 22,000 residents in the Esturro neighborhood; nine outposts were established there to distribute vaccines free of charge. Due to the high demand for the vaccines, the trial was modified so that citizens from outside Esturro could also be vaccinated. About 30,000 outsiders came, resulting in long queues and an average waiting time of about 85 minutes per dose.

We obtained information from the complete database of vaccinated individuals collected at the Esturro vaccination outposts, as well as household information collected from a sample of city-wide, in-person interviews conducted in the summer of 2005, to estimate travel cost models of the revealed demand for cholera vaccines among households informed of the trial. We estimated standard and zero-inflated household count models of vaccine demand and dichotomous choice models for the head of the household. To our knowledge, this is the first application of the travel cost method to estimate vaccine demand.

Our travel cost analysis showed that distance traveled and time spent in acquiring vaccines were critical determinants of coverage levels in the population. The quantity of vaccines obtained by households decreased as travel cost — in time and transport expenses — rose. Our best estimates of per capita willingness to pay for cholera vaccination are about US$1. These travel cost estimates are sensitive to the assumed value of time spent acquiring vaccines, and are somewhat lower than those obtained using the contingent valuation method (about US$1.40 per capita).

KEYWORDS: travel cost, willingness to pay, vaccine demand, cholera, Mozambique

ACKNOWLEDGEMENTS

This research is part of the Diseases of the Most Impoverished program (DOMI), administered by the International Vaccine Institute (IVI) with support from the Bill and Melinda Gates Foundation. The DOMI program works to accelerate the development and introduction of new-generation vaccines against cholera, typhoid fever, and shigellosis. The program involves a number of parallel activities, including epidemiological studies, social science studies, and vaccine technology transfer. The results will support public decision-making regarding immunization programs for cholera and typhoid fever.

Thanks are also due to my collaborators in this research: Marcelino Lucas, John Clemens and the rest of the IVI team and Dale Whittington. Finally, I am grateful to Donald Lauria, Dan Phaneuf, Gregory Characklis for their many helpful comments and criticisms of this work.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

INTRODUCTION

Cholera is endemic throughout large parts of Mozambique; the largest affected city is Beira (pop. 450,000). Residents in most parts of the city are vulnerable to regular disease outbreaks. The annual number of patients with severe diarrhea admitted to the city's Cholera Treatment Center (CTC) ranged from 4.5 to 11.2 patients per thousand residents over the period 2000-2004 (Lucas, 2005). Large cholera outbreaks have occurred often, most recently in 1999 and in early 2006. These epidemics often coincide with serious flooding of low-lying areas of the city, where many dwellings lack appropriate sanitation facilities (WHO, 2004).

The Ministry of Health and the International Vaccine Institute (IVI) launched a pilot vaccination trial in Beira from December 2003 through January 2004 to test the effectiveness of the oral, two-dose, recombinant toxin B subunit killed whole-cell rBS-WC cholera vaccine that was distributed free of charge.[1] A case-control study associated with this trial was conducted only for the 21,818 residents of Esturro *bairro* in central Beira. Eight of the nine outposts were spread around Esturro; the ninth was situated 0.2 km away at a primary school in the *bairro* of Matacuane. Household demand for the vaccines was so high, however, that the government decided that they should be made available at these outposts to all Beira residents. Ultimately nearly three-quarters (29,836) of the 41,011 recipients of the two-dose regime came from outside Esturro (Deen, 2004).

Residents of the city were informed of the cholera vaccination campaign via radio broadcasts, visits from health workers, and word of mouth. These publicity efforts were

[1] The vaccination trial was part of the Diseases of the Most Impoverished (DOMI) program, funded through contributions from the Bill and Melinda Gates Foundation. IVI conducted an epidemiological case-control study of cholera patients to determine vaccine effectiveness and found that it was 78% in the first year following the trial (Lucas et al., 2005).

particularly strong in Esturro and nearby areas, and misinformation about the trial increased with distance from the outpost locations. Many adults walked with family members from distant neighborhoods to acquire vaccines, and some used public transportation or private vehicles to travel to vaccination sites. On arrival at a vaccination outpost, people spent additional time in queues.

The expansion of the Esturro clinical trial offered us an unusual natural experiment to estimate private demand for cholera vaccines based on revealed behavior. In a normal vaccination campaign or clinical trial, vaccination outposts would be located as near as possible to the target population; one would never attempt to vaccinate individuals throughout a city with vaccination posts in only one neighborhood. Due to this atypical situation in Beira, households were confronted with very different monetary expenses and time costs to obtain the cholera vaccine. Among those informed of the trial, some chose to incur these costs, and others did not. In this paper we utilize a travel cost methodology that uses a survey methodology to identify the population of households informed of the vaccination effort, and takes advantage of the unusually large variation in the effective price, or travel cost, of obtaining the vaccine.

We initially designed the study in Beira to measure private demand for the cholera vaccine using the contingent valuation method (CVM), a stated preference approach. The vaccine trial experience in Beira therefore also offers a rare opportunity to compare the private health benefits of cholera vaccination using both revealed and stated preference methodologies. The details of the CVM study are reported elsewhere (Lucas, Jeuland et al., 2006). Discussion here is focused on the travel cost models we developed to estimate private willingness to pay (WTP) for cholera vaccines in this location, and our comparison of results from the two studies.

BACKGROUND - NONMARKET VALUATION TECHNIQUES

Revealed preference analysis seeks to explain household demand for nonmarket goods by observing demand for related goods such as substitutes or complements, or by considering markets in which prices may reflect differing levels of quality of the nonmarket goods (hedonic studies). The substitution model has typically been used in health applications because individuals frequently engage in treating, mitigating, or coping behaviors when faced with negative health impacts. In the vaccine literature, one such strategy is to calculate expected averted cost of illness (COI) as a lower bound for possible vaccine benefits (Jefferson, 1999; Fischer, Anh et al., 2005). However, critics have pointed out that expected COI cannot be a complete measure of the benefits from vaccination as it does not encompass individuals' risk of death or pain and suffering (Cropper, Haile et al., 2004).

The travel cost method (TCM) is another revealed preference approach, based on the insight that consumption of a nonmarket good may be linked with complementary consumption of private goods such as time and transportation. Inferences about demand for the former can be made by examining how individuals make decisions about the latter. The TCM has been widely used in the field of environmental and resource economics to estimate demand for recreational facilities (Freeman, 2003). Economists have acknowledged that time and transportation issues can affect choices that individuals make about their health care and often include travel expenditures in cost of illness estimation (Harris and Keane, 1999; Harris, Schultz et al., 2002; Leonard, Mliga et al., 2002), but the TCM has rarely been used in health economics. Clarke (Clarke, 1998; 2002) used the TCM to analyze the demand for mobile mammography examination units in Australia; Brown and Jewell (Brown and Jewell, 1996) used it to examine the demand for abortion services. To date, no attempt has been made to determine vaccine demand using the TCM.

Stated preference methods such as CV, on the other hand, have been applied to a wide variety of health-related valuation problems, including vaccine demand (Whittington, Matsui-Santana et al., 2002; Cropper, Haile et al., 2004; Canh, Whittington et al., 2006; Cook, Whittington et al., 2007; Kim, Canh et al., 2007), largely because of their flexibility. Meta-analyses of results from stated and revealed preference studies fail to show systematic and large discrepancies (Brookshire, 1982; Carson, 1996). Nonetheless, many economists and policymakers remain doubtful about the accuracy and reliability of survey results that report what people say rather than what they actually do. On the other hand, two challenges in implementing the TCM in this case are determining which households were sufficiently informed of the trial to make decisions regarding participation in the trial, and specifying the value of time spent traveling. A comparison of stated and revealed preference approaches is thus timely in the vaccine policy field.

FIELDWORK AND DATA SOURCES

Two types of demand data were used for our travel cost analysis: basic data from vaccination records of all participants collected at outposts during the trial, and survey data collected *ex post* during the summer of 2005 from a sample of nearly 1,400 residents in various neighborhoods of Beira and the neighboring town of Dondo, only some of whom had participated in the trial. This paper focuses primarily on analysis of the survey data. Appendix 1 presents an analysis of the data from the vaccination outposts, using a zonal travel cost modeling methodology.

Household Sampling and Data Collection

The purpose of the travel cost component of the household surveys was to collect information on participation in the vaccine trial. The data needed pertained to the knowledge households had about the trial and the reasons for participating or not participating in it, the costs of travel to vaccination outposts, the time needed to acquire vaccines, the mode of transportation used, and the household members who traveled to the Esturro outposts to be vaccinated (if anyone). A comprehensive list of households in Beira was not available for use as a sample frame for our 2005 household survey. We thus developed a three-stage sampling procedure. First, using the 1997 census of the city's 22 *bairros* (INE, 1998), we determined the proportion of the total sample to select from each of four previously classified urban areas. *Bairros* were randomly selected from each urban area – nine *bairros* in all – and sample sizes were assigned on the basis of population weights.

In the second stage, smaller neighborhood units *(unidades)* were randomly selected from each chosen *bairro*. The third stage involved the selection of households in the *unidades*. Project field staff scheduled interviews with every fifth house. Finally, we also conducted interviews with an additional 284 households living in the neighboring town of Dondo, a few of whom had traveled to Esturro for vaccines. The interviews conducted in Dondo and in one of these nine randomly chosen *bairros* were not included in our CVM study.

Project field staff selected households according to a set of inclusion criteria: interviewed households were required to have one or more children less than 19 years of age, and the interview had to be conducted with the head of household or his/her spouse. This household survey sample thus included participants in the Esturro trial as well as non-participants. Interviews were conducted in Portuguese; only 3% of respondents who were insufficiently fluent in Portuguese were interviewed in the local languages (Ndau and Sena).

The Survey Instrument

The questionnaire for the household survey had eight sections. Section 1 gathered general demographic information, confirmed whether study inclusion criteria had been met, and obtained the respondent's informed consent. Section 2 dealt with the respondent's perceptions, attitudes, and experiences related to cholera. Section 3 assessed the respondent's knowledge of vaccines and past experience with cholera vaccines. It also supplied respondents with information about the cholera vaccine and reminded them about the Esturro trial, the only cholera vaccination effort in the city, which had taken place 18 months previously. For everyone in the household who had participated in the Esturro trial, the enumerator recorded the number of doses received, vaccination outpost visited, and mode of transportation used.

In Section 4 the enumerator introduced to all respondents a hypothetical WTP scenario involving two sets of CV questions about hypothetical vaccine prices to be charged in a future vaccination program (actual vaccines in the 2003 trial were free of charge), questions which were used as the basis for the stated preference analysis reported elsewhere (Lucas, Jeuland et al., 2006). Section 5 focused on actual or anticipated travel costs during the vaccination trial, and required two different sequences of questions: one for households that reported members' having traveled to a vaccine outpost to obtain vaccines, and the other for households that did not travel. For the former, interviewers asked to see their green vaccination cards (which they had been told to keep by the health workers organizing the vaccination effort) — though some no longer had them. The interviewers then gathered information about how people had learned about the 2003 vaccination campaign, how they had traveled to the vaccination outpost, how long it took, and how long they had waited in line.

For those households that did not participate in the trial, interviewers asked whether or not they had been aware of the trial at the time it happened. Households that did not know about the trial were excluded from the travel cost analysis. Other households were excluded if the respondent explained household lack of participation as being due to journeying away from Beira, being sick or pregnant at the time of distribution, or being poorly informed about the vaccine, unsure about its safety, or thinking it was only for children or only for residents of Esturro. All other non-participating households were assumed to have made a conscious choice not to obtain vaccines based on having received full information about the trial. Finally, interviewers recorded respondent estimates of how much travel and waiting time would have been required to acquire the vaccines.

Section 6 recorded household socioeconomic characteristics. Section 7 recorded the interviewer's observations regarding visible conditions of the home and opinions on the quality of the interview. Section 8 probed the respondent's decision-making process in the CV experiment. Analysis and discussion below center on the travel cost information acquired in Section 5 of the questionnaire.

MODELING STRATEGIES

This research reports on household-level travel cost models for estimating private demand and WTP for cholera vaccinations in Beira, based on the data obtained from our surveys. We consider three model types: a standard count model, a zero-inflated count model, and finally a dichotomous choice *probit* model. An alternative zonal travel cost modeling approach using the aggregated data collected from individuals at outposts during the vaccination trial in 2003 is presented in Appendix 1.

Household-level analysis using a simple count model

The data obtained from respondents in the 2005 household survey include the number of persons (Y) that received two-dose vaccinations in each of the sampled households.[2] Y is a discrete variable whose values are nonnegative integers: 0, 1, 2, etc. Poisson and related models treat demand as a stochastic process. Such models are commonly used in travel cost applications because trips occur in indivisible quantities.

[2] Households with members who only received one dose were treated as having acquired zero two-dose vaccines unless they identified legitimate reasons for missing the second dose, in which case they were treated as having received the full regime. Legitimate reasons for missing the second dose were being away from Beira, being sick or pregnant when the second dose was given, or being poorly informed about the time and need for the second dose. Exclusion of some one-dose recipients may lead to underestimates of WTP.

Our models assume that the expected quantity demanded λ_k by any household k that was fully informed of the vaccination trial is a function of a linear combination (Z) of explanatory variables (X_{ik}), where λ is a continuous variable; specifically, $Z = \sum \beta_i X_{ik}$. The coefficients β_i of the variables are estimated by regression analysis. The continuous variable λ does not represent a feasible demand outcome (which must be integer-valued). In this sense, the model parameterizes the distribution of demand quantities given the observed realizations of the variables X_i (Haab and McConnell, 2002).

Most commonly, the relationship between the expected quantity demanded λ_k and the explanatory variables is assumed to be log-linear:

$$\lambda_k = \exp(\sum_i \beta_i X_{ik}). \qquad (1)$$

The assumption that the trips or demand quantities in the data follow a Poisson distribution typically does not hold. This statistical distribution assumes that the mean and variance of the data are equal (both equal to λ), an assumption that is commonly violated. Since the variance in our data exceeds the mean by a factor of more than two, we use the negative binomial distribution for the estimation of the household travel cost model, and parameterize the expected quantity demanded according to equation 1.

The explanatory variables in equation 1 pertain to households, some of which chose to acquire vaccines and others of which did not (recall that households not informed or misinformed of the trial were excluded). In our first specification of the household models, we hypothesize that the X vector is composed of separate variables for the total time spent traveling and queuing to acquire vaccines, and the pecuniary cost of transportation to reach the outposts, as well as income and assets, education, age group composition of the household, household health attitude variables, and variables pertaining to the manner in which the household was informed about the vaccination campaign. Though rarely used in travel cost applications because of multicollinearity problems, models that contain separate variables for both travel time and transportation expenses can be useful for expressing the value of time spent in money units. Since the coefficient of transport cost (β_{ct}) is expressed in inverse money units, and the coefficient of time spent queuing and traveling (β_{time}) is in inverse time units, we can derive a sample "average" value of time v in money units per time unit as the ratio of these two coefficients:

$$v = \beta_{time} / \beta_{ct}. \qquad (2)$$

This derived value of time pertains neither to any specific household in the sample, nor to any individual in those households, but is rather an averaged value based on the modeled impacts of the separate time and transportation variables on the household participation decision.

In our second specification of the travel cost count models, we replace the variables for distance traveled and cost of transportation by one full travel cost variable p, which is constructed as follows:

$$p = [value\ of\ travel\ time] + [value\ of\ queue\ time] + [cost\ of\ travel] \qquad (3)$$

For ease of exposition, we choose to separate p from the other X explanatory variables in equation 1 and refer to the remaining variables in the travel cost models as Ψ:

$$\lambda_k = \exp(p_k \beta_p + \sum_i \beta_i \Psi_{ik}). \qquad (4)$$

The values of the independent variables for the household model were easily obtained from our survey data, with the exception of travel cost p, which reflects the implicit cost of obtaining a vaccination to a household based on the sum of the travel costs shown in equation 3. It is the sum of the value of time spent traveling and queuing and the cost of traveling to an outpost, and is based on transportation mode decisions and the travel times experienced by households given those decisions.

Furthermore, the household models must account for the behavior of all households in the sample who knew about the 2003 trial, some of which participated, and others of which did not. However, for non-participants, actual transportation choices were not made and pecuniary transportation costs were not incurred. Additional assumptions are necessary to include such households, as described below.

For the third term of equation 3, we consider the *n* different modes of transportation available to households. Define c_{jk} as the pecuniary cost of transportation incurred by household *k* from using mode of transportation *j*, which may or may not be the same for all households. Also let $M_{jk} = 1$ if household *k* used (for participants) or reported that it would have used (for non-participants) mode of transport *j* and 0 otherwise. Then the third term of equation 3 for household *k* can be written as

$$\sum_j M_{jk} \cdot c_{jk}. \tag{5}$$

In order to test the influence on our model results of unreliability in reported experiences, we used two approaches to find the first two terms of equation 3 for households, which we term the "subjective" and "objective" travel cost approaches, respectively. The first "subjective" approach relied on the recalled experiences of participating (and non-participating) households' real (and anticipated) travel and queuing times. In other words, during the interview, participating households provided estimates of their actual travel times, t_k, and queue times, q_k. Meanwhile, non-participating households provided estimates of the time they thought they would have spent traveling $\bar{t}_k$ and queuing $\bar{q}_k$ to obtain vaccines. Let vt_k be the average unit value of travel time (e.g., dollars per hour) and vq_k be the average unit value of queue time for a person in household k; if these unit values were the same for all households, then subscripts k could be deleted. The parameters vt_k and vq_k may or may not be equal. Then

$$\textit{[value of travel time]} = t_k \cdot vt_k \text{ if household } k \text{ acquired vaccines;} \tag{6a}$$

$$\textit{[value of travel time]} = \bar{t}_k \cdot vt_k \text{ if household } k \text{ did not acquire vaccines;} \tag{6b}$$

$$\textit{[value of queue time]} = q_k \cdot vq_k \text{ if household } k \text{ acquired vaccines;} \tag{6c}$$

$$\textit{[value of queue time]} = \bar{q}_k \cdot vq_k \text{ if household } k \text{ did not acquire vaccines.} \tag{6d}$$

Thus, in the "subjective" travel cost modeling approach, if household k participated in the vaccination trial, the implicit cost it faced for the vaccine is obtained by adding terms 5, 6a and 6c:

$$p_k = t_k \cdot vt_k + q_k \cdot vq_k + \sum_j M_{jk} \cdot c_{jk}. \tag{7a}$$

Similarly, if household k did not participate, it faced the cost obtained by summing 5, 6b and 6d:

$$p_k = \bar{t}_k \cdot vt_k + \bar{q}_k \cdot vq_k + \sum_j M_{jk} \cdot c_{jk}. \tag{7b}$$

We refer to the costs in equations 7a and 7b as "subjective travel cost," because they rely on household recall and/or perceptions of time spent.

The "objective" travel cost approach for estimating the first two terms of equation 3 for households sought to avoid the problem of unreliable recall of the time required (or anticipated) for travel and queuing. First we focus on travel time. Let d_k represent the round-trip distance to a vaccination outpost for household k, and r_j denote the average reciprocal rate of travel (in time/distance) using mode j. Then $r_j \cdot d_k$ will be the round-trip time of travel to an outpost for household k using mode j. M_{jk} and vt_k are defined as above, and the value of time spent traveling for household k is

$$\sum_j M_{jk} \cdot r_j \cdot d_k \cdot vt_k. \tag{8}$$

For queue time, let $\overline{Qt}$ be the average sample-reported queue time. With vq_k defined as above, the value of queue time is

$$\overline{Qt} \cdot vq_k, \tag{9}$$

and adding terms 5, 8 and 9 yields the implicit cost to household k of obtaining a vaccine:

$$p_k = \sum_j M_{jk} \cdot r_j \cdot d_k \cdot vt_k + \overline{Qt} \cdot vq_k + \sum_j M_{jk} \cdot c_{jk}. \tag{10}$$

We refer to this cost as "objective travel cost," because it is based on non-subjective measures of distance traveled and incorporates sample average queue times, rather than reported times for travel and queuing.

Assuming the functional form for the demand equation (4) and having estimated values for the parameters, convenient expressions for the households' expected willingness to pay (WTP) for vaccines can be derived. We assume that the regression coefficient on travel cost can be applied over all prices to yield a demand function that relates vaccine quantities to prices. Then, integrating under the demand curve from a price of zero to infinity yields household k's WTP for cholera vaccines:

$$\mathrm{E}[WTP_k] = \int_0^\infty \exp(\textstyle\sum_i \beta_i \Psi_{ik} + \beta_p p)dp = -(1/\beta_p) \cdot \exp(\sum_i \beta_i \Psi_{ik}). \tag{11}$$

This WTP estimate for households is converted to per-capita WTP in the population by dividing by the average household size.

Household-level analysis – extension to the zero-inflated count model

In addition to estimation using a negative binomial distribution, we also test zero-inflated negative binomial models that seek to account statistically for the large proportion of non-participants in the vaccination campaign (zeros in our sample data set). These models require a two-stage maximum likelihood estimation procedure. In stage one, a *probit* specification is used to predict the probability of each household not participating in the vaccine trial, given the same set of explanatory variables as appear in the simple count model. In stage two, a negative binomial count model is fit to the data on vaccine quantities demanded. In functional terms, our zero-inflated negative binomial demand model can be written:

$$E[Q_k] = \begin{cases} \pi(\sum_i \beta_i X_{ik}) + \left[1 - \pi(\sum_i \beta_i X_{ik})\right] \cdot \exp(\sum_i \beta_i X_{ik}) & \text{if } y = 0, \\ \left[1 - \pi(\sum_i \beta_i X_{ik})\right] \cdot \exp(\sum_i \beta_i X_{ik}) & \text{if } y \geq 1. \end{cases} \qquad (12)$$

where $E[Q_k]$ is the expected quantity of vaccines obtained by household k, π is the probability of zero counts from the *probit* model with explanatory variables X_{ik}, and $\exp(\sum_i \beta_i X_{ik}) = \exp(\sum_i \beta_i \Psi_{ik} + \beta_p p_k)$ is the expected level of demand given p_k and Ψ_{ik}, similar to the continuous variable λ_k estimated in equation 4. Note that zero counts in this formulation can occur as a realization of the binary decision-making process or from the count process when the binary random variable takes on a value of 1.

Assuming that households do not obtain utility from non-participation in the trial, the calculation of expected WTP for household k from the zero-inflated negative binomial model must then incorporate the probability of participation as estimated in the first stage:

$$E[WTP_k] = \Pr(participation) \cdot WTP(participation) = (1-\pi) \cdot \int_0^{\infty} \exp\left(\sum_i \beta_i \Psi_{ik} + \beta_p p\right) dp = \left[(\pi - 1)/\beta_p\right] \cdot \exp\left(\sum_i \beta_i \Psi_{ik}\right). \tag{13}$$

This expression is equivalent to the WTP derived in equation 11 except that it is multiplied by $(1-\pi)$, the probability of participation.

One weakness of all the count models in this particular application is that they force the opportunity cost of time of all members of the household to be the same, which may not be very realistic, considering that the individuals in the household include infant children, school children, and more or less productive adults. In effect, most people would expect that children's time value is likely to be low. Balancing this concern, however, is the evidence from stated preference research in Beira, which suggested that demand was higher in households with greater numbers of children (Lucas, Jeuland et al., 2006).

The differential value of time within households is a particular problem in the standard count models which do not differentiate the participation decision from the decision about how many people should go to the outposts once a household has decided to participate. The zero-inflated models, besides allowing for a large number of zeros, therefore present a more plausible explanation of households' decision making process concerning the vaccination campaign, since travel cost (and time value) can influence participation and the number of participants to a differing extent.

Dichotomous choice model for head of the household

Finally, an additional model was estimated in which the primary accompanying adult's WTP in each household was estimated using a *probit* model. Of course, the interpretation of this estimated WTP amount is complicated, because it depends on the overall household composition. In other words, if the participation decision for the household head depends on the number of school children in the house, it is inappropriate to use this WTP as a pure measure of private benefits to the head of the household, since it will reflect some of the concern for those children who were brought along to the vaccination outpost.

We assumed that the accompanying adult's demand for the cholera vaccine Q_{adult} was an exponential function of the same independent variables X that were used previously in the count models for total household demand (i.e. distance/cost or travel cost, socio-economic variables, information variables, etc):

$$Q_{adult,k} = \exp(\sum_i \beta_i X_{ik}) \tag{14}$$

This demand amount Q_{adult} is a latent variable, we observed only a binary realization that was 1 if the household participated in the trial and 0 if it did not. Following estimation of the distance-cost model using a *probit* model, we were again able to derive an "average" value of time for the accompanying adult (which could be different from the value of time derived from the household models) using equation 2. The interpretation of this time value is somewhat different from that obtained in the household models. In the *probit* model, this "average" value of time is not a modeled household average, but rather a value obtained from the relative effect of time and transportation cost on the household heads' participation decisions. Using this value of time to calculate full travel costs for all

k accompanying adults in the sample, we calculated WTP similarly to equation 11, except that this WTP was for one accompanying adult only, and not the entire household.

We constructed this additional model for three reasons. First, it made empirical sense considering the fact that nearly all children under the age of 19 who traveled to the trial were accompanied by an adult. Second, one can question the realism of household models that assumed that one time value could be applied to all members of the household. We thus wanted to see if results for the head of the household were consistent with the overall household results. Finally, the individual WTP derived from this simple binary estimation problem can also be informative for comparison with per capita estimates for adults from the CVM study, in spite of the differences in interpretation mentioned at the beginning of this section.

Parameter estimation for the travel cost models

Rather than using continuous explanatory variables for income and education in the regression models, households were assigned to income quartiles (due to low income reporting rates in the survey) and education level dummy variables. The income quartiles were created from a predictive model based on asset ownership and other socioeconomic data collected in the household survey. Though the other explanatory variables were relatively easy to construct from our survey data, there were a large number of variables to be specified in constructing the travel cost variable p, as mentioned above.

We computed the subjective and objective travel cost for each household according to equations 7 and 10. The variables for reported times – t_k, $\bar{t}_k$, q_k, $\bar{q}_k$ in equation 7 – as well as for transport mode choice M_{jk} (equations 7 and 10) and travel distance d_k (equation 10) were obtained directly from survey responses. The pecuniary cost of transportation c_{jk} was assumed to be zero unless household k opted (or said it would have opted) to use public transportation to reach a vaccination outpost, for which a one-way fee of 5,000 Mts. was charged for any household in Beira, and 10,000 for households traveling from Dondo.[3]

Reciprocal travel rates r_i (in hours/km) in equation 10 were derived for each of the five available transportation modes (walking, bicycle, motorcycle, public bus, car) based on the average rates reported in the survey from households using the relevant modes. The average reported queuing time $\overline{Qt}$ for the entire sample was 83 minutes. Finally, although the value of time was allowed to vary when conducting sensitivity analysis, the unit value of travel time vt_k and the unit value of queue time vq_k were assumed to be the same, i.e. it was assumed that household members experienced no greater relative disutility from traveling than from queuing. The base case travel cost analysis used the value of time derived from the time-transport cost models for the negative binomial, zero-inflated and *probit* models, respectively.

Finally, to test the sensitivity of our willingness-to-pay estimates to this derived value of time, we conducted sensitivity analysis using a range of time values from zero to just over the full wage for the median respondent.[4] We did not use household-varying time costs due to the 1) low rates of income reporting in the survey, 2) the unobservable

---

[3] Use of private vehicles and bicycles for transportation does involve costs, so the method adopted here is likely to underestimate transportation cost. Nonetheless, as the proportion of households using such modes was low and the condition of the vehicles was unknown, we favored this conservative treatment of costs rather than transferring cost per mileage rates typically used in the United States or Europe.

[4] The median wage was 7,290 Meticais, or about US$0.30/hour ($1 US = 24,500 Mts in 2005).

nature of the true opportunity cost of time, and 3) the problem of heterogeneity in time values *within* households.

RESULTS

Sample characteristics

About 20 percent of respondents claimed to have had cases of cholera in their immediate households (Table 1), and 22 percent (302 households) reported that one or more household members had received a cholera vaccine during the Esturro trial. The aggregate number of cholera vaccine recipients among those households was 951. The majority of these (718, or 75 percent) said they had received both required doses, though far fewer (365, or 38 percent) still had the green vaccination cards to prove it. About two-fifths (41 percent) of respondent households that did not travel to Esturro did not know about the trial at the time it occurred; these households were excluded from the analysis since they could not make a choice about whether or not to participate in the campaign. Other households that were considered to have not made a choice about participation were those who thought the campaign was only for Esturro (164 households, or 15 percent of non-participants), were away from Beira at the time (35, 3.2 percent), did not have sufficient information about the campaign (32, 2.9 percent), thought the vaccine was experimental or unsafe (20, 1.8), thought it was only for children (11, 1.0), or were sick (9, 0.8) or pregnant and ineligible (3, 0.3) at the time of the vaccination effort.

Although there was no charge for the vaccine, there were a higher proportion of upper income quartile respondents in the subset of households that had obtained at least one vaccine. Similarly, we found higher vaccine coverage among households with private

Table 1: Summary statistics for households with no vaccines and with at least one cholera vaccine acquired in Esturro: mean (standard deviation), maximum/minimum range

	No one vaccinated	At least one vaccination	t-statistic (p value)
Travel cost behavior statistics			
Sample size	1,092	302	N/A
Number in household vaccinated	0	3.2 (2.0), range 1–14	N/A
Number in household receiving 2 doses	0	2.8 (2.2), range 0–14	N/A
Number of green cards in household	0	1.2 (1.8), range 0–7	N/A
Satisfied with vaccine	N/A	96%	N/A
Distance (km) to nearest outpost	9.8 (11.9), range 0.3–31	2.8 (3.9), range 0.3–31	16.6 (0.00)
Knew about trial	59%	100%	27.3 (0.00)
Excluded for lack of information or other legitimate reasons[5]	25%	N/A	N/A
Would have walked/Walked	43%	74%	10.5 (0.00)
Would have used/Used public transportation	53%	22%	11.1 (0.00)
Would have gone in private car/Went in private care	2.1%	3.3%	1.1 (0.86)
Anticipated/real 1-way travel time (minutes)	39.0 (25.1), range 0–240	30.0 (25.2), range 0–180	5.5 (0.00)
Anticipated/real queue time (minutes)	84.3 (85.1), range 0–420	80.0 (20.5), range 0–900	0.85 (0.40)
Average one dose travel cost (thousands of Mts)	11.5 (6.0)	17.8 (9.2)	14.3 (0.00)
Socioeconomic statistics			
Single head of household	16%	19%	0.96 (0.34)
Had cholera in household	20%	20%	0.18 (0.86)
Had cholera death in household	3%	2%	0.80 (0.43)
Have electricity	32%	45%	4.1 (0.00)
Have phone	36%	52%	4.8 (0.00)
Have private water	19%	31%	4.3 (0.00)
Have private toilet	24%	32%	2.6 (0.01)
Median education level	Primary	Primary	N/A
Households in income quartile 1	27%	14%	5.3 (0.00)
Income quartile 2	26%	21%	1.9 (0.06)
Income quartile 3	24%	30%	2.3 (0.02)
Income quartile 4	23%	34%	3.7 (0.00)

toilets, water connections, and utilities, usually indicators of higher levels of income and education. This distribution of vaccines to wealthier households was due in part to the

[5] Legitimate reasons for exclusion from the choice process were: thinking the vaccines were only for residents of Esturro (15%), traveling away from Beira at the time (3.2%), not having enough information about the trial (2.9%), thinking the vaccine was experimental and potentially unsafe (1.8%), thinking only children were eligible (1.0%), or being sick (0.8%) or pregnant (0.3%) and ineligible at the time of vaccination.

location of the trial in a *bairro* near central Beira, where incomes are higher and people were better informed about the trial. Figure 1 shows that over 90% of vaccine recipients came from *bairros* with borders less than 1.2 kilometers from Esturro, and 95% were from *bairros* within two kilometers of the vaccination outposts, despite the fact that those *bairros* contained only slightly over half of the city's population.

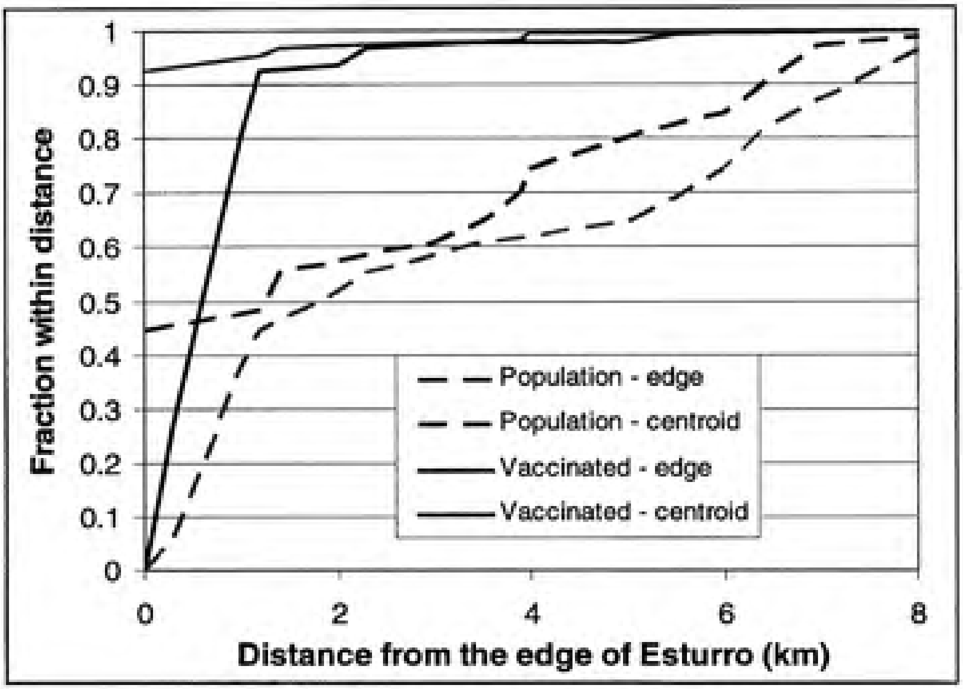

Figure 1: Cumulative distribution functions showing the percentage of population and vaccinated people living within a given distance of the edge of Esturro, measured from the edge and center of their own bairro.

Participant and non-participant households reported similar average queue times, although the standard deviation for non-participants was much higher, probably due to their greater uncertainty in estimating waiting times. There were other important differences between the two subsets of households. The average distance from the household's location to the nearest vaccine outpost for non-participant households — 9.8

kilometers — was much higher than the average distance for those that acquired vaccines (2.8 kilometers). This is partly due to the influence of including households from Dondo, since very few households traveled the roughly thirty kilometers from Dondo to Esturro. If Dondo households are excluded from the comparison, the average distance for households with no vaccinations drops to 3.4 kilometers and that for households with at least one vaccination drops to 2.3 kilometers, but the difference is still statistically significant. This higher average distance translated into higher average travel times for non-vaccinated households.

Households that did not participate were also much more likely to claim they would have used public transportation (55%) than did participants (22%). Again excluding Dondo, these percentages drop to 40% and 21%, respectively, and the differences remain statistically significant. This result largely reflects the effect of distance on transportation decisions and not inconsistency between actual and predicted behaviors: comparing households that traveled to the trial and households that did not know about the trial (thus not including those who knew about the trial and chose not to visit outposts), the actual and predicted transportation mode choices were very similar.

Of the three components that influence travel cost in the household models (distance or travel time, queuing time and transportation cost), transportation cost and distance are highly correlated (pairwise correlation coefficient $\rho = 0.81$). This result stems from the fact that people who lived far from the vaccination trial sites could not feasibly walk to the outposts and primarily took public transportation to get there. Reported total time spent traveling and queuing and distance are not as highly correlated ($\rho = 0.30$), as some people in middle-distance zones chose to walk and so spent larger amounts of time than others who came from further away. Total time and transportation cost are even less correlated (ρ = 0.12), as traveling on public transportation led to middle-range travel times, whereas a larger range of times were typically experienced by walking households (from households walking short to very long distances). This relatively low correlation provides motivation for modeling the quantity of vaccines acquired by households as a function of costs and time expenses rather than proceeding directly to full travel cost models for which time values must be assumed.

Household analysis: The determinants of demand

We investigated the factors influencing demand for vaccines using the household-level count models and the dichotomous choice model. Comparison of these models (Tables 2 and 3) generates a set of robust, statistically significant determinants of demand, though results related to the education variables in the zero-inflated models are difficult to interpret. In the *probit* portion of the zero-inflated models, the interpretation of coefficient signs is opposite to those obtained from the other specifications, because that step is explaining the probability of non-participation in the trial. Table 2 presents the distance-cost models, and Table 3 presents the full travel cost models with objective travel cost (calculated according to equation 10). We also discuss differences with the models that use subjective travel cost (using equation 7).

Table 2: Household-level models of vaccine demand with distance and cost variables[6]

Variable	Negative binomial	Zero-inflated negative binomial model		Probit [7]
		Probit [7]	Count	
Pecuniary cost of travel (thousands of Mts) [8]	-0.13*** 0.02	0.11*** 0.02	-0.04*** 0.01	-0.10*** 0.02
Total time spent in travel and queues (hr)	-0.43*** 0.10	0.41*** 0.12	-0.20** 0.08	-0.40*** 0.10
Number of adults age 19+	0.02 0.04	0.00 0.05	0.04 0.03	0.01 0.04
Number of children ages 6–18	0.22*** 0.03	-0.11** 0.05	0.21*** 0.03	0.13*** 0.04
Number of children age < 6	0.03 0.05	-0.04 0.07	0.03 0.04	0.02 0.06
Education level 2 – Primary schooling	0.14 0.22	-0.40* 0.23	-0.27* 0.14	0.33* 0.21
Education level 3 – Secondary schooling	0.23 0.23	-0.58** 0.26	-0.31* 0.17	0.44 0.23
Education level 4 – Post secondary schooling	-0.20 0.42	-1.49 1.89	-1.36** 0.60	0.18 0.39
Soap used in household	0.23* 0.13	-0.24 0.15	-0.04 0.11	0.18 0.13
Mosquito netting used	0.24* 0.12	-0.25* 0.14	0.01 0.09	0.21** 0.12
Income quartile	0.01 0.19	0.12 0.21	0.14 0.14	-0.09 0.17
Income quartile 3	-0.06 0.19	0.13 0.22	0.11 0.14	-0.06 0.18
Income quartile 4	0.06 0.20	0.14 0.25	0.17 0.15	-0.06 0.20
Own motor vehicle	-0.27 0.22	-0.58 0.42	-0.49** 0.21	0.08 0.25
Number of information sources	-0.29*** 0.08	0.38*** 0.09	-0.05 0.05	-0.34*** 0.08
Informed by health workers	0.75*** 0.12	-1.17*** 0.21	0.07 0.09	0.93*** 0.15
Constant	0.60 0.38	-0.51 0.44	1.12*** 0.26	0.40 0.38
n	641	641		641
Pseudo-R^2	0.0953	N/A		0.1805
Log-likelihood	–920.6	-860.6		-360.5
Vuong statistic [9]	N/A	6.98 (0.000)		N/A
Value of time (Mts/hr) (fraction of median hourly wage)	3242 (0.44)	3622 [10] (0.50)		3883 (0.53)

***Significant at 1% level   **Significant at 5% level   *Significant at 10% level

[6] Coefficients on first line, standard errors on second line

[7] Choice of 0 vaccines in zero-inflated probit, choice to participate in simple probit.

[8] 1,000 Mts. = US$0.04.

[9] The Vuong test is highly significant, indicating that the zero-inflated negative binomial model explains the raw data much better than the standard negative binomial model.

[10] Note that the time value is derived based on the first stage (*probit*) estimation.

Table 3: Household-level models of vaccine demand with objective travel cost and value of time derived from Table 2 [11]

Variable	Negative binomial	Zero-inflated negative binomial model		Probit[12]
		Probit[13]	Count	
Full travel cost (thousands of Mts)[13]	-0.07*** 0.01	0.06*** 0.01	-0.02*** 0.01	-0.06*** 0.01
Number of adults age 19+	0.03 0.04	0.00 0.05	0.04 0.03	0.02 0.04
Number of children ages 6–18	0.22*** 0.03	-0.10** 0.05	0.21*** 0.03	0.13*** 0.04
Number of children age < 6	0.03 0.05	-0.04 0.06	0.04 0.04	0.02 0.06
Education level 2 – Primary schooling	0.13 0.22	-0.39* 0.23	-0.27* 0.15	0.31 0.20
Education level 3 – Secondary schooling	0.22 0.23	-0.57** 0.26	-0.32* 0.17	0.43* 0.23
Education level 4 – Post secondary schooling	-0.16 0.42	-1.26 1.32	-1.27** 0.55	0.20 0.39
Soap used in household	0.24* 0.13	-0.25* 0.15	-0.05 0.11	0.19 0.13
Mosquito netting used	0.24* 0.12	-0.26* 0.14	0.01 0.09	0.21* 0.12
Income quartile	0.03 0.19	0.09 0.21	0.12 0.14	-0.06 0.17
Income quartile 3	-0.02 0.19	0.07 0.21	0.11 0.15	-0.01 0.18
Income quartile 4	0.11 0.20	0.07 0.24	0.18 0.15	0.00 0.20
Own motor vehicle	-0.25 0.22	-0.64 0.43	-0.48* 0.21	0.12 0.25
Number of information sources	-0.28*** 0.08	0.37*** 0.09	-0.04 0.05	-0.33*** 0.08
Informed by health workers	0.78*** 0.11	-1.19*** 0.21	0.08 0.09	0.95*** 0.15
Constant	0.14 0.30	-0.04 0.33	0.85*** 0.20	-0.07 0.29
n	641	641		641
Pseudo-R^2	0.0934	N/A		0.176
Log-likelihood	-922.5	-863.9		-362.3
Vuong statistic	N/A	7.01 (0.000)[14]		N/A
WTP (Mts/capita)	0.59	0.99		1.00

***Significant at 1% level **Significant at 5% level *Significant at 10% level

[11] Coefficients on first line, standard errors on second line.
[12] Choice of 0 vaccines in zero-inflated probit, choice to participate in simple probit.
[13] 1,000 Mts. = US$0.04.
[14] The Vuong test is highly significant, indicating that the zero-inflated negative binomial model explains the raw data much better than the standard negative binomial model.

The models all show that both higher pecuniary transportation cost and objective measures of time spent acquiring vaccines (obtained according to the separate components in equation 10) are highly significant and of the expected sign (1% level). Calculation of the ratio of these coefficients leads to a value of time ranging from about 3240 Meticais (Mts) per hour (US$0.13 per hour) in the negative binomial model to 3880 Mts per hour (US$0.16 per hour) in the *probit* model. Consistent with expectations, therefore, the derived time value from the simple *probit* model (which pertained only to the accompanying adult in the household) was somewhat higher than the value derived from the count models for all household members. These estimates are 0.44-0.53 of the median hourly wage recorded in the survey and are used for the base analysis of the full travel cost models. Such estimates are also consistent with findings from the travel mode choice literature (Mackie, Diaz et al., 2001).

Households with more school-age children were more likely to participate and more likely to obtain larger numbers of vaccines (adding one such child to households led to an approximately 20% increase in the number of vaccines acquired). Greater numbers of adults and young children have positive coefficients but these are not statistically significant. Considering that the trial happened 18 months prior to the household survey, any child under 2.5 years of age (about half of the age group) would have been ineligible for vaccination at the time, which would greatly diminish the effect of the number of young children. The risk aversion proxy measures (use of mosquito nets, presence of soap in the house) were weakly linked (at 10% level of significance) to higher demand for vaccines, and the zero-inflated model suggests that these behaviors only affected the participation decision and not the number of vaccines acquired. Households that were directly informed of the vaccination trial by visits from health workers (rather than public announcements, etc.) were much more likely to participate in the trial (at 1% level of significance).

In the standard negative binomial models, level of education is not a statistically significant determinant of vaccine demand, though the higher education levels generally have positive coefficients. The zero-inflated models suggest that higher levels of education are significantly (at the 5% level) related to higher probability of participation in the vaccine trial. Income did not have a statistically significant effect on demand (which is not particularly surprising since the vaccines were given free of charge). This result suggests that even perfectly known, household-specific wages would be a poor approximation for the opportunity cost of time, or that income is highly correlated in space with other variables such as education or distance from outposts.

The full travel cost models suggest that an additional 1000 Mts (US$0.04) spent acquiring vaccines reduces quantity demanded by 6-8 percent. These effects are based on models that use the objective travel cost measure, but using subjective travel cost proved troublesome. In effect, relying on the latter measure led to the time variable being insignificant in the distance-cost models, which in turn made it impossible to derive a time value. Recall that reported queue times varied widely among non-participants in the trial (Table 1); this wide variation clearly created problems during estimation. Assuming that the value of the subjective time was similar to that derived from the objective time models did not help much either. In the zero-inflated model, for instance, the travel cost variable given such assumptions was statistically significant (p-value = 0.00) in raising the probability of non-participation in the trial, as expected, but was not statistically significant in lowering the number of vaccines acquired (though it had the expected sign). This led to unstable welfare estimates, since the travel cost coefficients were unstable. It also provides additional motivation for testing the sensitivity of welfare estimates to the derived value of time.

Household analysis: benefit measures

We used the travel cost models presented in Table 3 to calculate average per-capita WTP for the cholera vaccine offered in the trial. The negative binomial model produced estimates of 14,450 Mts (US$0.59), but the travel cost coefficient appeared to be biased due to the inability of the count model to properly fit the many zeros in the data. The zero-inflated negative binomial WTP estimates are higher than the standard negative binomial model estimates, and we believe that the two-stage model better explains the data structure of the problem. Calculated on that basis, WTP is US$6.01 per household, yielding per capita estimates of US$0.99. These estimates are somewhat lower than per capita estimates obtained from the CV survey, which were US$1.40 (Lucas, Jeuland et al., 2006). The zero-inflated negative binomial travel cost results are consistent with WTP estimates using a zonal travel cost methodology (see Appendix).

The *probit* model for the household head's decision yields a WTP estimate of US$1.00. This WTP estimate can be interpreted as the average demand among household heads who might accompany other family members to the trial, but it does also reflect household demand to the extent that adult's decisions to participate were influenced by the presence of children or other family members. It is not substantially different from WTP estimates derived from CV survey responses for adults in Beira, which were

US$1.20 on average. Also, if WTP for vaccines for children and other household members is similar to the household head's demand for vaccines, the average household WTP would be about 6 times this amount (since household size was 6.02 people on average), which is similar to the result from the zero-inflated negative binomial model. There are good reasons why these might diverge, though, since the household head's decision is in part based on household composition, and based on evidence from the CVM study that demand for vaccines for children is higher than for adults.

Exploring value of time assumptions

In order to better understand the sensitivity of the WTP estimates to time values in the various modeling approaches, the assumed value of time for households was varied from zero to the full median wage (Figure 2).

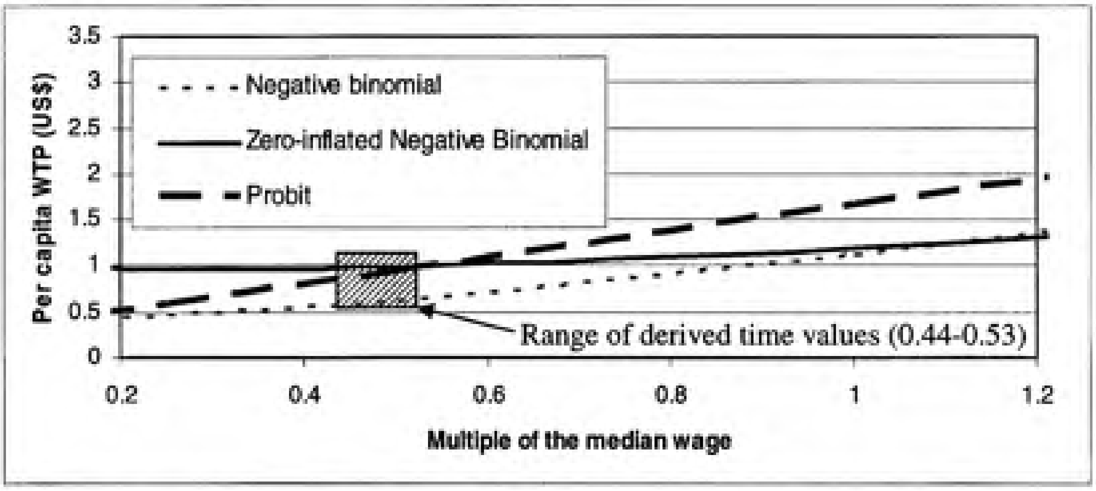

Figure 2: Sensitivity of WTP estimates from different models to the assumed value of time

Assuming a zero value of time for all households and age groups yields an absolute lower bound for the WTP estimates, since it only allows transportation expenses to enter the equation for travel cost. The lower bound per capita WTP estimates are US$0.22, $0.35, $0.97 for the *probit*, negative binomial and zero-inflated negative binomial, respectively.

Several interesting observations emerge from this analysis. First, WTP derived from the count models increases more slowly as a function of the assumed value of time than WTP obtained from the *probit* model, particularly for the zero-inflated model. As can be seen in Figure 2, the zero-inflated one predicts the highest WTP at low time values and the lowest WTP at high time values. The reason this model is less sensitive to the assumed value of time is that the majority of the effect of travel cost goes into explaining the probability of household participation. The predicted percentage of participating households changes very little as a function of travel cost – from 47 to 46.7 percent over the range of time values. Therefore, the travel cost coefficient is forced to adjust to reflect the change in the assumed value of the variable. The smaller coefficient on travel cost in the count estimation does not need to change as much as in the other two models to explain the choice of number of vaccines, and this in turn reduces the effect of travel cost on the WTP calculation.

Second, for the range of time values that appear to be most reasonable (a quarter of the median wage to half median wage), the WTP estimates from the three models fall between US$0.50 and US$1. Considering this sensitivity analysis, the US$1 per capita estimates obtained at the derived value of time seem quite reasonable.

DISCUSSION

The revealed preference (travel cost) data from our study in Beira suggest that the best estimate of average individual and household willingness-to-pay for cholera vaccines in 2005 was roughly US$1 per capita and US$6 per household. In our studies, private vaccine demand decreases as the cost of vaccination to the household goes up, whether expressed as a user fee for vaccination (as described in the CVM study in Lucas, Jeuland et al., 2006) or as a travel cost. The influence of travel cost on vaccine demand is not a surprising result; economic theory and common sense would suggest that people are less likely to participate in a vaccination campaign if they have to travel long distances and engage in costly behaviors to acquire the vaccines. These results suggest that plans for mass vaccination campaigns that aim to cut costs by providing fewer vaccination outposts with lower staffing levels are likely to significantly reduce participation. In a context like Beira, it appears that very few people were willing to travel for vaccines once distances increased beyond two kilometers. Likewise, increases in waiting times also greatly reduced demand (1 hour of additional time spent to obtain vaccines translated to a 40% reduction in quantity of vaccines demanded).

Both our CV survey analysis and our travel cost analysis show that household demand is positively influenced by higher numbers of school-age children and higher household levels of education. The stated preference results from the CV survey (Lucas, Jeuland et al., 2006) suggest average household WTP of about US$8 for a household with 6 members (US$1.40 per capita, around $1.20 for adults and $2 for school-age children).

Estimates from our revealed analyses are slightly lower than the CV survey results. WTP estimates based on zero-inflated individual-level household travel cost models ranged from US$1 to $1.25 per capita, depending upon whether time was valued at a derived time value or the full median reported wage. From these results, WTP estimates of about US$1 per capita seem plausible and conservative. WTP estimates obtained from a *probit* model (US$1) aimed at explaining adult demand for vaccines were very similar to the CV estimates for adults.

A striking feature of our raw travel cost data is that the households with at least one vaccinated individual were on average wealthier than the non-participating ones (Table 1). The multivariate models show that higher income did not in fact increase demand, but that this effect was due to the wealthier households being located closer to the vaccination outposts. Public health planners might assume that the free provision of cholera vaccines would enable poorer households to participate. Our results suggest that poor households were often unable to spend the necessary time and money to go to receive vaccinations, simply because they were located farther away from outposts. And they were generally less well informed about the trial in the first place: only about 55% of the respondents in the lowest two income quartiles of our survey reported having known about the trial at the time it occurred, whereas more than 70% of households in the top two income quartiles were aware of the vaccine distribution. It is thus not sufficient to provide "free" vaccines to ensure that the poor will be served. The number and location of benefit delivery sites must be carefully planned to keep travel and queuing costs to poor households low and to enable relevant information to reach them. This principle is well established in immunization program design in many countries, where distribution of vaccines through community outposts and mobile units is considered standard practice.

REFERENCES

Brookshire, D. S. Valuing Public Goods: A Comparison of Survey and Hedonic Approaches. American Economic Review 1982; 72: 1; 165-177.

Brown, R. W. and R. T. Jewell. The impact of provider availability on abortion demand. Contemporary Economic Policy 1996; 14: 2; 95-106.

Canh, D. G., D. Whittington, L. T. K. Thoa, N. Utomo, N. T. Hoa, et al. Household demand for typhoid fever vaccines in Hue City, Vietnam. Health Policy and Planning 2006; 21: 241-255.

Carson, R. T. Contingent Valuation and Revealed Preference Methodologies: Comparing the Estimates for Quasi-Public Goods. Land Economics 1996; 72: 1; 80-99.

Clarke, P. M. Cost-benefit analysis and mammographic screening: a travel cost approach. Journal of Health Economics 1998; 17: 6; 767-787.

Clarke, P. M. Testing the convergent validity of the contingent valuation and travel cost methods in valuing the benefits of health care. Health Economics 2002; 11: 2; 117-127.

Cook, J., D. Whittington, D. G. Canh, F. R. Johnson and A. Nyamete. Reliability of stated preferences for cholera and typhoid vaccines with time to think in Hue, Vietnam. Economic Inquiry 2007; 45: 1; 100-114.

Cropper, M. L., M. Haile, J. Lampietti, C. Poulos and D. Whittington. The demand for a malaria vaccine: evidence from Ethiopia. Journal of Development Economics 2004; 75: 1; 303-318.

Deen, J. Coverage of mass cholera immunization in Beira, Mozambique, December 2003 to January 2004. Unpublished project document 2004.

Fischer, T. K., D. D. Anh, L. Antil, N. D. L. Cat, P. E. Kilgore, et al. Health care costs of diarrheal disease and estimates of the cost-effectiveness of rotavirus vaccination in Vietnam. Journal of Infectious Diseases 2005; 192: 10; 1720-1727.

Freeman, A. M., III. (2003). The measurement of environmental and resource values: Theory and methods. Washington, D.C., Resources for the Future.

Haab, T. C. and K. E. McConnell (2002). Valuing environmental and natural resources: The econometrics of non-market valuation. Cheltenham, U.K. and Northampton, Mass., Elgar; distributed by American International Distribution Corporation, Williston, Vt.

Harris, K., J. Schultz and R. Feldman. Measuring Consumer Perceptions of Quality Differences among Competing Health Benefit Plans. Journal of Health Economics 2002; 21: 1; 1-17.

Harris, K. M. and M. P. Keane. A Model of Health Plan Choice: Inferring Preferences and Perceptions from a Combination of Revealed Preference and Attitudinal Data. Journal of Econometrics 1999; 89: 1-2; 131-157.

INE (1998). Second National Census. Maputo, Mozambique.

Jefferson, T. Do vaccines make best use of available resources? (in other words are they cost effective?). Vaccine 1999; 17: 69-73.

Kim, D., D. G. Canh, C. Poulos, L. T. K. Thoa, J. Cook, et al. Private demand for cholera vaccines in Hue,Vietnam. Forthcoming in Value in Health 2007.

Leonard, K. L., G. R. Mliga and D. H. Mariam. Bypassing Health Centres in Tanzania: Revealed Preferences for Quality. Journal of African Economies 2002; 11: 4; 441-471.

Lucas, M. The burden and etiology of diarrheal disease in Beira, Mozambique: Application for review of research using human subjects. Unpublished draft 2005.

Lucas, M., J. L. Deen, L. Von Seidlein, X.-Y. Wang, J. Ampuero, et al. Effectiveness of Mass Oral Cholera Vaccination in Beira, Mozambique. New England Journal of Medicine 2005; 352: 8; 757-767.

Lucas, M., M. Jeuland, J. Deen, N. Lazaro, M. MacMahon, et al. Private Demand for Cholera Vaccines in Beira, Mozambique. Vaccine 2006; 25: 14; 2599-2609.

Mackie, P., Jara-Diaz, S., Fowkes, A. The value of travel time savings in evaluation. Transportation Research 2001; 37:91-106.

Whittington, D., O. Matsui-Santana, J. J. Freiberger, G. Van Houtven and S. Pattanayak. Private demand for a HIV/AIDS vaccine: evidence from Guadalajara, Mexico. Vaccine 2002; 20: 19/20; 2585-2591.

WHO. Mozambique mass campaign tests the theory: Vaccines could limit the impact of cholera. WHO Week in Review 2004.

APPENDIX: TRAVEL COST ANALYSIS OF THE AGGREGATE DATA

The aggregate data

During the vaccine trial in 2003, the International Vaccine Institute and the staff of the Centre for Environmental Hygiene and Medical Examination in Beira established a database of individuals who received one or both doses of the cholera vaccine. Along with vaccine administration details and dates, the basic information recorded at vaccination outposts included the participant's name, age, sex, *bairro* address, and father's name. Beneficiaries received green vaccination cards with this information, which they were told to keep and present upon arrival for the second of the two doses to expedite the vaccination process. Addresses were only consistently recorded for residents of the town of Beira itself, and did not include streets or even neighborhood unit descriptors smaller than *bairros*.

Zonal Demand Modeling

Zonal travel cost models require researchers to divide the spatial domain of interest into a series of zones at varying distances from the site under investigation and to analyze the variation in population visitation rate across zones. The city of Beira is comprised of 22 *bairros* (Figure A1), each with a distinct name, boundaries, and known population. Since our data on all vaccine trial participants are identified at the *bairro* level, we define each *bairro* as a zone.

The fraction of the population of each zone (*bairro*) that was vaccinated with two doses was calculated from the data collected at outposts:

$$V_j = Q_j / Pop_j, \qquad \text{(A1)}$$

where V_j is the visitation rate, or fraction of population in zone j that was vaccinated with two doses, Pop_j is the population of zone j and Q_j is the number of persons from zone j who were vaccinated. The values V_j vary from one zone to another, for example, zones located far from the outposts tended to have lower values of V_j than those closer to the outposts, presumably in part because it was more costly or difficult for people to get to the outposts, or because they were poorly informed about the campaign, or for other reasons.

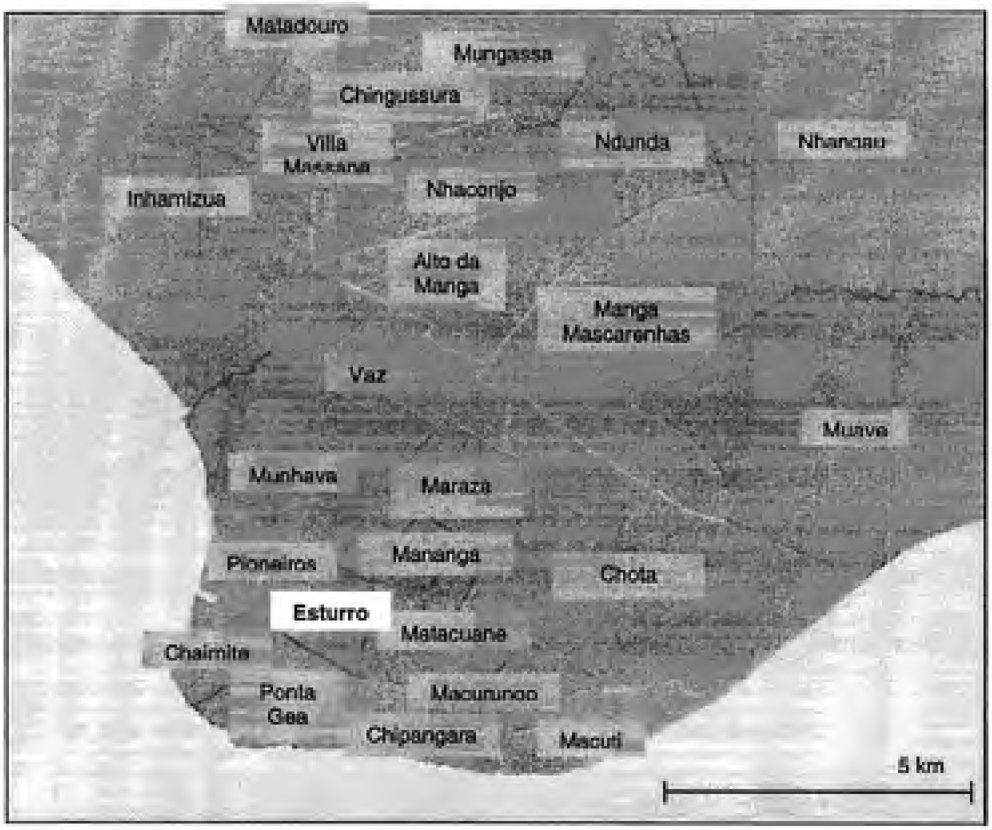

Figure A1: Map of Beira, depicting the vaccination bairro (Esturro) and other neighborhoods

The first modeling task is to explain the variation in the visitation rates, which was done by fitting a regression model to $\{V_j\}$, the set of values for the dependent variable, and $\{Z_j\}$, a set of zonal values for a vector of variables hypothesized to explain V_j. Z is comprised of several explanatory variables, including diarrhea incidence in each zone (I); the fraction of population in each zone that knew about the vaccination campaign (K), median household income in each zone (Y), a dummy variable b that is 1 for the two zones containing outposts and 0 otherwise to reflect differences in the way the campaign was promoted in these target zones, and the average cost of traveling TC from each zone to an outpost. The regression model is shown in equation A2, where $Z = [I, K, Y, b, TC]$, and ε is the portion of V not explained by Z.

$$V = f(Z, \varepsilon). \tag{A2}$$

We also consider one regression model in which the travel cost variable TC in Z is replaced with travel distance d.

We further develop an extension of the zonal model in equation A2 that seeks to account for the difference in travel cost across age groups from the same *bairros*. In this extended model, rather than using one visitation rate V_j for each zone, we define visitation rates V_{kj} that apply to the subset of people in age group k and zone j. We use three age groups: children under age five $(k=1)$, school-aged children 5-18 years old $(k=2)$, and adults $(k=3)$. The explanatory variable TC_{kj} then varies across age groups k and zones j, rather than simply zones j. In these age-group models, we also include dummy variables C_j and S_j to account for differences across age groups in the visitation rates V_{kj} that are not explained by the other explanatory variables in the model. C_j is equal to 1 for all observations with $k=1$, and 0 otherwise, and S_j is equal to 1 for all observations with $k=2$, and 0 otherwise. In this formulation of the zonal model, vector $Z = [I, K, Y, b, TC, S, C]$.

Obtaining numerical values for the variables in Z for each zone was straightforward with the exception of TC, the average value of travel cost, defined as the sum of the value of time spent and the cost of transportation incurred to receive vaccines, as shown in equation A3.

$$TC = [\textit{value of travel time}] + [\textit{value of queue time}] + [\textit{cost of travel}] \qquad \text{(A3)}$$

To calculate the value of travel time for any zone, we had to consider the n different modes of travel (e.g. public transport, walking, etc.) available to vaccine recipients. Let d_j represent the round-trip distance from zone j to a vaccination outpost, and r_i be the average rate of travel (velocity) using mode i. Then $r_i \cdot d_j$ is the round-trip time of travel from zone j to an outpost using mode i. If s_{ij} is the fraction of recipients from zone j that used travel mode i, then $\sum_{i=1}^{n} s_{ij} \cdot r_i \cdot d_j$ is the average time spent traveling for one dose of the vaccine by a typical (average) recipient in zone j. Let vt_{kj} denote the average unit value of travel time for a person in age group k from zone j (e.g., dollars per hour). If this unit value were assumed to be the same for all age groups (or zones), then subscript k (or j) could be deleted.

Hence

$$\sum_{i=1}^{n} s_{ij} \cdot r_i \cdot d_j \cdot vt_{kj}. \qquad \text{(A4)}$$

is the average value of time in monetary units spent by a recipient in age group k from zone j traveling to get vaccinated.

The estimation of the value of queue time, the second term in equation A3, begins with times reported in the household survey. Let t_q be the average reported queue time of all recipients in the survey sample, without distinguishing the zones from which they came. Let vq_{kj} denote the average unit value of queue time for a person in age group k from zone j (e.g., dollars per hour), which may or may not be the same as vt_{kj} and which may or may not be different from one age group (or *bairro*) to another. Hence,

$$t_q \cdot vq_{kj} \tag{A5}$$

is the average value of time spent waiting for one dose by a vaccine recipient in age group k from zone j, in monetary units.

The final term in equation A3 represents the average pecuniary cost of traveling. Let c_{ij} denote the pecuniary cost of round trip travel from zone j to an outpost using mode i. Recall that s_{ij} is the fraction of all the vaccine recipients from zone j that used travel mode i. Hence,

$$\sum_{i=1}^{n} s_{ij} \cdot c_{ij} \tag{A6}$$

is the average pecuniary cost incurred by a vaccine recipient in zone j traveling to an outpost to get vaccinated, and is assumed to be constant across age groups. Thus, the average total travel cost (TC_{kj}) for a typical vaccine recipient in age group k from any zone j in monetary units is the sum of equations A4, A5 and A6, which is shown in equation A7:

$$TC_{kj} = \sum_{i=1}^{n} s_{ij}\ r_i\ d_j\ vt_{kj} + t_q\ vq_{kj} + \sum_{i=1}^{n} s_{ij}\ c_{ij}. \tag{A7}$$

With data for all the explanatory variables in Z and data for the dependent variable V in all zones, the model in equation A2 was fitted using OLS. Several different functional forms were employed and are presented in the upper portion of Table A2.

The next step in the analysis was to predict vaccine demand at various prices and derive welfare estimates. Using the fitted equation from the regression analysis, we predicted the fraction of persons vaccinated in age group k and zone j: $\hat{V}_{kj} = f(Z_{kj})$. This prediction is based on the fact that recipients in zone j were provided vaccinations free of charge (at an implicit price $P_{kj} = TC_{kj}$, for all age groups k and zones j). They therefore "paid" only TC_{kj} for vaccination. Thus, $\hat{V}_{kj} \cdot Pop_{kj}$ is the predicted quantity of vaccinations demanded by persons in age group k and zone j given that users paid TC_{kj}. Thus, $[\hat{V}_{kj} \cdot Pop_{kj}, TC_{kj}]$ is one point on the demand function for age group k and zone j, at a vaccine user fee of zero. If users were charged some positive amount, it would be possible to estimate the demand function for age group k and zone j, provided that the fee and TC had the same influence on quantity demanded. Thus, in zonal travel cost demand models, the same estimated coefficient of TC is assumed to apply to a complete (travel + user fee) price $P = TC + f$, where f is the user fee and P can range from 0 to infinity. In the context of the vaccine trial, a price of zero was not a feasible outcome, since all zones had positive travel costs TC. Hence, the demand model for vaccinations of people living in zone j is

$$Q_{kj} = Pop_{kj} \, V_{kj} = Pop_{kj} \, f(\Psi_{kj}, P_{kj}), \tag{A8}$$

where Q_{kj} denotes the quantity of vaccinations demanded by people in age group k and zone j, and $f(.)$ is the fitted regression model that includes all explanatory variables Ψ_{kj} in Z_{kj} for age group k and zone j except TC_{kj}, and TC_{kj} is replaced with P_{kj}. Inserting different values for P results in unique predictions of Q_{kj} for age group k and zone j. A separate demand function like equation A8 was developed for each zone from which people traveled to be vaccinated.

We next define the WTP, for one dose of the vaccine for people in age group k and zone j as the area under the zonal demand curve, from a price of zero to infinity:

$$WTP_{kj} = \int_0^{\infty} \left[Pop_{kj} \cdot f(\Psi_{kj}, P_{kj})\right] dP_{kj}. \tag{A9}$$

Summing over all of the age groups and zones, total WTP for one vaccine dose during the Beira trial is:

$$WTP = \sum_j \sum_k \int_0^{\infty} \left[Pop_{kj} \cdot f(\Psi_{kj}, P_{kj})\right] dP_{kj}. \tag{A10}$$

For example, using the non-age group specific log-linear functional form (B1) in Table A2 leads to an analytical form of *WTP* that is similar to that described in the household modeling problem, but with several important differences. In the household approach, the quantity demanded was modeled for all households in the sample, and was treated as a stochastic variable. As such, the explanatory variables in equation 1 of this paper pertained to households, some of which chose to acquire vaccines and others of which did not. In the zonal approach, a separate equation for the quantity of vaccinations demanded is developed for each zone, and the quantities are treated as deterministic continuous variables, which are explained by a small number of zone-specific variables.

Substituting functional form B1 with estimated parameters into equation A10 leads to

$$WTP = \sum_j Pop_j \int_0^{\infty} \left[\exp(\hat{\beta}_0 + \hat{\beta}_1 P_j + \hat{\beta}_2 I_j + \hat{\beta}_3 K_j) \right] dP_j, \tag{A11}$$

which, upon integration, yields equation A12, the population's willingness to pay for each dose of the vaccine:

$$WTP = -\left[\exp(\hat{\beta}_0) / \hat{\beta}_1 \right] \cdot \sum_j \left[Pop_j \cdot \exp(\hat{\beta}_2 \, Inc_j + \hat{\beta}_3 K_j) \right]. \tag{A12}$$

We convert this to per capita WTP for the two-dose coverage (the private benefits of protection against cholera with this particular vaccine) by multiplying equation A12 by 2 and then dividing by the total population of Beira. In fact, this may underestimate WTP since about one fifth of first-dose recipients did not return for the second dose and are not included in the analysis. We take this approach because the aggregate data (unlike the more detailed household data which allowed us to correct for legitimate absences) do not allow us to know why these people did not return, whether for lack of information, dissatisfaction with the first dose, or other reasons.

Parameter estimation for the zonal travel cost models

One limitation unique to the zonal (as opposed to household) modeling analysis in this application was the difficulty in finding adequate explanatory variables *I*, *K* and *Y* to use for the 22 zones in equation A2. We discuss the composition of these variables in this section, and then explain the calculation of travel cost *TC*. Most of these variables are shown in Table A1. For zonal cholera incidence *I*, we used diarrheal incidence rates measured at the cholera treatment center (a proxy for cholera incidence). This proxy is problematic for two reasons. First, the epidemiology of cholera disease is not particularly well represented by other diarrheal incidence rates, and second, the CTC recorded rates are likely dependent on a selection process which encourages higher visitation from patients living near the facility (in central Beira).

Knowledge of the vaccination trial *K* was calculated for each zone in the household survey data that includes respondents who did and did not participate in the campaign. Unfortunately, surveys were only conducted in 9 zones, so information rates for the other zones had to be transferred from similar ones in the sample. Income *Y* suffered from a similar problem, as no *bairro* level statistics could be obtained. Furthermore, because households were often reluctant to disclose income, we resorted to using the predicted income quartile (recall that income quartiles were also used in the household models because of this limitation in the data) of the median respondent from a zone rather than some continuous measure of income. In general, zonal model performance can be improved if something is known about the heterogeneity of individuals within zones, but we did not have sufficient information to allow for such extensions.

Table A1: Data for zonal regression models, by bairro (sorted by distance from Esturro)

Bairro	Population[15]	# of Vaccines	Distance to Esturro (km)		Travel cost[16] ('000 of Mts)	Diarrheal incidence[17]	% HHs informed[18]
			Edge	*Center*			
Esturro	27,804	11,175	0	0.3	5.93	15.3	95
Pioneiros	8,595	1,521	0	1	5.93	18.2	80
Mananga	22,821	1,084	0	1	6.91	17.1	75
Munhava	36,624	5,138	0	1	7.27	14.7	67
Chaimite	17,671	921	0	1	7.78	12.9	70
Matacuane	33,847	9,286	0	1	5.93	12.2	86
Ponta Gea	28,139	3,943	0	1	7.39	9.7	83
Chipangara	29,628	4,863	0	1.2	7.06	11.2	76
Maraza/Chota	33,217	511	1.4	2	9.37	10.3	68
Macurungo	17,033	1,235	1.2	2.3	11.30	6.1	80
Vaz	7,223	126	3	3.5	13.89	13.2	60
Macuti	16,417	346	2.4	3.5	12.71	4.9	74
Inhamizua	18,167	53	3.5	5	17.00	5.6	50
Alto da Manga	20,264	590	4	5.5	13.14	4.4	60
Manga Mascarenhas	24,613	29	3.9	6	17.28	6.6	56
Muave	8,442	32	5.5	6.5	14.19	8.5	60
Nhaconjo	29,365	30	5.1	6.5	18.12	4.8	50
Chingussura	26,374	64	7	8	19.45	7.1	50
Ndunda	8,947	4	6	8	18.75	6.3	50
Villa Massane	26,561	51	7	8	19.45	5.7	50
Mungassa	4,593	1	7	9	19.45	3.7	50
Matadouro	13,416	8	9	14	20.86	6.9	50
Total	**459,761**	**41,011**					

Household surveys were conducted in the neighborhoods in **bold**

We used equation A7 to compute average zonal travel cost *TC*. For the first term, the value of travel time, we started with travel rates for the *i*th mode of transportation r_i.

[15] Population projected from Census (1997) at about 2.8% per annum rate of increase.
[16] Travel cost for neighborhoods obtained using equation A7 and edge–to–edge distance to Esturro.
[17] Incidence of severe diarrhea per thousand people, measured by visitation rates to the Cholera Treatment Center in Beira.
[18] % reporting they knew about the trial in the survey; for neighborhoods that were not part of the survey, similarly distant neighborhood rates were averaged, except for those furthest away, where no surveys were done, and for which 50% is assumed, noting that about 47% of people had heard of the trial in the neighboring town of Dondo.

These rates were calculated as in the household analysis, and weighted by s_{ij}, the proportion of people from zone j who reported using mode i. As with income and knowledge of the trial, for zones not included in the household survey, the s_{ij} weights had to be transferred from similar zones. Next, because the vaccination outposts were located around the edges of Esturro, one set of models used distance d_j (in kilometers) from the edge of each zone to the edge of Esturro; the other used the zonal centroid to the edge of Esturro. In reality the true average travel distance was probably somewhere between these two distances, since there would have been higher participation rates from households nearest the outposts, on the edges nearest the vaccination zone. Finally, in the basic zonal model, we used half the median wage for the value of travel time vt_j (where subscript k is omitted because the model contained no age groups), which is similar to the time values derived from the household models.

To value queue time (term two in equation A7), we used the average reported queuing time t_q for the entire sample (83 minutes), obtained from the survey, multiplied by $vq_{kj} = \ vt_{kj}$, which assumes that the value of time spent queuing and traveling was equivalent. Finally, for the third term in equation A7 (the cost of travel), we weighted the cost of transportation c_{ij} of taking mode i from zone j by the same proportion s_{ij} of people from zone j who reported using mode i as we used to determine average travel times. Transportation modes were assumed to have zero monetary cost c_{ij} except for public transportation, for which there is a uniform city-wide, one-way fare of 5,000 Mts (about US$0.20) within Beira. The *bairro*-specific transportation cost thus depended proportionately upon the degree to which residents opted for public transportation among other available choices.

In the age group models, we computed travel cost based on a different assumed value of time $vq_{kj} = \ vt_{kj}$ for each of three age groups – young children under age 5, school-aged children under 18 and adults – in each zone. In these models, we imposed a time value $vq_{1j} = 0$ for young children, and $vq_{2j} = (vq_{3j}/2)$, i.e. half the value of adults' time for school children. In the base model children's time value was thus one quarter of the median wage.

Finally, as with the household models, to test the sensitivity of our willingness-to-pay estimates to this assumed value of time, we conducted sensitivity analysis using a range of time values spanning from zero to twice the full wage for the median respondent. In the age group models, the value of time assumed for adults also ranged from zero to twice the median wage. Thus, for school children time value in the sensitivity analysis ranged from zero to the median wage, and young children were always assumed to have zero value of time.

Zonal model estimation results

The second particular limitation of the zonal travel cost models in this application arises during estimation and is due to the problem of multicollinearity between the explanatory variables. The models in Table A2 were constructed to minimize such problems while still retaining information about important asymmetries in knowledge of the campaign across zones. Those models thus favor the vaccination bairro dummy variable b (which can be interpreted to represent the additional publicity efforts and door-to-door campaigns undertaken in targeted zones) to the explanatory variable K, which causes more coefficient instability. And while model coefficients for the different functional form specifications in Table A2 cannot be readily compared, the models generally show that the percentage of the population that received the cholera vaccine decreased with increasing distance and travel cost.

Table A2: Regression models for alternative specifications of population percentage vaccinated[19]

Functional forms tested

Distance:	$V_j = \beta_0 + \beta_1 d_j + \beta_2 I_j + \beta_4 b_j$
A:	$V_j = \beta_0 + \beta_1 TC_j + \beta_2 I_j + \beta_4 b_j$
B:	$ln(V_j) = \beta_0 + \beta_1 TC_j + \beta_2 I_j + \beta_4 b_j$
B1^c:	$ln(V_j) = \beta_0 + \beta_1 TC_j + \beta_2 I_j + \beta_3 K_j$
B2^c:	$ln(V_j) = \beta_0 + \beta_1 TC_j + \beta_2 I_j$ (excluding the two bairros with $b_j = 1$)
B3^c:	$ln(V_j) = \beta_0 + \beta_1 TC_j + \beta_2 I_j + \beta_4 b_j + \beta_5 Y_j$
B4^c:	$ln(V_{kj}) = \beta_0 + \beta_1 TC_{kj} + \beta_2 I_j + \beta_3 K_j + \beta_5 Y_j + + \beta_6 C_j + \beta_7 S_j$[20]
C:	$ln(V_j) = \beta_0 + \beta_1 ln(TC_j) + \beta_2 I_j + \beta_4 b_j$
D:	$V_j = \beta_0 + \beta_1 ln(TC_j) + \beta_2 I_j + \beta_4 b_j$
E:	$V_j = \beta_0 + \beta_1 /TC_j + \beta_2 I_j + \beta_4 b_j$

Variable	Distance		A	B[21]	C	D	E
	Edge	Centroid					
Travel Cost *TC* (thousands of Mts)			−0.81** (0.29)	**−0.42*** (0.05)**	−5.4*** (0.78)	−12.0*** (3.5)	154*** (38.0)
Distance (km)	-1.1** (0.47)	-0.77 (0.39)					
Diarrheal incidence *I* (cases/1000 people)	0.43 (0.31)	0.53 (0.32)	0.21 (0.34)	**−0.05 (0.06)**	−0.09 (0.08)	0.0 (0.3)	−0.22 (0.34)
Vaccination *bairro b* (=1 for Esturro/ Matacuane)	23.5*** (3.5)	23.7*** (3.6)	22.3*** (3.4)	**0.61 (0.64)**	0.03 (0.74)	20.6*** (3.3)	18.1*** (3.3)
Constant	4.3 (4.1)	3.4 (4.3)	13.3** (6.5)	**6.1*** (1.2)**	14.3*** (2.5)	34.7*** (11.4)	−7.3*** (2.1)
Adjusted R^2	0.825	0.811	0.841	**0.885**	0.878	0.861	0.880

***Significance at 1% level **Significant at 5% level *Significant at 10% level

Table A2 presents two distance-only models, in which the full travel cost variable *TC* is replaced with one of its components, the distance variable *d* (only one piece of

[19] Travel time, based on distance (km) from edge of *bairro* to edge of Esturro *bairro*; Standard errors are shown in parentheses.
[20] *C* and *D* are age group dummy variables, $C_j = 1$ if $k=1$and 0 otherwise, where V_{1j} is the visitation rate for young children from *bairro j*, and $S_j = 1$ if $k=2$and 0 otherwise, where V_{2j} is the visitation rate for school children from *bairro j* (V_{3j} is the visitation rate for adults from *bairro j*).
[21] The results of the alternative specifications of model B are shown in Table A2

travel cost in equation A7), which is alternatively a measure of the distance (in km) from the edge of zone j to the edge of Esturro, or from the center of zone j to the edge of Esturro. Unlike the household models, time and pecuniary cost of transportation are too highly correlated in the zonal application to allow inclusion of both in one model; therefore, these models do not permit calculation of an implicit time value based on the relative size of the regression coefficients. The "edge-to-edge" model gives a slightly better fit than the "center-to-edge" model and is used for the other travel cost models.

In addition, residents in one of the two vaccination outpost *bairros* (Esturro and Matacuane) were more likely to be vaccinated. Due to high correlation of this dummy variable with travel cost (ρ = -0.41), this effect is not statistically significant in all specifications. Still, as mentioned above, the correlation of b and TC caused fewer problems than some other explanatory variables: a) percent of households knowing about the campaign K and travel cost TC (ρ = -0.91) and b) income quartile of the median household Y and travel cost TC (ρ = -0.64). Based on the model results and fits presented in Table A2, and because of its functional similarity to the household models, we chose to analyze the demand for vaccine in more detail using the log – linear travel cost specification.

Figure A2 compares the actual and predicted number of vaccines from Model B1 in Table A3. In the base case (value of time equal to half the wage of the median respondent), the estimated per capita WTP for two doses with this specification is US$1.14; for a model that uses the vaccination bairro dummy instead of the percent informed variable (Table A2 Model B), the estimate of US$1.09 is a bit lower.

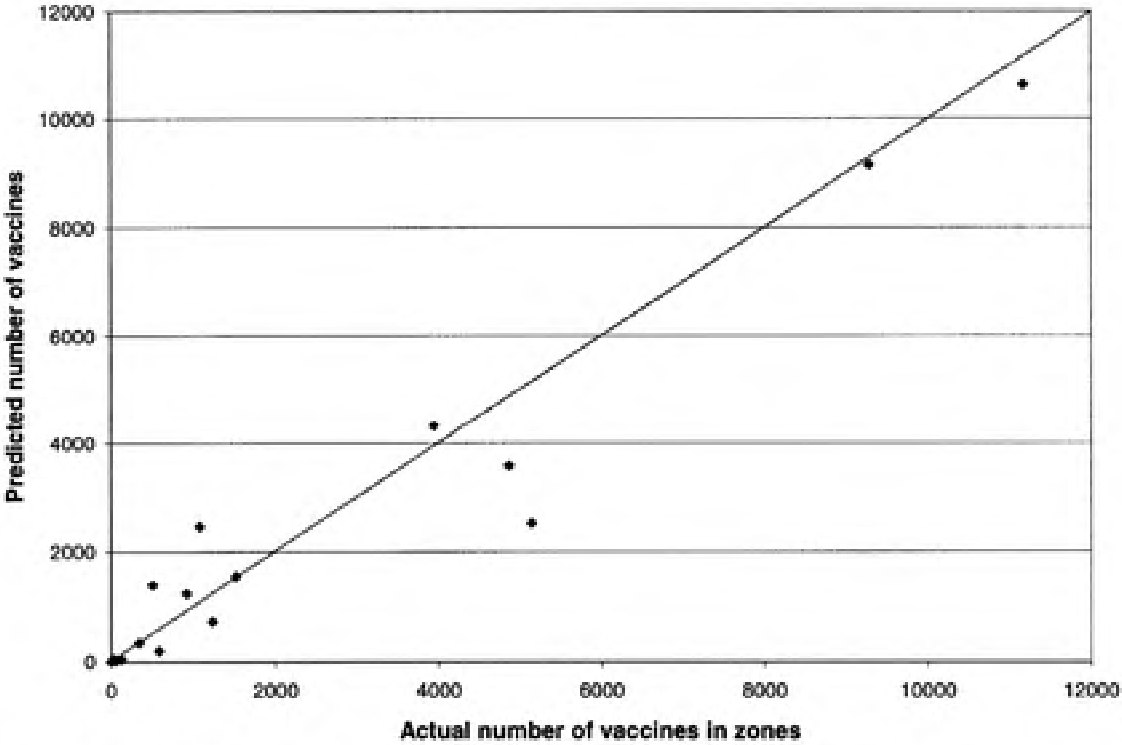

Figure A2: Actual number of vaccines obtained and number predicted from model specification B1 in Table A3.

Two alternative specifications were attempted as well. In the first (Model B2 in Table A3), the two *bairros* with vaccination outposts (Esturro and Matacuane) were excluded, because they may have been influenced by confounding factors related to information not included in the model. With this specification, we estimate slightly higher total WTP of US$1.25 per capita for two vaccine doses. The second specification (Model B3) added the median income quartile variable. Given the multicollinearity problems described above and the fact that this was a rather approximate strategy, it is not surprising that the adjusted R^2 statistic decreased and the income variable was not significant in the model. Nonetheless, the estimate for total benefits hardly changed (US$1.10 per capita).

Table A3: Regression models for percentage of population vaccinated in each *bairro* as a function of full travel cost (log-linear specification with half median wage valuation)[22]

Variable	Model B	Model B1	Model B2	Model B3	Model B4
Full travel cost *TC*[23] (thousands of Mts)	-0.42*** (0.05)	−0.31*** (0.09)	−0.30*** (0.04)	−0.39*** (0.07)	−0.27*** (0.06)
Diarrhea incidence *I* (cases/1000 people)	-0.05 (0.06)	−0.02 (0.06)	−0.03 (0.07)	−0.03 (0.07)	0.01 (0.04)
Vaccination *bairro b* (=1 if Esturro/ Matacuane)	0.61 (0.64)			0.63 (0.65)	
% Knowledge of vaccine trial *K*		0.05 (0.03)			0.06*** (0.02)
Income quartile of median household *Y*				0.23 (0.33)	0.18 (0.18)
Young child dummy *C*					−1.78*** (0.47)
School child dummy *S*					−0.89*** (0.31)
Constant	6.1*** (1.2)	1.4 (3.2)	6.6*** (1.4)	4.9** (2.1)	−0.64 (1.78)
Number of *bairros*	22	22	20	22	63
Adjusted R^2	0.885	0.894	0.864	0.881	0.887
Consumer surplus (billions of Mts)[d]	12.2	12.8	12.1	12.4	11.2
Per capita consumer surplus (US$)	1.09	1.14	1.25	1.10	0.99

***Significant at 1% level **Significant at 5% level *Significant at 10% level

Finally, in the age group zonal model (Model B4 in Table A3), which incorporates variation in the assumed value of time across age groups, estimated per capita WTP was found to be $0.99. The infant and child dummy variables are statistically significant and negative, suggesting that those age groups' visitation rates are lower than expected based on their travel costs and the other explanatory variables in the model. In addition, the coefficient of this variable for the youngest children is larger than for the school children.

[22] Standard errors in parentheses.
[23] 2005US$1= Mts. 24,500; 1 billion Mts = US$40,815.

Two explanations for these age group results seem plausible. One is that children's participation, and particularly that of the youngest children, was also dependent on adults' higher travel cost since they had to be accompanied to outposts (in all but 18 percent of surveyed households). In fact, including an interaction term of the age group dummy variables with adults' travel costs leads to the interaction term being significant, and the dummy variables lose significance. The average WTP estimate for the model with such an interaction term does not change. Another explanation is that the assumed values of time for the two groups of children were relatively too low (or alternatively that adults' value of time was relatively too high) and that the dummy variables were picking up these systematic biases. It thus becomes particularly important to investigate the implications of value of time assumptions in these models. The sensitivity of the WTP measures to this value is shown in Figure A3 below; given the value of time judged reasonable in the main part of the report, estimates around $1 seem reasonable.

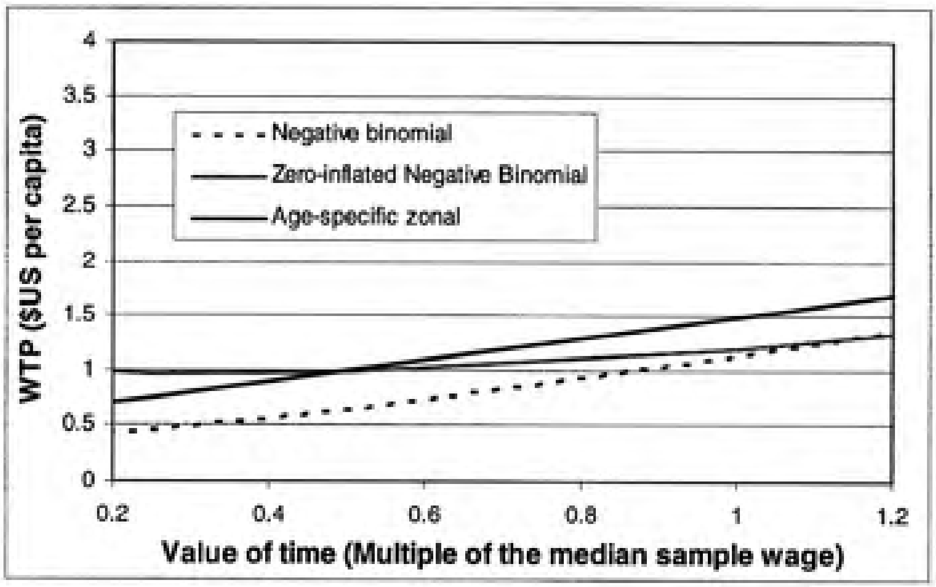

Figure A3: Sensitivity of WTP estimates from different models to the assumed value of time

Discussion: Comparing the zonal and household models

Despite the differences between the zonal and household modeling approaches, the welfare estimates obtained from them (about US$1) are remarkably similar in this application. Nonetheless, it is important to consider differences. The most obvious difference between the zonal and household models of vaccine demand is in the number and type of observations. The zonal models seek to explain the percentage of residents of a zone that were vaccinated, and so do not account for variation and heterogeneity within that zone. Nor do they incorporate many of the socio-economic factors that should impact the number of vaccinations obtained. The other factors that are included in the zonal models (information about the trial, income and diarrheal incidence) are only known imprecisely. The zonal models have few observations ($N = 22$, or N=63 when age groups are included) and are vulnerable to aggregation bias. On the other hand, the household models incorporate heterogeneity and socio-economic factors much more effectively, but are vulnerable to problems of recall bias, particularly in the subjective formulations. Both could yield biased welfare estimates if the assumed value of time is incorrect.

Another striking difference between models is that the coefficients on travel cost are much larger in the zonal models than in the household models. For example, a one unit increase in travel cost (1,000 Mts) can be interpreted to lower the percentage of people vaccinated by about 30% in the zonal models, leading to a very quick dropoff in the number of people vaccinated with small increases in travel costs. In contrast, the household models predict that a one unit increase in travel cost would lead to about a 6-9% decrease in the expected number of vaccines acquired for each household (or a 7% decrease in the probability of participation and a 2% decrease in the expected number of vaccines acquired for participating households). There are a number of possible explanations for this difference.

First, the household models allow for statistical control of education, income, and other household socioeconomic factors that tend to be correlated with travel cost in space, but cannot adequately be dealt with in the zonal models. We know that this is in fact true: income, household composition, education and knowledge of the trial are all highly correlated with travel cost in the zonal model. Because the statistical control for these variables is imperfect, the travel cost variable tends to absorb their entire effect and has a particularly large (biased) coefficient. In a few formulations (not shown), the opposite effect occurs; the travel cost coefficient becomes unstable while other independent variables predict the entire change in visitation rates. Another issue related to the lack of sufficient information and variation in the zonal models is that the edge-to-edge distance used in the models represents a lower bound travel cost for households from their particular zones. The edge models were preferred because of their improved fit, but using a centroid-to-edge model (which would tend to overestimate costs and yielded travel costs 35% higher than the edge-to-edge models) decreased the coefficient on travel cost by 10-15%, without seriously affecting the per capita WTP estimate.

A second possibility is that the coefficient on travel cost in the household analysis is sensitive to the way in which the sample is constructed in space. Only a limited number of interviews (8%) were conducted in the most heavily vaccinated *bairros* of Matacuane and Esturro. As a result, the average number of vaccines acquired by households in the survey sample may have been artificially low. Because bairros nearest the trial outposts were not included in the analysis, the effect of the travel cost variable in the household models may have been attenuated by this aspect of the sample.

A third factor that could have impacted the effect of the travel cost variable is the fact that the zonal model contained the whole population of Beira, while the household model only included households with children under the age of 19 who knew about the vaccination trial. The selection of households with children was used in order to maintain consistency with protocols for the CVM demand studies conducted in other countries, since policy-makers were particularly interested in the demand among households with children. However, it may have led to creation of a sample of people that were willing to travel longer distances than normal in order to have their children protected, which would also have diminished the effect of the travel cost variable. Plus, there was no way to know how many of the people in different zones really knew about the vaccination trial (in order to scale the eligible population fractions V_{kj} appropriately), so we tried household models that did not exclude respondents who were unaware of the trial in order to test how much bias this could introduce. In the zero-inflated model, the magnitude of the travel cost coefficients in the participation and number of vaccines acquired steps increased by 18% and 27%, respectively, which combined would lead to a 50% overestimate of the travel cost coefficient.