HEALTH ECONOMICS
Health Econ. **14**: 687–701 (2005)
Published online 24 December 2004 in Wiley InterScience (www.interscience.wiley.com). **DOI:**10.1002/hec.959

Preferences for hospital quality in Zambia: results from a discrete choice experiment

Kara Hanson[a,*], Barbara McPake[a], Pamela Nakamba[b] and Luke Archard[a]
[a] *Health Policy Unit, London School of Hygiene and Tropical Medicine, UK*
[b] *Department of Economics, University of Zambia, Zambia*

Summary

This study reports on the results of a discrete choice experiment undertaken in Zambia to assess the factors influencing the demand for hospital care in Zambia, in particular the role of (perceived) quality and trade-offs between price and quality. Valuations of quality were evaluated for the treatment of two acute medical conditions, cerebral malaria in adults and acute pneumonia in children. Marginal utilities and willingness-to-pay for attributes of quality of hospital care were estimated, together with the influence of socioeconomic characteristics on these valuations and the extent of non-linearities in valuations of time and money. We find the technical quality of care, as represented by the thoroughness of examination, to be the most important quality attribute, followed by staff attitudes and drug availability. Valuations of examination thoroughness increase with increasing socioeconomic status. The disutility of cost was found to decrease with higher socioeconomic status, as was the value of drug availability. The implications of the findings for Zambian hospital sector reforms are discussed. Copyright © 2004 John Wiley & Sons, Ltd.

JEL classification: I11 (Analysis of health care markets)

Keywords hospitals; quality; discrete choice experiments; Zambia

Introduction

Stated preference techniques are increasingly used in health economics as a way of taking patient preferences into account when making health care planning decisions, particularly where it is not feasible to use revealed preference data or where the key behaviours of interest are obscured, for example because of the influence of policy variables such as universal health insurance [1] [Lancsar E, Savage E. Deriving welfare measures from discrete choice experiments: inconsistency between current methods and random utility and welfare theory. *Health Econ*, in press]. Stated preference is a form of conjoint analysis, in which goods or services are described in terms of their characteristics, and individual valuations are assumed to depend on the levels of these characteristics [Mark TL, Swait JD, Using stated preference and revealed preference modeling to evaluate prescribing decisions. *Health Econ*, in press] [2]. Although a number of different formats exist, including graded pairs and ranking approaches [1,3], within the field of health economics the discrete choice method has been used most widely. A recent review of discrete choice experiment (DCE) literature identified 31 studies applying the method in health economics [4]. DCE has

*Correspondence to: Health Policy Unit, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK. E-mail: kara.hanson@lshtm.ac.uk

Received 28 March 2003
Accepted 11 August 2004

been used to examine a variety of issues in the health sector, including preferences for health outcomes [1,5,6] [Lancsar and Savage, in press], characteristics of service delivery [7–11] [Slothuus Skjoldborg U, Gyrd-Hansen D. Conjoint analysis. The cost variable: an Achilles' heel? *Health Econ*, in press], management of specific health conditions [2,12–19], health sector resource allocation [20,21], and provider preferences for job characteristics [8]. All of the published literature is from high-income settings, though we are aware of at least two studies currently underway in low-income countries (Sally Lake, Phare Mujinja, personal communication).

The aim of this study is to understand the factors affecting the demand for hospital services in a low-income setting, urban Zambia, in particular the role of (perceived) quality and the trade-offs between price and quality. It was not possible to use revealed preference data because the available large scale household survey data from the 1998 Zambia Living Conditions Monitoring Survey (LCMS) [22] contained too few inpatient episodes in those markets where there is effective competition among hospitals, and no information was collected about the quality of services in the facilities which form the relevant choice set. The discrete choice form of stated preference study seemed most appropriate because it is relatively simple and best mimics the type of choice that individuals make in the hospital care market.

The policy context for this study is provided by the hospital reforms currently underway in Zambia, which are part of the overall health sector reforms which have been implemented since the early 1990s [23]. At the hospital level, the reforms are designed to strengthen incentives for efficiency and quality by increasing autonomy over managerial decision making [24–26]. Hospitals are encouraged to augment their revenue base by generating revenue from paying patients. This implies that they are 'competing' for patients, which should make them more responsive to patient preferences. The reforms have been accompanied by significant institutional changes, the most salient of which for the purposes of this analysis is the introduction of contracting arrangements between the Central Board of Health and autonomous hospital boards for secondary and tertiary level facilities [27].

The incentive regime faced by hospitals is driven by the ways they earn their revenue. By means of their contract with the Central Board of Health, hospitals receive a grant from central government that is calculated on the basis of the number of 'approved' beds (although there is tremendous variation between budgeted grant revenue and the amounts actually received). The second largest revenue component is from patient fees, where a distinction is made between 'low cost' patients, who pay a relatively small fee, and 'high cost' patients who use separate facilities within the hospital. The demand response to these price-quality differences has implications for the allocation of resources within the hospital, and the equity implications of two-tier services [McPake B, Hanson K, Adam C. '*Two-tier*' *charging strategies in public hospitals*: *implications for intra-hospital resource allocation and equity of access to hospital services*. Unpublished Manuscript, 2003].

Most studies using revealed preference data on the role of quality in the demand for care have either assessed quality from the provider's perspective, or have failed to distinguish among different dimensions of perceived quality [Hanson K, Yip WC, Hsiao W. *The impact of quality on the demand for outpatient services in Cyprus*. *Health Econ*. in press]. Two papers which have looked at multiple dimensions of perceived quality have found patients to be more responsive to non-technical quality attributes such as interpersonal quality, amenities and opening hours [Hanson *et al*., 2003] [Yip W. *Choice of outpatient provider in Egypt*: *Does patient satisfaction matter for low income countries residents*? Unpublished Manuscript, 2001]. However, both of these studies focused on outpatient care, and little is known about preferences for hospital care.

Theoretical framework and estimation strategy

The theoretical underpinnings of discrete choice experiments are rooted in Random Utility theory [1,28]. In choosing who to consult in the case of illness, individuals face a number of options, each of which yields indirect utility, Y^*. Y^* is a latent variable which is not directly observed: all we observe is whether an option is chosen or not. Individuals are assumed to choose the option that yields the highest indirect utility, i.e. $Y_i = 1$ if $Y_i^* = \text{Max}(Y_1^*, Y_2^*, Y_3^*, \ldots, Y_m^*)$.

The indirect utility yielded by an option is assumed to be a function of choice-specific attributes. The residual, ε, captures unobserved variation in the characteristics of the different options, and errors in measurement and optimisation by the consumer [29]. In other words,

$$Y_{iq} = X_i\beta_i + \varepsilon_{iq}$$

where Y_{iq} is the indirect utility of individual q for option i and X_i is a vector of attributes of the ith choice.

Making specific assumptions about the distribution of the error term, the choice can be modelled using a logit or probit specification. Because each individual is asked to make multiple choices the error terms cannot be assumed to be independent, and panel data estimation techniques are required.

Typically, the regression model is specified in terms of differences in attributes between the two choices being analysed:

$$\Delta Y = \beta_0 + \beta_1(X_{1i} - X_{1j}) + \beta_2(X_{2i} - X_{2j}) + \cdots(\varepsilon_i - \varepsilon_j)$$

The estimated parameters can be interpreted as the marginal utility from a change in the level of the attribute as one moves from Option 1 to Option 2. The ratio of any two parameters is the marginal rate of substitution between them. If cost is included as an attribute, then the ratio of an estimated quality parameter to the cost parameter can be interpreted as the marginal willingness-to-pay for that attribute.

Study setting

Fieldwork was undertaken in the city of Kitwe, located in Copperbelt Province which is the commercial centre of the country and the site of the majority of the copper mines, an important source of export earnings for Zambia. There are a variety of inpatient facilities in Kitwe town, including Kitwe Central Hospital which has 'low cost' and 'high cost' wards. There are also hospitals operated by mining companies within a short distance.

Our analysis of the data from the 1998 Living Conditions Monitoring Survey [22] indicates that Copperbelt province is the most urbanised province after Lusaka, with 77% of the population being urban. Sixty-five percent of households fell below the poverty line in 1998. Just over 9% of the population surveyed had reported illness in the previous 2 weeks. Of these, 41.4% consulted a health care provider, 36.5% self-treated, and 22.0% did nothing. The rate of consultation increases with socioeconomic status (measured as expenditure per adult equivalent), while lower socioeconomic groups are more likely to have reported doing nothing in face of illness. For individuals who reported spending one or more nights as an inpatient, the mean expenditure for hospitalisation was K27 961. This is 66% of mean monthly per capita expenditure (equivalence-adjusted).

Methods

We chose to study health problems which would be perceived as sufficiently serious to warrant immediate hospitalisation. Different questionnaires were developed to elicit patient preferences for two different types of acute medical problem, with each respondent receiving only one of the two questionnaires (choice of hospital for an adult relative with symptoms of complicated malaria, and for a child with signs of severe pneumonia).

Establishing attributes and levels

Focus group discussions were used to obtain information about the dimensions of the quality of hospital services that are important to individuals when choosing a hospital [30]. Data were collected in Lusaka and Kitwe districts, the two areas of focus for the broader study of autonomous hospitals policy. In each district, eight focus groups were held, with a minimum of six participants per group. Separate discussions were held with males and females for each of the following groups: market sellers, church members from a high density residential neighbourhood, church members from a low density residential neighbourhood, and university students. Discussions were conducted by a sociologist from the Institute of Economic and Social Research and were tape recorded, with transcripts supplemented by detailed handwritten notes. Manual content analysis was used to group the data according to major themes and to identify the quality attributes most frequently mentioned by participants.

Following the analysis of the focus group discussions, a total of five quality attributes were chosen for inclusion in the questionnaire (Table 1). All five (or minor modifications thereof) were among the top ten positive attributes identified in the focus groups. Other issues which were frequently mentioned by discussants but not included in the questionnaire because of conceptual overlap with included dimensions were: availability of diagnostic facilities, quality of nursing care, staff dedication, and availability of staff.

The levels of attributes were chosen to reflect the range of situations that respondents might expect to experience. In choosing the price levels our aim was to reflect the different hospital options available. These are 'low cost' (public) wards at public hospitals, 'high cost' (private) wards at public hospitals, and private hospitals. A standard package of care for the two types of illness was determined and priced using price lists for a public high-cost ward and a private hospital. Standard charges apply for public wards in public hospitals. From this information we arrived at the range of prices of K20 000 (£3.70), K100 000 (£18.50) and K250 000 (£46.30).

Generating the questionnaire

The five attributes and levels indicated in Table 1 give rise to a total of 216 possible scenarios ($3^3 \times 2^3$) for each of the two illness conditions. A fractional factorial design was used to reduce this to a feasible number. Speed v. 2.1 software was used to generate an orthogonal main effects design, which produced a total of 16 scenarios. One of these was randomly chosen as a constant comparator, which gave a total of 15 choice pairs for each questionnaire, which is at the top of the range (9–16) which can be managed by respondents before boredom sets in [4].

The resulting experimental design appears to satisfy three of the standard criteria for efficient designs: orthogonality, minimal overlap, and level balance [Zwerina K, Huber J, Kuhfeld WF. *A general method for constructing efficient choice designs*. Unpublished Manuscript, Fuqua School of Business, Duke University, 1996]. Although the use of design software to generate the 16 scenarios was aimed at producing an orthogonal fractional factorial design, orthogonality is no longer guaranteed once scenarios are paired [7]. Orthogonality in attribute differences was therefore verified using χ^2 tests of association. All

Table 1. Attributes and levels included in the stated preference study

Attribute	Variable name	Levels, definition and regression coding
Likelihood that the child will receive all the drugs s/he needs at the hospital	DRUG	Drugs available[a] = 1 Likely to have to seek some drugs elsewhere = 0
Likelihood that the hospital staff will examine the child properly	EXAM	Child examined thoroughly[a] = 1 Child given a superficial examination = 0
Attitudes of staff[b]	STAFF_FRIEND STAFF_RUDE	Staff are indifferent[a] = 1 Staff are friendly = 2 Staff are rude = 0
Cleanliness of wards and toilets	CLEAN	Cleaned rarely[a] = 0 Cleaned often = 1
Waiting time between arrival at OPD and admission to the ward	WAIT	5 hours[a] = 5 2 hours = 2 30 min = 0.5
Cost to you of the child's examination and treatment	COST	K100 000[a] = 100 K20 000 = 20 K250 000 = 250

[a] Value taken in constant comparator, Choice 1.
[b] Because it takes three values, the effect of staff attitudes has been modelled using 2 dummy variables. STAFF_FRIEND = 1 if the difference is equal to –1 (option 1 = indifferent, option 2 = friendly). STAFF_RUDE = 1 if the difference is equal to 1 (option 1 = indifferent, option 2 = rude).

correlations were found to be statistically insignificant. The average degree of overlap between attributes within scenarios was 40%, with overlap ranging from 1 out of 6 to 5 out of 6 attributes. Exact level balance is achieved for the three attributes with 2 levels (with each occurring in eight scenarios); for the 3-level attributes, the distribution is unbalanced (4,4,8). Because the focus groups did not directly investigate the relative importance of different attributes (e.g. using ranking methods), it is not possible to comment on whether the design meets the criterion of utility balance.

Measuring socioeconomic status

The questionnaire included a set of socioeconomic indicators including education and ownership of assets. The choice of assets was informed by analysis of the 1998 Living Conditions Monitoring Survey to determine which assets were significantly correlated with monthly household expenditure in Copperbelt Province (household ownership of: motor vehicle, TV, video, radio, electric iron, fridge, phone, sewing/knitting machine, stove, building). Principal components analysis was used to establish the weights for an index of assets owned by the respondent's household [31]. This produces a continuous variable that was then divided into three equal-sized groups (low, medium, and high socioeconomic status). Three rather than the conventional five groups were used because of 'bunching' of the scores at particular values which made it impossible to divide the group into five equal-sized groups. Using three groups also avoided problems of small cell-sizes for sub-group analyses, and simplified the interpretation of interaction terms in the regression model (see below).

Pre-testing, sampling procedures and administering of survey

The questionnaire was pre-tested on 30 respondents. Following pre-testing minor modifications were made, which included small changes to the scenario, and revisions to the price vector to make it more consistent with the prices people would expect to pay in Kitwe.

To establish the required sample size, a literature review was undertaken which found substantial variation in the number of individuals interviewed and little technical justification for the chosen sample sizes, which ranged from 57 [13] to 3893 [7]. We settled on a sample size of 300 questionnaires per condition (a total of 600), which was in the range of other studies in the literature, and was influenced by constraints on feasibility and the availability of resources, given the need for interviews to be conducted face-to-face in this setting. Resources were not available for a fully random sampling process, so our sampling process was designed to balance the needs of representativeness and convenience. The sample was stratified by neighbourhood, with areas chosen to reflect a range of housing types, infrastructure and facilities. One-third of the sample was chosen from two each of 'high density', 'medium density', and 'low density' neighbourhoods. Upon arriving at the sampling site, researchers randomly selected households. The refusal rate was not logged, but the number of refusals was negligible. Households that refused were replaced with the next nearest household.

Three university undergraduate students conducted the interviews. Training involved a general review of the method, a detailed review of the questionnaire, role-playing and administration of the questionnaire during the pilot study. During training standard translations of the questions into the local language (Bemba) were agreed, and interviews were administered in either English or Bemba. To aid in comprehension, a picture board with symbolic representations of the attributes and their levels was used, while the interviewer read out the different scenarios. Figure 1 shows an example of a choice presented to respondents. In general, respondents were very enthusiastic about answering the questions.

Data analysis

The survey results were analysed using the random-effects probit estimator in STATA v. 7.0. The coding of variables for analysis is shown in Table 1. The baseline empirical model (Models 1 and 3 in Table 3) was:

Choice 1

	Hospital 1	**Hospital 2**
The likelihood your relative will receive all the drugs s/he needs at the hospital.	It is likely that **all** the drugs will be **available**.	It is likely that you will have to look for **some** of the drugs **elsewhere**.
The likelihood that the hospital staff will examine your relative properly.	It is likely the examination will be **thorough.**	It is likely the examination will be **superficial.**
The attitudes of staff:	The staff treat you with **indifference**: they are neither friendly nor rude.	The staff treat you with **indifference**: they are neither friendly nor rude.
The cleanliness of the wards and toilets.	The wards and toilets are cleaned **rarely**.	The wards and toilets are cleaned **rarely**.
Waiting time between arrival at the out-patients department and admission to the ward	You have to wait for about **5 hours** before admission to the ward.	You have to wait for about **5 hours** before admission to the ward.
The cost to you of your relative's examination and treatment.	You have to pay K100, 000 for examination and treatment.	You have to pay K100, 000 for examination and treatment.

Figure 1. Example of choice pairs from questionnaire

$$\begin{aligned}\Delta Y = {} & (\alpha_{02} - \alpha_{01}) + \alpha_1(\text{d\_EXAM}) + \alpha_2(\text{d\_DRUG}) \\ & + \alpha_3(\text{d\_WAIT}) + \alpha_4(\text{d\_STAFF\_FRIEND}) \\ & + \alpha_5(\text{d\_STAFF\_RUDE}) + \alpha_6(\text{d\_CLEAN}) \\ & + \alpha_7(\text{d\_COST}) + \varepsilon + \mu\end{aligned}$$

where $\text{corr}(\varepsilon, \mu) = \rho$, which takes account of the correlation among an individual's choices, and

d_EXAM = difference in examination between option 1 and 2

d_DRUG = difference in drug availability between option 1 and 2

d_WAIT = difference in waiting time between option 1 and 2

d_STAFF_FRIENDLY and d_STAFF_RUDE = dummy variables for difference in staff attitudes between option 1 and 2

d_CLEAN = difference in cleanliness between option 1 and 2

d_COST = difference in cost between option 1 and 2

Models 2 and 4 include quadratic terms in (differences in) cost and waiting time in order to test empirically for non-linear effects of these variables [3].

In Models 5 and 6 the socioeconomic status variable was interacted with the quality attributes to determine the degree to which socioeconomic differences influence the valuations of quality and money [1,32]. A likelihood ratio test was used to exclude insignificant interaction terms from the reduced model, and to test the final reduced model against the restricted (no socioeconomic status (SES) interactions) model.

Theoretical validity of the valuations was assessed by determining whether the estimated quality parameters were of the anticipated sign, and whether the interaction terms with socioeconomic status met theoretical expectations. The sign of the difference variables depended on the value taken by the constant comparator, thereby influencing the expected sign of the estimated coefficients. Better drug availability, more thorough examination, indifferent (vs rude) staff, were all expected to increase utility (positive coefficient); indifferent (vs friendly) staff, rarely cleaned toilets (vs often), longer waiting time and higher cost were all expected to decrease utility (negative coefficient). Quality was expected to be a normal good, with higher utility for higher income groups. Cost was expected to have a more negative effect

for lower socioeconomic groups, implying that the interaction terms with group 2 (middle) and group 3 (high) should be positive. Drug availability was expected to give greater utility to lower socioeconomic groups, because of the added costs incurred when drugs have to be purchased elsewhere, implying negative coefficients on the interaction terms with medium and high socioeconomic groups. Higher income groups should have a higher opportunity cost of time, and therefore a greater disutility from waiting time, implying negative coefficients on interaction terms with medium and high socioeconomic groups.

The internal consistency of responses was investigated by including two choice pairs in which one option was superior to the other on all characteristics, assuming that people prefer a lower cost, more thorough examination, a shorter waiting time, and better drug availability, cleanliness and staff attitudes. Individuals who failed to choose the superior options may have misunderstood the questionnaire, or have been unable to give consistent answers because of problems of communication or translation by the interviewers.

Given the low-income setting for this study, we were particularly concerned about the possibility of dominant preferences over the price attribute, which would lead to individuals failing to trade among attributes. A limited test for dominance in preferences over cost was undertaken by identifying those individuals who always chose the lowest priced option. The socio-demographic characteristics of individuals demonstrating these preferences were evaluated, and econometric models were estimated with and without these observations in order to observe the effect on estimated coefficients.

Results

The target number of 300 questionnaires for each of the illness conditions was achieved. A substantial number of respondents stated that they could not afford the price of any of the options (49 for the child questionnaire and 31 for the adult questionnaire). This response was recorded on the questionnaire, and these observations were excluded from the analysis. No follow-up questions were asked of these individuals about why they could not pay, as this response had not been anticipated for hypothetical questions and had not arisen in the pre-testing. For both questionnaires, this group had significantly lower ownership of all assets (with the exception of a radio in the child questionnaire group) (see Appendix). It is likely that even the lowest payment amount was simply unaffordable to them, and they were not able to answer the questions in a hypothetical context.

A total of 29 respondents for the child questionnaire and 33 for the adult questionnaire failed to choose the three superior options (approximately 10% in each group), and these observations were also excluded from the econometric estimation. A comparison of those who did not choose the superior option with the rest of the sample revealed a mixed picture, with lower ownership of some but not all assets (see Appendix).

Characteristics of respondents

Table 2 shows the summary statistics for those who responded to each of the questionnaires. The group responding to the adult illness questionnaire had slightly higher levels of asset ownership than the child illness group, but in other respects the groups were very similar. The level of ownership of assets is higher than in the general urban population of Copperbelt province. Although high density areas, where poorer households generally live, were included in the sample, the equal weighting of high, medium, and low density areas appears to have resulted in a sample that over-represents higher income areas.

Test for dominant preferences

In the test for dominant preferences over price, we identified a small number of individuals who always chose the lowest price option ($n = 9$ for child; $n = 13$ for adult). These individuals exhibited lower ownership of most assets (see Appendix). We estimated a model including interaction terms for whether the individual demonstrated dominant preferences, but the model failed to converge for both adult and child datasets. The baseline model was therefore estimated with and without data from these individuals (results not shown). Visual comparison of the coefficients and of their 95% confidence intervals suggests that the results do not differ materially when these individuals are excluded. However, because of concerns about the

Table 2. Characteristics of sample and comparison with general urban population of Copperbelt province

	Respondents to child illness questionnaire	Respondents to adult illness questionnaire	General population (urban residents of Copperbelt province (LCMS)
Percent of households owning			
Motor vehicle	28	27.3	7.5
TV	68	75.3	57.8
Video	47	56	15.7
Radio	88.7	90.7	71.8
Iron	64.7	72	41.9
Fridge	56	61.7	26.9
Telephone	43.3	45	7.4
Sewing/knitting machine	29.3	29.3	19.2
Stove	66.7	70.7	50.2
Building	10.3	15	4.2
Education			
None	6.7	2.3	
Primary	3.7	5.3	
Some secondary	15	20.3	
Finished secondary	54	52.3	
College/university	20.7	19.3	
Age: mean (median)	27 (25)	28.1 (25)	
Household size: mean (median)	6.7 (7)	6.8 (7)	
Sex: percent male	31	46.3	
N	100	300	2308

source of the discontinuity (strong preferences vs complexity of the decision task, see [9]), the remaining discussion focuses on models which were estimated excluding these observations.

Magnitude and statistical significance of attributes

Table 3 presents the results from the baseline random effects models (Models 1 and 3) for the adult illness and child illness, respectively. The models appear to fit the data well, with more than 85% of observations correctly predicted. The statistical significance of rho confirms the appropriateness of using the panel data estimator.

The estimated coefficients are all of the anticipated sign, and are all statistically significant with the exception of one of the staff attitude dummies in the adult model, and cleanliness in both models. The effects are similar across the two illness conditions. The coefficients can be interpreted as the effect of the difference between option 1 and option 2 on the likelihood of choosing option 1 over option 2, with the sign reflecting whether the level of the attribute was higher or lower in option 1. Compared with a superficial examination, a thorough examination increases the likelihood of choosing an option, as does better drug availability. The coefficients of the two staff attitude variables are of opposite sign as anticipated (with indifferent vs friendly staff associated with a negative coefficient, reflecting a lower probability of choosing option 1 (indifferent) over 2 (friendly), and indifferent vs rude with a positive coefficient, reflecting a higher probability of choosing option 1 (indifferent) over 2 (rude)). However, the effect of the difference between indifferent and rude staff is significantly larger than the difference between indifferent and friendly staff, as indicated by both the magnitude of the coefficients, and the statistical insignificance of the latter in the adult model. This implies greater disutility associated with rude staff than utility associated with friendly staff, when compared with indifferent staff. Longer waiting time is associated with higher disutility, as is higher cost. The estimated coefficients represent the utility associated with a one hour difference in waiting time, and a K1000 difference in costs. The coefficients on cleanliness were not statistically significant in either model.

Table 3. Random effects probit, baseline models

	Adult		Child	
	Model 1: baseline, excluding dominant preferences	Model 2: including quadratic cost and time terms	Model 3: baseline, excluding dominant preferences	Model 4: including quadratic cost and time terms
EXAM	2.78*** (2.53, 3.03)	2.76*** (2.51, 3.01)	2.37*** (2.17, 2.57	2.39*** (2.17, 2.62)
DRUG	0.91*** (0.68, 1.14)	0.89*** (0.67, 1.12)	0.87*** (0.70, 1.05)	0.90*** (0.71, 1.08)
WAIT	−0.20*** (−0.24, −0.16)	−0.57*** (−0.80, −0.34)	−0.20*** (−0.24, −0.17)	−0.92*** (−1.18, −0.66)
WAIT2		0.09*** (0.04, 0.14)		0.15*** (0.10, 0.21)
STAFF_FRIEND	−0.13 (−0.30, 0.05)	−0.08 (−0.28, 0.12)	−0.19** (−0.36, −0.03)	−0.44*** (−0.68, −0.20)
STAFF_RUDE	1.51*** (1.30, 1.72)	1.82*** (1.52, 2.10)	1.17*** (0.98, 1.35)	1.31*** (−0.68, −0.20)
CLEAN	−0.02 (−0.25, 0.21)	−0.09 (−0.32, 0.13)	−0.05 (−0.23, 0.13)	−0.29*** (−0.48, −0.09)
COST	−0.009*** (−0.01, −0.008)	−0.007*** (−0.01, −0.005)	−0.008*** (−0.009, −0.007)	−0.01*** (−0.01, −0.009)
COST2		0.000008 (−0.00001, 0.00002)		−0.00003*** (−0.00005, −0.00008)
CONSTANT	−0.94*** (−1.19, −0.69)	−1.10*** (−1.47, −0.72)	−0.85*** (−1.09, −0.61)	−0.69*** (−0.99, −0.38)
N	3344	3344	3193	3328
Log likelihood	−851.54	−845.34	−956.55	−1003.08
Prob(Chi^2)	<0.001	<0.001	<0.001	<0.001
Rho	0.30 (0.22, 0.38)	0.30 (0.23, 0.39)	0.34 (0.28, 0.42)	0.34 (0.27, 0.41)
Proportion 1s correctly predicted	88.6	88.6	86.0	86.0
Proportion 0s correctly predicted	87.7	87.7	87.4	87.4

$p<0.05$,*$p<0.01$.

Testing for non-linear effects of money and quality attributes

Models 2 and 4 in Table 3 show the results of including squared terms for cost and waiting time in the estimated utility function. This allows a test of whether the assumption of a constant marginal willingness-to-pay is valid. For the adult illness condition the cost-squared term is not statistically significant, while for the child condition it is negative and significant, indicating that the disutility of cost depends on the level of cost, increasing (i.e. becoming more negative) as the size of the cost difference between the two options increases. For both conditions, the disutility of waiting time becomes less negative with the level of difference. The coefficient on cleanliness becomes negative and significant in the child illness model when the quadratic terms are included.

Comparing the marginal utilities across socioeconomic groups

The effects of interacting all quality and cost coefficients with dummy variables for socioeconomic status are shown in Table 4. The excluded group is the lowest socioeconomic group. For both the adult and the child data, interactions with socioeconomic status significantly improved model fit for examination, drug availability and staff attitudes and cost (all $p<0.01$). The interactions were statistically insignificant for waiting time and cleanliness. Higher socioeconomic groups have a significantly higher marginal utility from a thor-

Table 4. Random effects probit, including interactions with SES, reduced models

	Adult	Child
	Model 5	Model 6
SES_MED	0.15 (−0.41, 0.72)	0.29 (−0.20, 0.78)
SES_HIGH	−0.08 (−0.69, 0.53)	0.09 (−0.46, 0.64)
EXAM	2.08*** (1.72, 2.44)	1.88*** (1.57, 2.19)
EXAM*SES_MED	0.96*** (0.42, 1.51)	1.20*** (0.73, 1.67)
EXAM*SES_HIGH	1.99*** (1.32, 2.66)	0.67*** (0.20, 1.14)
DRUG	1.43*** (1.05, 1.81)	1.52*** (1.21, 1.82)
DRUG*SES_MED	−0.35 (−0.87, 0.16)	−0.85*** (−1.27, −0.44)
DRUG*SES_HIGH	−0.99*** (−1.54, −0.43)	−1.00*** (−1.44, −0.56)
WAIT	−0.24*** (−0.28, −0.20)	−0.22*** (−0.26, −0.18)
WAIT*SES_MED		
WAIT*SES_HIGH		
STAFF_FRIEND	0.54*** (0.18, 0.90)	0.13 (−0.19, 0.45)
STAFF_FRIEND*SES_MED	−0.76*** (−1.23, −0.28)	−0.54** (−0.98, −0.10)
STAFF_FRIEND*SES_HIGH	−1.12*** (−1.65, −0.59)	−0.39 (0.86, 0.08)
STAFF_RUDE	1.65*** (1.26, 2.04)	1.07*** (0.76, 1.38)
STAFF_RUDE*SES_MED	−0.09 (−0.60, 0.43)	0.09 (−0.33, 0.51)
STAFF_RUDE*SES_HIGH	−0.003 (−0.53, 0.52)	0.27 (−0.18, 0.74)
CLEAN	−0.07 (−0.33, 0.18)	−0.06 (−0.25, 0.13)
CLEAN*SES_MED		
CLEAN*SES_HIGH		
COST	−0.01*** (−0.01, −0.009)	−0.009*** (−0.01, −0.008)
COST*SES_MED	0.003*** (0.0005, 0.004)	0.001 (−0.0004, 0.003)
COST*SES_HIGH	0.0007 (−0.002, 0.003)	0.003*** (0,0009, 0.005)
CONSTANT	−1.01*** (−1.44, −0.58)	−0.99*** (−1.37, −0.61)
N	3344	3193
Log likelihood	−764.47	−889.29
Prob(Chi^2)	<0.001	<0.001
Rho	0.36 (0.28, 0.45)	0.38 (0.31, 0.46)
Proportion 1s correctly predicted	89.7	86.8
Proportion 0s correctly predicted	84.3	85.1

*$p<0.10$, **$p<0.05$,***$p<0.01$.

Table 5. Willingness to pay estimates

	Adult				Child			
	All	Low SES	Med SES	High SES	All	Low SES	Med SES	High SES
EXAM	306 667	208 000	434 286	437 634	296 250	208 889	385 000	425 000
DRUG	101 111	143 000	*154 286*	47 312	108 750	168 889	83 750	86 667
WAIT	22 222	24 000	34 286	*25 806*	25 000	24 444	27 500	36 667
STAFF_FRIEND	14 444	54 000	31 429	62 366	23 750	*14 444*	51 250	*43 333*
STAFF_RUDE	167 778	165 000	*222 857*	*177 387*	146 250	118 889	*145 000*	*223 333*
CLEAN	2222	*7000*	*10 000*	*7527*	6250	*6667*	*7500*	*10 000*

Note: WTP estimates are calculated using results of Models 1 and 3 (all) and Models 5 and 6 (differences by SES). Figures in italics derived from attribute coefficients that were not statistically significant. In the case where the reduced model did not include interaction terms with a specific attribute, the variation is provided by the cost-interactions.

ough examination. The valuation of drug availability is higher for lower socioeconomic groups. While increased cost is always associated with negative marginal utility, this effect is less negative for higher socioeconomic groups.

The patterns of utility associated with staff attitudes are complex: In the case of the adult condition, the marginal utility associated with friendly over indifferent staff is greater for higher income groups; but the disutility associated with rude staff over indifferent staff is lower. For the child's condition, the same pattern holds for friendly staff, but when the difference is between indifferent and rude staff, the disutility is greater for higher income groups.

Willingness-to-pay

Table 5 shows the estimated willingness-to-pay for the different attributes for the whole sample (estimated using Models 1 and 3 for adult and child's condition, respectively), and by socioeconomic group (Models 5 and 6), estimated by taking the ratio of the estimated coefficient for each quality attribute to that of the cost coefficient.[a] These results help to assess the relative valuation of different attributes. The highest willingness-to-pay is for a thorough examination, followed by (avoiding) rude staff. The third most important attribute is drug availability. The levels of WTP for friendly (compared with indifferent) staff, and for cleanliness, are considerably smaller in magnitude. The waiting time results can be interpreted as willingness-to-pay for a one hour reduction in waiting time to admission.

For both illness conditions, willingness-to-pay for a thorough examination in the lowest socioeconomic group is only about half that in the highest group; and willingness-to-pay for better drug availability is considerably higher in the lowest than the highest group. Willingness-to-pay to avoid rude staff increases dramatically with SES in the case of children's illness, but this relationship is non-monotonic in the case of adults.

Discussion

The results presented above provide new information about how consumers in a developing country setting value the quality of hospital care. We find the technical quality of care, as represented by the thoroughness of examination, to be the most important quality attribute (indicated by the highest willingness-to-pay). The next two most important attributes are avoiding rude staff and drug availability. A study of consumer preferences for hospital care in Australia found the complication rate, together with waiting times, to be the most highly valued attributes [11].

Differences in the valuation of the quality and cost attributes by socioeconomic status provide insights into the equity implications of two-tier hospital policies. The Zambian health financing policy allows for a 'two-tier' pricing system which encourages hospitals to create private wards with higher levels of hotel services, provided that costs can be fully covered by the fees. The idea is that any profits that can be generated from these patients be used to cross-subsidise the services offered in the public wards. Our results reveal the dilemma facing policymakers: The most pronounced and consistently significant socioeco-

nomic gradient in willingness-to-pay was for thoroughness of clinical examination, which is an attribute clearly associated with clinical quality rather than hotel services. At the same time, it is precisely this characteristic that attracts the highest overall willingness-to-pay, and therefore offers the greatest opportunity to differentiate services into two tiers. Willingness-to-pay for hotel services, such as cleanliness, was both low and undifferentiated by socioeconomic group. This suggests to us that it is not possible to achieve an equitable two-tier charging policy which takes account of the preferences of different socioeconomic groups over quality.

A second equity implication is revealed by the differences in willingness-to-pay for drug availability. The finding that the lowest socioeconomic group has the highest willingness-to-pay for reliable drug supply is probably explained by the fact that when drugs are not available they must be purchased outside the hospital, which has more negative consequences for lower socioeconomic groups. This 'quality' attribute may therefore actually be more akin to a cost attribute. This result suggests that there is a highly regressive impact of a drug supply system which allows drugs to be frequently out-of-stock, and that ensuring a reliable drug supply in the low-cost areas of the hospital is an important pro-poor policy measure.

In the case of waiting time, it is interesting that there was no significant difference between socioeconomic groups in the marginal valuation (Table 4). While theoretically, people with higher income would be expected to have a higher opportunity cost of time, empirically it is likely that there is significant within-group variation as to whether time taken off from work to seek health care would result in a loss of income or not (e.g. this might depend on whether the individual is self-employed, is paid an hourly wage or is salaried). This heterogeneity could be responsible for the failure to find a significant time effect. The non-monotonicity in willingness-to-pay, particularly for staff attitudes, may be more related to sociological than economic phenomena – both the asymmetry between friendly and rude, and the fact that people in the middle SES group appear to experience more disutility from rude staff than those in either the low or high groups.

Alternatively, the non-monotonicity could be due to errors in measurement of SES. However, given the procedure for choosing the assets to include (based on LCMS analysis) and their demonstrated correlation with consumption expenditure used as a 'gold standard' for SES measurement, we are confident that our SES measure is a valid one.

There are a number of methodological weaknesses that we would try to remedy in a future study. First, it appears *ex post* that our sampling strategy was biased towards higher socioeconomic groups. This can be seen from the higher levels of asset ownership in our sample than in the general population and in the greater propensity of those with lower asset ownership to respond that they were unable to pay any amount for hospital care. It is also possible that within households, the nature of the questionnaire attracted the most literate respondent which could be a further source of bias. This may limit the generalisability of estimates of willingness-to-pay for quality of hospital care. Given the exploratory nature of the study and the limited resources available, it did not seem prudent at the outset to invest in a fully representative sampling procedure, however, future studies in these contexts should seek to use a more appropriately random sampling strategy. It was not possible to check the results for interviewer bias, although the relatively high proportion of respondents who did not choose the superior option may be an indication of problems with the explanation of the task and the questionnaire.

Second, we did not include a 'choose neither' option, and consequently have effectively estimated an unconditional demand curve [4]. While those who stated that they could afford neither option were excluded from the analysis, it is not clear that the implied willingness-to-pay for quality attributes would therefore be translated into actual willingness-to-pay. Participants in the focus group discussions reported that in many cases even those who are referred to the hospital choose to self-treat because of prior unsatisfactory experiences with the quality of hospital care, in particular problems with drug availability [30]. The choice of life-threatening illness conditions which were sufficiently severe that they could not be managed at home was an effort to mitigate this problem. However, in future studies it would be important to include the status quo option.

Third, we had limited opportunity for follow-up debriefing with respondents about their experience of using the questionnaire. Issues that could be addressed in such a debriefing would include the validity of the choice sets (whether the combina-

tions of attributes were credible); whether particular attributes were assumed to reflect a very specific or wider range of quality characteristics (for example whether 'thorough examination' was assumed to imply the availability and use of laboratory testing facilities); and some reflection on the complexity of the choice task.

Fourth, we used a very limited form of test for dominant preferences, which may not be sufficiently powerful to detect this violation of the standard axioms. In addition to using the criterion of whether an individual always chose the scenarios with the lowest cost, it would have been useful to incorporate additional information about the relative importance of the different attributes [9]. The association between dominant preferences and socioeconomic status suggests that dominance may have been generated by strong preferences (working through, for example, real life budget constraints) rather than features of the choice process such as task complexity and decision making heuristics [9]. A more comprehensive criterion, together with information about respondents' perceptions of task complexity (see above) might have helped to confirm this. The actual range of prices chosen may also explain non-trading of this characteristic. Our choice of price vector was guided by a desire to have prices reflect the choices actually available in the market, but the resulting price range may have been too large. These limitations notwithstanding, only a very small proportion of respondents exhibited this behaviour ($<5\%$). There is also a methodological debate about whether the data from non-traders should be included in analysis [10]. In the absence of better information about the cause of non-trading behaviour, and armed with the evidence that the econometric results are not affected materially either way, we have chosen to base our interpretation on models which excluded data from these individuals.

The implied willingness-to-pay for quality is very high, even in the lowest socioeconomic group, and considerably exceeds the reported expenditure on inpatient episodes from the LCMS. This may arise from the abovementioned bias in the sampling procedure. While this may limit the generalisability of the results, there are still indications that the estimates are valid. The price levels in the questionnaire were based on the actual price lists from these facilities, implying that these levels of prices are realistic for these facilities (see [33] for more detail). Furthermore, the conditions for which we estimated willingness-to-pay were life threatening. It is to be expected that willingness-to-pay exceeds normal expenditure levels in these circumstances.

There is, however, a certain amount of divergence between the utilities estimated for the quality attributes and the levels of those attributes that would be available in the highest-priced option. For instance, there was a view expressed among focus group participants that private hospitals did not have the full range of specialists and diagnostic facilities required for treatment of serious illness. In other words, the highest priced option would not necessarily provide the quality attribute (thorough examination) that people are most willing to pay for. This may also explain why in practice, utilisation of private inpatient services is much lower than would be predicted from the implied willingness-to-pay.

The choice of the quality attributes to include in the questionnaire, and their interpretation by respondents, also raises important issues for study design and interpretation of results. We chose attributes that encompassed a broad range of the quality dimensions that were identified in the focus group discussions, thus minimising the unmeasured quality variation between the options. However, this may complicate the interpretation of the coefficients, which may reflect a broader construct in the mind of the respondent. The large coefficient on the thoroughness of examination, for example, may capture the effect of other attributes of importance to respondents such as the availability of diagnostic facilities and the experience and specialised qualifications of staff. Both of these were among the ten most important attributes that emerged from the qualitative work.

The design of our experiment follows the empirical literature in assuming independence among quality attributes. It has been suggested that the main effects typically explains over 80% of the variation in preferences (Pearmain *et al.*, 1991, cited in [4]). Theory provides little guidance as to which interactions might exist. Designing DCEs to empirically investigate such interactions would seem to us to be a priority.

In sum, we have shown that it is feasible to undertake DCE studies in low-income contexts, and that the results can be used to inform health financing policy. Our findings suggest that policy-makers will have difficulties in reconciling the demands of an equitable financing strategy with

those of a greater market orientation and will need to regulate hospital responses to more competitive environments. Together, the incentive regime in the current policy environment and the characteristics of demand for hospital services may encourage hospitals to segment demand by offering distinctive clinical qualities of care.

Acknowledgements

KH works for the Health Economics and Financing Programme, and BM is the Head, Health Systems Development Programme, both of which are funded by the UK Department for International Development (DFID). At the time this work was conducted, LA worked for the Health Economics and Financing Programme. This research was funded under a further project grant from DFID's HP-ACORD programme and a second project grant from the Swedish International Development Agency. The views expressed are those of the authors and cannot be attributed to either of the funding agencies. Data from the 1998 Living Conditions Monitoring Survey are used with the permission of the Central Statistical Office, Republic of Zambia. We are grateful to an anonymous reviewer for helpful comments and suggestions.

Note

a. In the case of attributes for which the SES interactions were significant (waiting time and cleanliness), the different WTP amounts arise from differences in the cost coefficient.

References

1. Johnson FR, Banzhaf MR, Desvousges WF. Willingness to pay for improved respiratory and cardiovascular health: a multiple-format, stated-preference approach. *Health Econ* 2000; **9**: 295–317.
2. Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. *Br Med J* 2000; **320**: 1530–1533.
3. Louviere JJ, Hensher DA, Swait JD. *Stated Choice Methods*: *Analysis and Applications*. Cambridge University Press: Cambridge, 2000.
4. Ryan M, Gerard K. Using choice experiments to value health care programmes: current practice and future challenges. In *Proceedings of the 3rd International Health Economics Association Conference*. University of York, UK, 22–25 July 2001.
5. Hakim Z, Pathak DS. Modelling the Euroqol data: a comparison of discrete choice conjoint and conditional preference modelling. *Health Econ* 1999; **8**: 103–116.

Appendix

Socioeconomic and demographic differences between those who could not pay, did not choose the superior options, and exhibited non-trading behaviour (p-values of Chi^2 test for binary variables and t-test for continuous variables).

	Those who said they could not pay any amount vs those who could		Those who failed to choose superior option vs those who did		Those who exhibited lexicographic preferences over price vs those who did not	
Ownership of:	Adult questionnaire	Child questionnaire	Adult questionnaire	Child questionnaire	Adult quesetionnaire	Child questionnaire
Motor vehicle	<0.001	<0.001	0.67	0.63	0.03	0.30
TV	<0.001	<0.001	0.007	0.005	<0.001	0.08
Video	<0.001	<0.001	0.11	0.001	0.003	0.08
Radio	<0.005	0.83	0.49	0.09	0.06	<0.001
Electric iron	<0.001	<0.001	0.005	<0.001	<0.001	0.001
Fridge	<0.001	<0.001	0.02	0.004	<0.001	0.003
Phone	<0.001	<0.001	0.07	0.003	0.001	0.13
Sewing/knitting machine	0.001	<0.001	0.28	0.52	0.02	0.91
Stove	<0.001	0.001	0.003	0.001	<0.001	0.005
Building	0.02	0.009	0.98	0.53	0.14	0.75
Sex	0.26	0.28	0.53	0.67	0.36	0.55
Age	0.75	0.38	0.23	0.83	0.06	0.48
Education	0.66	0.02	0.007	0.03	0.003	0.44
Household size	0.13	0.03	0.55	0.09	0.27	0.41

6. van der Pol M, Cairns J. Estimating time preferences for health using discrete choice experiments. *Social Sci Med* 2001; **52**: 1459–1470.
7. Scott A, Stuart Watson M, Ross S. Eliciting preferences of the community for out-of-hours care provided by general practitioners: a stated preference discrete choice experiment. *Social Sci Med* 2003; **56**: 803–814.
8. Scott A. Eliciting GPs preferences for pecuniary and non-pecuniary job characteristics. *J Health Econ* 2001; **20**: 329–347.
9. Scott A. Identifying and analysing dominant preferences in discrete choice experiments: an application in health care. *J Econ Psychol* 2002; **23**: 383–398.
10. McIntosh E, Ryan M. Using discrete choice experiments to derive welfare estimates for the provision of elective surgery: implications of discontinuous preferences. *J Econ Psychol* 2002; **23**: 367–382.
11. Jan S, Mooney G, Ryan M *et al*. The use of conjoint analysis to elicit community preferences in public health research: a case study of hospital services in South Australia. *Aust NZ J Public Health* 2000; **24**(1): 64–70.
12. Bryan S, Buxton M, Sheldon R *et al*. Magnetic resonance imaging for the investigation of knee injuries: an investigation of preferences. *Health Econ* 1998; **7**: 595–603.
13. van der Pol M, Cairns J. Establishing patient preferences for blood transfusion support: an application of conjoint analysis. *J Health Services Res Policy* 1998; **3**(2): 70–76.
14. Ryan M, Bate A, Eastmond CJ *et al*. Use of discrete choice experiments to elicit preferences. *Qual Health Care* 2001; **10**(Suppl I): i55–i60.
15. Ryan M, Hughes J. Using conjoint analysis to assess women's preferences for miscarriage management. *Health Econ* 1997; **6**: 261–273.
16. Ryan M. Using conjoint analysis to take account of patient preferences and go beyond health outcomes: an application to *in vitro* fertilisation. *Social Sci Med* 1999; **48**: 535–546.
17. Salkeld G, Ryan M, Short L. The veil of experience: do consumers prefer what they know best? *Health Econ Lett* 2000; **9**: 267–270.
18. Michaels J, Brazier J, Palfreyman S *et al*. Cost and outcome implications of the organisation of vascular services. *Health Technol Assess* 2000; **4**(11): 1–191.
19. Ratcliffe J, Buxton M. Patients' preferences regarding the process and outcomes of life-saving technology. An application of conjoint analysis to liver transplantation. *Int J Technol Assess Health Care* 1999; **15**(2): 340–351.
20. Farrar S, Ryan M. Response-ordering effects: a methodological issue in conjoint analysis. *Health Econ Lett* 1999; **8**: 75–79.
21. Ratcliffe J. Public preferences for the allocation of donor liver grafts for transplantation. *Health Econ* 2000; **9**: 137–148.
22. Republic of Zambia. *Living Standards Monitoring Survey*. Central Statistics Office: Lusaka, 1998.
23. Lake S, Musumali C. Zambia: the role of aid management in sustaining visionary reform. *Health Policy Plann* 1999; **14**(3): 254–263.
24. McPake B. Public autonomous hospitals in sub-Saharan Africa: trends and issues. *Health Policy* 1996; **35**: 155–177.
25. Nakamba P, Hangoro C, McPake B *et al*. Zambia case study. In *Health Sector Reforms: What About Hospitals?* Eriksson P, Diwan V, Karlberg I (eds). Nordic School of Public Health: Goteborg, 2002.
26. Nakamba P, Hanson K, McPake B. Markets for hospital services in Zambia. *Int J Health Plann Manage* 2002; **17**(3): 229–247.
27. Kamwanga J, Hanson K, McPake B *et al*. *Autonomous hospitals in Zambia and the equity implications of the market for hospital services*. Phase 1 Report: A Description of the Current State of Hospital Autonomy Policy. London School of Hygiene and Tropical Medicine: London, 1999.
28. McFadden D. *Conditional Logit Analysis of Qualitative Choice Behaviour*. University of California at Berkeley: Berkeley, 1973.
29. Maddala GS. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge University Press: Cambridge, 1983.
30. Kamwanga J, McPake B, Hanson K. *Community Perceptions of Hospital Quality of Care: Sub-study of a Project on Autonomous Hospitals in Zambia*. London School of Hygiene and Tropical Medicine: London, 2001.
31. Filmer D, Pritchett L. Estimating wealth effects without expenditure data–or tears: an application to educational enrollements in states of India. *Demography* 2001; **38**(1): 115–132.
32. Telser H, Zweifel P. Measuring willingness-to-pay for risk reduction: an application of conjoint analysis. *Health Econ* 2002; **11**: 129–139.
33. McPake B, Nakamba P, Hanson K. *Incentives and the Allocation of Resources within Public Tertiary Hospitals in Zambia*. Health Economics and Financing Programme WP 05/04. London School of Hygiene and Tropical Medicine: London, 2004.

