74489

USER GUIDE WITH CASE STUDIES

How to Conduct a Discrete Choice Experiment for Health Workforce Recruitment and Retention in Remote and Rural Areas

USAID FROM THE AMERICAN PEOPLE
CapacityPlus
World Health Organization
THE WORLD BANK

How to Conduct a Discrete Choice Experiment for Health Workforce Recruitment and Retention in Remote and Rural Areas:

A USER GUIDE WITH CASE STUDIES

WHO Library Cataloguing-in-Publication Data

How to conduct a discrete choice experiment for health workforce recruitment and retention in remote and rural areas: a user guide with case studies.

1.Rural health services 2.Health personnel. 3.Delivery of health care. 4.Medically underserved areas. 5.Personnel turnover. I.World Health Organization. II. United States. Agency for International Development. III.World Bank.

ISBN 978 92 4 150480 5 (NLM classification: WA 390)

Printed in France
Graphic design : Atelier Rasmussen / CH 2012

Foreword

An acute shortage of qualified health workers in remote and rural regions is a serious and widespread problem that is affecting many countries across the globe, but it affects the low income developing countries especially severely. In many developing countries, these shortages and distributional imbalances in health workforce present a major obstacle in the achievement of health-related Millennium Development Goals, and significantly hamper their progress toward Universal Health Coverage. Thus, finding effective and feasible policy solutions to this problem is a priority concern for many developing countries.

The Discrete Choice Experiment (DCE) methodology described in this User Guide is a quantitative research method that can measure the strength of preference and trade-offs of the health workers toward different job characteristics that can influence their decision to take up rural postings. This User Guide offers step-by-step advice on the application of DCE in identifying policy interventions appropriate to that country context in addressing health workforce shortages in remote and rural areas. The User Guide also includes two case studies from Tanzania and Uganda that demonstrate the application of this method in a real life context. It is intended to be used as part of a training program for the researchers involved in the design and implementation of DCE, but it could also be used as part of a broader training program on health services research and evaluation.

This User Guide is a product of close collaboration among the three agencies – The World Bank, World Health Organization and the USAID-funded Capacity*Plus* project – and represents our shared commitment to supporting policy-relevant research on critical topics related to Human Resources for Health. It is our hope that this User Guide will contribute to a systematic analysis and a deeper understanding of the factors that inhibit recruitment and retention of qualified staff in remote and rural regions, and will help countries develop their own unique solutions to lifting this problem that are well-adapted to their country context.

(Jointly signed by)

Nicole Klingman
Acting Director
Health, Nutrition and Population
The World Bank

Wim Van Lerberghe
Director
Department for Health Systems Policies and Workforce
World Health Organization

Kate Tulenko
Director, Capacity*Plus*
IntraHealth International

Acknowledgments

This User Guide was prepared as a joint effort by the USAID-funded Capacity*Plus*, World Health Organization (WHO), and The World Bank. The primary authors are: Mandy Ryan, Health Economics Research Unit, University of Aberdeen and Consultant for the World Bank; Julie R. Kolstad, University of Bergen, Uni Rokkan Centre and Consultant for the World Bank; Peter C. Rockers, Harvard University and Consultant for Capacity*Plus*/IntraHealth International; and Carmen Dolea, World Health Organization. The funding for this publication was provided by Norwegian Development Agency (NORAD), USAID through Capacity*Plus*, WHO, and the World Bank.

The material presented in this User Guide is based on teaching material Professor Mandy Ryan and Dr Verity Watson have developed for their 3 day residential workshop, "Using Discrete Choice Experiments in Health Economics: Theoretical and Practical issues".

The staff at the National Institute for Medical Research (NIMR), Tanzania, provided valuable contribution in developing and conducting the Discrete Choice Experiment in Tanzania. The Ministry of Health of Uganda also contributed significantly to the conduct of Discrete Choice Experiment in Uganda which forms the basis for the Ugandan case study.

Julie R. Kolstad, Margaret Kruk, Lindsay Mangham, Nonglak Pagaiya, and Peter Rockers undertook short surveys on logistical issues facing those conducting Discrete Choice Experiments to address human resource issues in low and middle income countries.

The following persons provided detailed comments and advice on the earlier drafts of this User Guide, and raised valuable issues during its preparation: Marco Alfano, Edson Araújo, Peter Berman, Duane Blaauw, Jean-Marc Braichet, Arne Risa Hole, Ayako Honda, Wanda Jaskiewicz, Margaret Kruk, Manuel M. Dayrit, Akiko Maeda, Kamolnat Muangyim, Nonglak Pagaiya, Krishna Rao, Amani Siyam, Laura Stormont, Marko Vujicic and Verity Watson. Comments were also received from Christoph Kurowski, Magnus Lindelow, Christophe Lemiere, Christophe Herbst, Jishnu Das and Rajeev Ahuja.

The authors would like to thank all the participants at the meeting organized by WHO, the World Bank and Capacity*Plus* in October 2010: "Tools for Implementing Rural Retention Strategies: Towards a 'How To' Guide for 'Discrete Choice Experiments' – A Methods Workshop", who set the basis for developing this User Guide.[1]

[1] WHO, 2010. "Tools for Implementing Rural Retention Strategies: Towards a 'How To' Guide for 'Discrete Choice Experiments'-A Methods Workshop". Meeting report. World Health Organization, October 2010, Geneva, Switzerland. Available online at: http://www.who.int/hrh/resources/DCE_report.pdf.

Table of Contents

Figures

Boxes

Tables

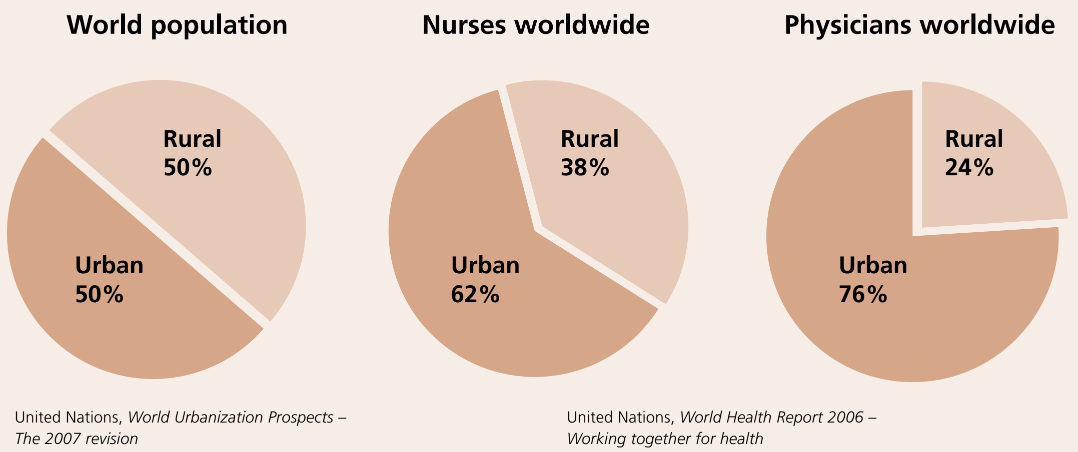
Acronyms

DCE	Discrete Choice Experiment
MXL	mixed logit
R	South African Rand
T Sh	Tanzania Shilling
USAID	United States Agency for International Development
U Sh	Uganda Shilling
WHO	World Health Organization
WTP	Willingness To Pay

Introduction

Skilled and motivated health workers in sufficient numbers at the right place and at the right time are critical to deliver effective services and improve health outcomes (WHO 2010). However, a shortage of qualified health workers in rural areas is common in both developed and developing countries. In most developing countries in the world more than half of their populations live in rural areas. The shortage is more pronounced in the 57 countries that the *World Health Report* of 2006 categorizes as being in a human resources for health crisis (WHO 2006). Globally, rural areas are served by only 38% of the total nursing workforce and by less than a quarter of the total physician workforce (figure 1). An estimated 1 billion people worldwide do not have access to health workers.

Figure 1 Urban/rural distribution of doctors and nurses worldwide



Source: WHO 2010.

The shortage of qualified human resources in low- and middle-income countries impedes access to health-care services for a significant percentage of the population, slows progress toward the Millennium Development Goals, and challenges the aspirations of achieving health for all (WHO 2010).

In 2010 the World Health Organization launched global policy recommendations on increasing access to health workers in remote and rural areas through improved retention of workers (WHO, 2010). These recommendations provide evidence-based guidelines on the effectiveness of interventions that can increase attraction, recruitment, and retention of health workers in remote and rural areas. The World Health Organization, the World Bank, and Capacity*Plus* (USAID's flagship global human resources for health project) are producing a road map for implementing the global recommendations. The road map will contain tools to help design, select, implement, monitor, and evaluate rural retention strategies. As well as this user guide, the road map will include a decision tool to choose the most appropriate strategies, a costing tool, and a monitoring and evaluation framework.

Objectives of the discrete choice experiment user guide

This guide aims to provide easy-to-read information and step-by-step advice on a quantitative research method that can help identify appropriate policy responses to health workforce shortages in remote and rural areas. This method can provide policy makers with quantitative measures of the relative importance of different job characteristics that influence the choice of health workers for rural postings.

The guide gives details about the types of questions such a method can (or cannot) answer, and the type of data, analysis, and analytical skills required to perform the research. It will use two case studies to illustrate the challenges and the ways to overcome them in conducting the work. Finally, it will provide information on the logistical and scientific requirements to perform such research and will offer links to resources for more detailed scientific and academic materials that can be consulted by advanced researchers.

What are discrete choice experiments?

Addressing the shortage of health workers in remote and rural areas has traditionally been approached from the perspective of *qualitative* surveys of small samples of students, practitioners, or key informants. Such surveys would identify a long list of factors that influence health worker decisions to go to, stay in, or leave a rural and remote practice location. These factors often followed a "laundry list" format—ranking salary, incentives, housing, career development, management style etc.—a format that proved daunting for policy makers, who would not know how to prioritize such factors.

Discrete choice experiments (DCEs)—a *quantitative* method for valuing different factors that influence job choices—has recently emerged as a very attractive method for researchers and policy makers alike, because it provides quantitative information on the relative importance of various jobs characteristics that influence the job choices of health workers, as well as the trade-offs between these factors and the probability of take-up of defined jobs. This method goes beyond the traditional qualitative assessments and provides quantifiable data that can better guide the selection of the most appropriate strategies for recruitment and retention in underserved areas. It also goes beyond the traditional ranking and rating exercises that do not provide information on strength of preference, trade-offs, or probability of take-up.

What can policy makers get out of DCEs?

DCEs are a quantitative methodology, or technique, for assessing stated preferences of health workers for a job, as a function of the job's characteristics. DCEs are useful for policy makers who want to investigate optimal packages of incentives or policy options to induce students in health professions or health workers to work in underserved areas, such as remote and rural locations.

More specifically, a well-conducted DCE can answer the following types of questions:

- What is the range of feasible and affordable policy interventions or incentive packages to address the rural shortages? (This information should be collected during the qualitative research phase, before starting a DCE, as part of the situation analysis).
- What characteristics of a job (or conditions of employment) are most important for newly graduated health workers to take up rural postings?
- What characteristics of a job (or conditions of employment) are most important in encouraging practicing physicians or nursing to remain in rural or remote locations?
- How much of the salary are rural health workers willing to trade for nonmonetary incentives or for other characteristics of the job?
- How do individual characteristics (such as sex, rural origin, socioeconomic status, intrinsic and extrinsic motivations) affect the preference for rural postings?
- What proportion of health workers will accept defined job postings, if given specific incentives?

When to conduct a DCE?

Policy makers who want to assess or introduce new strategies on increasing recruitment and retention in remote and rural areas can use a DCE to get answers to the above questions. Before conducting a DCE, however, a thorough situation analysis should be conducted to inform the current level and conditions of employment, as well as the barriers and opportunities for filling rural positions. A DCE should therefore be seen as one component of broader policy review and planning of human resources for health (WHO, 2010).

Unlike studies of revealed preference (which means actual choices) DCEs can also be used to estimate the effect of policies yet to be implemented, such as salary increases for uptake of rural posts. This makes DCEs helpful for planning future policy reforms. Ideally, DCEs should be followed by real-world experiments that present the preferred package to the target population.

Who is this guide intended for?

- Policy and decision makers (including health care managers and public health program leaders) who are interested in examining the question of recruitment and retention in more detail and want to use this information to improve the design of their staffing and recruitment programs.
- Researchers in the field who want to familiarize themselves with the technical requirements of conducting such a DCE. Certain sections of the guide are relatively straightforward (some steps of the method, for example) while others require a higher level of statistical and methodological skill.

How to use this guide

Section 1 (the step-by-step guide) gives an overview of the stages involved in conducting a DCE, and the issues that can arise. For each step, a summary box highlights the key messages, followed by detailed explanations of the concepts, data requirements, and statistical approaches. This section uses examples from published DCEs to illustrate the various elements of each step. It offers good practical guidance to the reader, along with further details, useful technical resources, and references to software packages.

A small subsection pulls together information on logistical requirements and challenges in conducting a DCE. Section 1 then offers some concluding comments.

Sections 2 and 3 provide two case studies—one each from Tanzania and Uganda—to illustrate in more depth the practicalities of conducting the work. The first, a study by Kolstad (2011), describes a DCE in Tanzania that included students in clinical officer training programs. The second, by Rockers et al. (2012) describes a DCE in Uganda that included three in-service health worker cadres: medical officers, nursing officers, and laboratory technicians.

A reference list is provided at the end, broken down into areas. Useful websites for software and other statistical methods or tools are also given.

Important points to note before conducting a DCE

While DCEs can help policy makers devise appropriate incentive packages to attract and retain health workers, a number of important points should be noted before a DCE is implemented:

- This User Guide provides guidance on using DCEs to understand the preferences, and potential resulting costs, of creating incentives to induce health works to move to remote/rural areas. No statement is made concerning the benefits of health workers relocating to remote and rural areas as in some contexts it may not be the most efficient or optimal intervention. Future research is encouraged in this area.

- Related to the above point, while increasing the number and availability of health workers in rural/remote regions could lead to better access to health services in these under-served regions, it is not the only solution for addressing the low density of health workers in these areas, and it may not even be a feasible solution in some country contexts. There are alternative options, such as restructuring the health care delivery model to expand on outreach services that will involve health professionals in a centralized location (e.g., district hospital) to make periodic visits to outlying communities which may have community health workers and midwives. Such alternative care models are being tested in a number of countries, e.g., Nepal and Yemen, and will require further validation, but it could reduce the density requirements of health professionals in some countries where the highly dispersed nature and remoteness of communities make it infeasible and/or unproductive to increase the number of health professionals in these areas.

- The stability of the policy environment should be taken into account when considering using a DCE. For example, in India there are currently a series of policy changes addressing HRH issues. These include considering: creating a new institution to regulate standards for health education; introducing a Bachelor of rural health care, a shorter course than the MBBS with the aim of creating a cadre of rural health professions who are trained to provide primary health care in rural areas; and the introduction of a compulsory rural posting for all MBBS students. The timing of the DCE needs to be considered alongside such policy initiatives, as well as the way the DCE can inform such policy initiatives.

- Attributes included in the DCE should be amenable to change. For example, recommended incentive strategies emerging from a DCE, such as increased compensation or improved employment conditions, should be able to be implemented within the given country context.

- Further, policy makers and health planners concerned with implementing HRH policy change would need to be aware of the usefulness of the technique, and play an active role in seeing its benefit and promoting its use. Efforts should be made by the DCE practitioner to communicate and translate the results of a DCE to the policy makers.

- The DCE is only one of the approaches that could be used to examine the impact of factors that influence recruitment and retention of health workers in remote and rural regions. Qualitative information is very useful in informing policy makers about the factors that determine job choices. It also has the appeal that the results would be immediately understandable by policy makers. However, DCEs do provide extra information, providing evidence not just on what is important, but on the strength of preference for given policy changes, trade-offs between given job attributes, and probability of take-up of specified jobs. Such information cannot be collected in qualitative research. If actual data sets exist regarding job choices then the information obtained from a DCE could also be obtained from such data sets. However, data sets on actual choices are often limited, either they don't exist or the information provided is incomplete. DCEs are therefore often the only option for gathering information on strength of preference, trade-offs and probability of take-up.

- Stated preference methods (where individuals state their preferences to hypothetical choices) may be preferred to revealed preference data (where actual behaviour is observed) when testing out new policies, especially since a revealed preference data would not be available without implementing the reform. However, given DCEs rely on responses to hypothetical choices it is important that interventions that aim to implement the findings from DCE studies validate such findings through subsequent monitoring and evaluation of policies.
- DCE is a method for looking at preferences, and thereafter devising incentives to move health workers from urban to rural areas. The use of DCEs does not exclude the need for a HRH labour market analysis, indeed it reinforces it. A labour market analysis is required to estimate key elements of the health workforce in the country (e.g., current number and distribution of health workers by skills, geographical location, gender, distribution across public and private sectors, the existence of dual practice) and identify the main factors leading to a mismatch between demand and supply of health workers in rural and remote areas.
- Additionally, DCE focuses on supply side behaviour and the data collected aims to inform policy design to change such a behaviour. A labour market analysis is then necessary to provide information on the labour market demand (the funding to hire health workers). In some contexts the need to increase the health workers supply may be restricted by the limited funds to hire them (insufficient demand). Therefore, prior to conduct any DCE a labour market analysis is necessary to estimate the supply and demand patterns of health workers in rural versus urban areas.

1. Step-by-step guide on how to conduct DCEs in low- and middle-income settings

This section gives to the reader an overview of discrete choice experiments (DCEs), the stages involved in its implementation, and the policy-relevant output that can be generated from a DCE. Given the focus on applying DCEs to addressing the shortage of medical staff in remote and rural areas in low- and middle-income countries, the points in this section are illustrated using examples from the studies listed in the appendix.

1.1 Discrete choice experiments: what are they and what can practitioners do with them?

SUMMARY OF 1.1

This section introduces DCEs through examples of studies addressing issues of health workforce rural retention. It explains what kind of research method it is, and what information it can provide.

DCEs have been applied in both the developed and developing world to assess how job attributes influence job choice. Increasingly, they are being used to inform policy making related to the attraction and retention of health workers in underserved areas.

In a DCE, respondents are presented with a number of hypothetical job choices that vary with respect to attributes and levels, such as salary, housing, and opportunities for career advancement. The responses can be used to obtain the following information for the policy level:

- Which job attributes are important and how important one attribute is in comparison to another.
- How much salary a health worker would be willing to give up for improvements in others attributes of a job, that is, how much health workers value other attributes of a job relative to their salary.
- The probability of respondents taking up a job with specified attributes.

1.1.1 DCEs: an attribute-based measure of value

DCEs are an *attribute*-based measure of benefit. They assume that individual decisions with regard to a good or service are determined by the attributes or characteristics of that good or service. For example, they have been used in transport economics to look at how transport choices are determined by attributes such as "mode of travel", "cost of travel", "travel time," and "comfort" (Hensher et al. 2005, and see box 1.1). The technique has also been used increasingly in health economics to look at attribute importance in delivering health care, with consideration both to the aspect of the patient experience and health outcomes, as well as to trade-offs between these and willingness to pay (WTP) for different attributes (see de Bekker-Grob et al. 2012 for a summary).

Box 1.1 General reading

DCEs have become a common technique in economics, addressing a wide range of policy questions in transport economics (Hensher et al. 2005), environmental economics (Hanley et al. 2001), and health economics (Ryan et al. 2008; de Bekker-Grob et al. 2012). Papers by Lancsar and Louviere (2008) and by Bridges et al. (2011) provide solid general guidance on good practice when conducting a DCE, and Train (2009) provides some additional guidance on conducting a DCE.

The technique has been applied to address issues around recruitment and retention of health workers in developed countries (Gosden et al. 2000; Scott 2001; Ubach et al. 2003; Wordsworth et al. 2004; Scott et al. 2007; Sivey et al, 2010) and in low- and middle-income countries (see appendix). Within low and middle income countries studies have been concerned with health worker preferences (Penn-Kekana et al. 2005; Mangham and Hanson 2008; Hanson and Jack 2010; Vujicic et al. 2010a); student preferences (Chomitz et al, 1998; Kruk et al. 2010; and Blaauw et al. 2010; Kolstad 2011;) and both health workers and students preferences (Vujicic et al. 2010b; Jaskiewicz et al., 2012; McAuliffe et al., forthcoming; Rao et al., forthcoming; Rockers et al., 2012)

For applications to low- and middle-income countries, the paper by Mangham et al. (2009) discusses some of the issues applied in designing a DCE, and the paper by Lagarde and Blaauw (2009) reviews the application and contribution of DCEs to inform human resources policy interventions. The book by Bennett and Birol (2010) also considers the application of DCEs to these countries. Although the book discusses environmental and agricultural economics, some of the issues raised are relevant for eliciting preferences of medical students and health workers on working in remote and rural areas.

Within the context of the health workforce, DCEs are a useful technique because job choice is known to be determined by the characteristics or attributes of the job. They have been applied in both the developed and developing world to assess how job attributes influence job choice (See table 1.1 and the appendix for examples of the application of DCEs to workforce issues). In the earliest study published in a low- and middle-income country context, Chomitz et al. (1998) applied the technique to develop incentives for doctors to serve in Indonesia's rural remote areas. Attributes (or factors determining choice) included province, remoteness, total monthly income, length of contract, probability of subsequent appointment to the civil service, and probability of subsequent specialist training. The range of objectives of a DCE is illustrated in table 1.1.

Table 1.1 Range of objectives of human resource DCE studies

Study	Country	Study objectives: to investigate…
Chomitz et al. 1998	Indonesia	doctors' preferences regarding various possible incentives, in particular to attract them to rural or remote places
Gosden et al. 2000	England	preferences of general practitioners (GPs) for practice and job characteristics, to understand what factors might improve GP recruitment in underserved areas
Scott 2001	United Kingdom	GPs' preferences for financial and nonfinancial incentives
Ubach et al. 2003	Scotland	the strength of hospital consultants' preferences for various aspects of their jobs, to improve recruitment and retention
Wordsworth et al. 2004	Scotland	the relative value given by sessional GPs to various job characteristics, to inform issues on recruitment and retention of GPs
Penn-Kekana et al. 2005	South Africa	the relative importance of various job characteristics, to explain staff dynamics
Mangham and Hanson 2008	Malawi	the range and relative importance of various factors that affect nurses' job choices in the public sector

Study	Country	Study objectives: to investigate…
Hanson and Jack 2010	Ethiopia	the effects of possible policy interventions, to improve the supply of doctors/nurses in rural areas
Kolstad 2011	Tanzania	clinical officers' job preferences, to understand how rural jobs can be made more attractive
Blaauw et al. 2010	Kenya, South Africa, Thailand	the effects of various policy incentives to attract staff to rural areas
Kruk et al. 2010	Ghana	
Sivey et al. 2010	Australia	what influences the choice of senior medical students for rural practice posts
		the preferences of junior doctors for the different attributes of specialties
Kruk et al. 2010		what influences the choice of senior medical students for rural practice posts
Sivey et al. 2010	Australia	the preferences of junior doctors for the different attributes of specialties

Source: Blaauw and Lagarde 201O, WHO, 2010.

Within a DCE respondents are presented with hypothetical job choices that vary by job attributes and levels of those attributes (section 1.2.2, figures 1.1–1.3). The responses are analyzed using regression techniques and can be used to obtain different pieces of information that are useful at the policy level (section 1.2.5).

Relative importance of job attributes

Information is provided on whether the attributes are important (statistically significant), the direction of importance (sign of the estimated parameter), and relative importance (size of the estimated parameter). For example, Lagarde and Blaauw (2009) note that studies applying DCEs to inform human resource policy interventions have shown that nonmonetary incentives are significant determinants of job choice, and sometimes more important than financial ones.

But while the above information is very useful, the real value of DCEs is in looking at the trade-offs that respondents are willing to make between attributes as well as the probability of take-up of defined posts. This type of information cannot be obtained from detailed focus groups or interviews, nor indeed from current longitudinal data when new policies are being implemented, since such data will not exist.

Trade-offs

Information on trade-offs is useful to policy makers because health workers can rarely have the best levels of all the factors important to them (owing to limited resources). Estimation of trade-offs allows policy makers to estimate how much of one attribute a health worker would be willing to give up to have an improvement in another. To estimate trade-offs, a continuous attribute must be included in the DCE. Within the job choice literature, this continuous variable is commonly salary. Inclusion of this attribute allows estimation of how improvements in aspects of a job, such as housing and education opportunities, can compensate for lower wages—that is, how much salary an individual would be willing to give up for improvements in others attributes of a job. For example, Mangham and Hanson (2008) were interested in the extent to which nurses in Malawi were willing to trade monetary for nonmonetary benefits.

Probability of take-up

DCEs also allow estimation of the probability of individuals taking up a job with specified attributes. Such information is very useful when policy makers look at job choice because policy is concerned with creating jobs to encourage medical students and health workers to take up posts in remote and rural areas. For example, Hanson and Jack (2010) found that doubling wages in areas outside the capital would increase the share of doctors willing to work there from about 7% to 50%; providing high-quality housing would increase physician supply to about 27% (equivalent to paying a wage bonus of about 46%); doubling wages paid to nurses for work in rural areas outside cities would increase their labour supply from 4% to 27%; and that the nonwage attribute that is most effective in inducing them to relocate to rural areas is the quality of equipment and drugs.

All the above information can be estimated for the total sample, or for subgroups of the population. Subgroup analysis is common in analyzing DCE data. For example, Kruk et al. (2010) looked at how preferences differed between women and men, finding that women's preferences were more influenced by supportive management and men's preferences by superior housing.

1.1.2 DCEs: a survey-based measure of value

As well as being attribute-based, DCEs are *survey*-based. That is, they rely on what respondents *say* they will do—also referred to as stated preference data—rather than what they *actually* do—referred to as revealed preference data. The literature shows a healthy skepticism for relying on stated preference data compared with revealed preference data, but there are good reasons for researchers' interest in stated preference data (Ryan et al. 2008). A key advantage of this hypothetical approach is that it allows preferences to be elicited for job profiles that do not exist (hence revealed preference data are unavailable). Indeed, Lagarde and Blaauw (2009) note that all the studies they reviewed used one of the key strengths of a DCE—including job attributes that do not currently exist but that could potentially influence choice of job, such as salary ranges, beyond that which exists and delinking the *opportunity* to specialize from the *obligation* to do so in the public sector (Chomitz et al. 1998).

It is useful to contrast DCEs with randomized experiments for public policy, which would be a form of eliciting revealed preferences. Randomized experiments would be constrained by the range of job opportunities available. While a DCE commonly presents individuals with a number of hypothetical choices (often between 16 and 32), it would be hard to offer individuals such a range of job choices in reality.

The hypothetical nature of DCEs also allows the independent variables to be identified in advance (via experimental design methods—see section 1.2.2 on experimental design), which allows identification of all effects of interest. This contrasts with revealed preference data, which cannot be controlled a priori so that model identification cannot be guaranteed because multi-collinearity may be present. Moreover, the use of revealed preference data is limited in most developing countries given the lack of data. Stated preference methods also allow large quantities of data to be collected at moderate cost.

1.2 Stages in conducting the DCE

The DCE has several key stages:

- identification of attributes and assignment of levels (section 1.2.1);
- experimental design: deciding what choices (job profiles) to present to individuals (section 1.2.2);
- development and administration of the survey (data collection) (section 1.2.3);
- data input (section 1.2.4); and
- analysis and interpretation (section 1.2.5).

Given that DCEs involve responses to hypothetical choices, it is crucial that each stage of a DCE is carried out well. Failure to do so may result in numbers that lack validity. The following sections describe in detail the stages involved.

1.2.1 Identification of attributes and assignment of levels

SUMMARY OF 1.2.1

The first stage in conducting a DCE is to identify the attributes of the job and levels of those attributes that are important to health workers in the local setting.

The selection of attributes and their levels should be:

- Informed by the literature on human resources for health and job choice
- Realistic and actionable by policy, as well as informed by policy makers and managers
- Based on qualitative work such as focus group discussions or interviews with health workers in the local setting.

The first stage in conducting a DCE is to identify the attributes of the job and levels of those attributes that would influence a health worker's decision to work in a remote or rural area. One example is shown in table 1.2.

Table 1.2 Attributes and levels within a (labeled) DCE

	Level	
Attribute	**Rural position**	**Urban position**
Type of facility	1. Clinic 2. Hospital	Clinic Hospital
Annual salary	1. R120,000/year 2. R120,000 + an additional R12,000/year 3. R120,000 + an additional R24,000/year 4. R120,000 + an additional R36,000/year	1. R120,000/year
Provision of subsidized housing	1. Basic: single room with a shared kitchen and shared toilet. 2. Superior: small two bedroom house for you and your family.	1. None 2. Basic: single room with a shared kitchen and shared toilet.
Time to wait before getting study leave to specialize	1. Normal: 6 years 2. Improved: 2 years	1. Normal: 6 years
Car allowance	1. None 2. R500 per month	1. None
Number of years to be spent in the facility until being eligible for promotion	1. Normal: 2 years 2. Improved: 1 year	1. Normal: 2 years 2. Improved: 1 year
Workplace management and culture	1. Hierarchical: this facility is formal and structured. The managers emphasise stability, following rules, and keeping things running smoothly. 2. Relational: this facility is personal and supportive. The managers emphasise teamwork, loyalty, and developing the full potential of staff.	1. Hierarchical: this facility is formal and structured. The managers emphasise stability, following rules, and keeping things running smoothly. 2. Relational: this facility is personal and supportive. The managers emphasise teamwork, loyalty, and developing the full potential of staff.

Source: Based on Blaauw et al. (2010). R = South African rand.

Identifying the attributes of the job and their levels is a key step in conducting the DCE, because this will inform the subsequent formulation of job choices. This step involves literature reviews and qualitative research, such as in-depth interviews and focus groups involving students, health workers, and policy makers (box 1.1).

Box 1.2 The importance of qualitative work

The DCE literature is increasingly recognizing the importance of qualitative work to derive the attributes and associated levels (for example, Coast and Horrocks 2007; de Bekker-Grob et al. 2012) and such methods have been applied in the workforce area.

Lievens et al. (2009) provide a useful guide to the use of qualitative research methods when understanding health workforce issues. Rao et al. (2010) report on detailed qualitative research that fed into deriving attributes and levels for the DCE concerned with how to attract health workers to rural areas in India.

Mangham and Hanson (2008) conducted in-depth interviews with 20 registered nurses working in three different districts in Malawi, as well as in primary, secondary, and tertiary health facilities, to establish attributes. Kolstad (2011) conducted 20 in-depth interviews with clinical officer students in one rural and one urban location to inform the development of attributes and levels (see the Tanzania case study).

Kruk et al. (2010) conducted seven focus groups with medical students when eliciting rural practice preferences among medical students in Ghana. Similarly, Rockers et al. (2012) conducted several focus groups in Uganda with members of medical, nursing, pharmacy, and laboratory cadres to inform DCE design for a study of job preferences among students and in-service workers in the country (see the Uganda case study).

Before embarking on qualitative work, the practitioner should seek the advice of an experienced qualitative researcher (or ideally have him or her involved in the project). Existing focus group discussions or in-depth interviews specific to the context may exist and can be used. Literature on job choice and the global shortage of health workers, as well as context-specific policy documents, can also inform what attributes to include (Mangham and Hanson 2008).

Inclusion of the concerns of policy makers and managers is also important. For example, Mangham and Hanson (2008) included an attribute concerned with the provision of government housing because the Ministry of Health was interested in how the availability and quality of government housing affected the retention of health personnel. Such inclusion will also increase the chances that the results are taken into account at the policy level.

The number of attributes to include in the DCE is important. When individuals respond to the choices, it is assumed that they are considering all the attributes, and making trade-offs among them. It is this assumption that allows such trade-offs to be estimated, and therefore monetary values to be estimated. One concern is that, if too many attributes and levels are included, individuals will not consider all the information, but adopt simple decision-making strategies (such as always choosing the option with the highest pay). If this is the case, estimated trade-offs will not be valid.

The question is then raised: What is too many attributes? Applications of DCEs in health economics have included anywhere between two and 24, with a mode of six (de Bekker-Grob et al. 2012). Within the context of applications to workforce issues in low- and middle-income countries, attribute numbers have ranged from five to eight (see appendix). When conducting a DCE, it is important to investigate the acceptable number of attributes within the pilot work as this is likely to be context specific, although eight is generally seen as approaching the maximum.

Levels may be defined *continuously* or *categorically*. Continuous variables can have any numeric value. From table 1.2, examples of continuous attributes include annual salary, amount of car allowance, time to wait before getting study leave to specialize, and number of years to be spent in the facility until being eligible for promotion. As mentioned, inclusion of continuous variables allows estimation of trade-offs, a very useful output at the policy level.

Categorical variables refer to variables where the levels belong to categories. Such variables may be described descriptively. For example (again from table 1.2), provision of subsidized housing is defined as “basic” (single room with a shared kitchen and shared toilet) or superior (small, two-bedroomed house for you and your family); and workplace management and culture is defined as “hierarchical” (this facility is formal and structured, the managers emphasize stability and follow rules while keeping things running smoothly) or “relational” (this facility is personal and supportive, the managers emphasize teamwork, loyalty, and developing the full potential of staff). Some categorical variables may be defined as ordinal—here it is known that one level is better (or worse) than another, but not by how much (as distinct from continuous variables where, for example, four years is twice two years). For example, “provision of subsidized housing” is ordinal since superior provision of housing is better than basic provision of housing.

When describing categorical variables it is important to define clearly what is meant by the levels since they must be interpreted by respondents in the way intended by those who designed the questionnaire. For example, given the focus of studies on inducing students and medical staff to work in rural areas, it is crucial that respondents have a common understanding of what “rural” and “urban” mean. Studies have defined this in different ways (see appendix).

For example, Chomitz et al. (1998) defined a “remoteness” attribute (non-remote, remote, or very remote); Mangham and Hanson (2008) a “place of work” attribute (city or district town); Hanson and Jack (2010) a “location” attribute (Addis Ababa versus regional capital); and Kolstad (2011) a “location” attribute (Dar es Salaam; regional headquarters; district headquarters and “a three-hour or more bus ride from the district headquarters”).

Blaauw et al. (2010) used what is known as a labeled design, using “rural” and “urban” labels. Here, rather than define a location attribute, each choice was labeled as either “rural job” or “urban job” (table 1.2). When using a labeled design, as with defining categorical attributes, the researcher would need to define the labels well. Blaauw et al. (2010) defined rural facilities as located in small villages or remote rural areas where infrastructure is poorly developed and access to services such as shops and schools may be limited. Urban facilities were defined as located in cities or large towns with well-developed facilities and good access to all services.

Qualitative pilot work will be invaluable in helping define levels for the attributes and levels.

Researchers often define attributes and levels within realistic levels that are potentially actionable by policy. In essence, there is no point offering respondents a salary that is so high it is unrealistic, nor nonmonetary benefits that cannot be implemented at the policy level. Having said that, the hypothetical nature of DCEs means that it is possible to extend options beyond the current policy space—as long as the boundaries can realistically be extended.

Constructing the attributes and levels for a DCE is as much an art as science, with the researcher trying to capture as much relevant information from the qualitative work in the attributes and levels. These need to be put together such that the job descriptions reflect an actual job choice the health worker could face.

1.2.2 Selection of experimental design and construction of choice sets

SUMMARY OF 1.2.2

It is often not possible to present respondents with all the choices of hypothetical DCE job scenarios (the "full factorial" design). Experimental design methods are used to select a reduced sample of choices (a "fractional factorial" design) for respondents.

Experimental design methods consider (to varying degrees) the following when identifying choices to present to respondents:
- There should be minimal correlation between different attribute levels as they appear in the DCE (orthogonality).
- Each attribute level should appear roughly an equal number of times in the DCE (level balance).
- Two job scenarios that appear together in a choice set should rarely have the same attribute levels (minimal overlap).

Also important is whether interaction terms will be included, the appropriateness of an opt-out option, a generic versus labeled design, and the number of DCE questions a respondent can answer before becoming tired, bored, or unmotivated.

Constructing the experimental design
Having established the relevant attributes and their levels, hypothetical job choices with different combinations of attributes and levels must be constructed and presented to individuals.

Figures 1.1–1.3 show examples of choices presented within a DCE. Figures 1.1 and 1.2 are generic choices, where the job label (Job A or Job B; Job 1 or Job 2) do not mean anything in themselves. Figure 1.3 is a labeled choice (where the jobs are defined as rural facility or urban facility); it is a choice set derived from the set of attributes and levels in table 1.2.

Figure 1.1 Example of a choice from Kolstad (2011)

Job A						
Availability of equipment and drugs	**Housing**	**Education opportunities/ possibility of upgrading qualifications**	**Workload**	**Infra-structure**	**Salary and allowances**	**Location**
Sufficient	No house is provided	Education offered after 6 years of service	Normal: Nearly enough time to complete duties. One hour of extra work per day	The place has mobile coverage, electricity and water	T Sh 350,000 per month	Regional headquarters

Job B						
Availability of equipment and drugs	**Housing**	**Education opportunities/ possibility of upgrading qualifications**	**Workload**	**Infra-structure**	**Salary and allowances**	**Location**
Insufficient	A decent house is provided	Education offered after 2 years of service	Heavy: Barely enough time to complete duties. Three hours of extra work per day	The place has unreliable mobile coverage, no electricity or water	T Sh 500,000 per month	A 3-hour or more bus ride from the district head-quarters

Considering your current situation, which of the two jobs would you choose?

Job A: ☐ **Job B:** ☐

Figure 1.2 Example of a choice from Vujicic et al. (2010a)

In this section of the questionnaire we want to try and understand what type of nursing jobs you most prefer.

We will be doing this by presenting you with two different nursing jobs and then asking you tell us which you prefer. You will see that each job has advantages and disadvantages and you will nee to carefully trade-off the advantages and disadvantages in telling us which job you prefer.

For each pair of jobs, we would also like to know whether you would accept this job over your current job if the Ministry of Health offered it to you.

You can assume that the length of service in all jobs is 3 years.

Job 1		Job 2	
Location	Urban	Location	Urban
Equipment	Inadequate	Equipment	Adequate
Total pay	160	Total pay	240
Transportation	Yes	Transportation	Yes
Housing	No	Housing	Yes
Workload	Heavy	Workload	Normal

Which of these two jobs do you prefer? ☐ Job 1 ☐ Job 2

Would you accept this job over your current job? ☐ Yes ☐ No

Figure 1.3 Example of a choice from Blaauw et al. (2010)

Which of these two public sector facilities would you choose to work in?

	RURAL Facility	URBAN Facility
Type of facility	Clinic	Hospital
Monthly salary	R120,000 per year	R120,000 per year
Rural allowance	An additional R12,000 per year	None
The number of years you would have to work before getting study leave to specialise	2 years	2 years
The housing provided	You can choose to stay in the subsidised accommodation which is a single room with a shared kitchen and shared toilet	None
The number of years you would have to work before being eligible for promotion	2 years	2 years
The car allowance offered	None	None
The workplace culture and style of management	This facility is formal and structured. The managers emphasise stability, following rules, and keeping things running smoothly.	This facility is personal and supportive. The managers emphasise teamwork, loyalty, and developing the full potential of staff.
Which facility would you choose?	**Rural facility** ☐	**Urban facility** ☐

From the total number of possible combinations of attributes and levels the question then arises, which job descriptions (choices) should be presented? The number of possible job descriptions is determined by the number of attributes and levels. For example, if there are 3 attributes, all at 2 levels, the total number of job profiles is 8 (2^3). For simplicity, assume 3 attributes, with 2 levels (a reduced set from Penn-Kekana et al. 2005): *equipment* (fully equipped or poorly equipped); *staffing* (well-staffed or understaffed); *social amenities* (underdeveloped or developed). The total number of job profiles is 8 (figure 1.4).

Figure 1.4 Experimental design for 3 attributes with 2 levels each (full factorial)

		Equipment	Staffing	Social amenities
Job profiles	1	poorly equipped	well staffed	underdeveloped
	2	poorly equipped	well staffed	developed
	3	poorly equipped	Understaffed	underdeveloped
	4	poorly equipped	Understaffed	developed
	5	fully equipped	well staffed	underdeveloped
	6	fully equipped	well staffed	developed
	7	fully equipped	Understaffed	underdeveloped
	8	fully equipped	Understaffed	developed

(numeric representation)

		Equipment[1]	Staffing[2]	Social amenities[3]
Job profiles	1	0	1	0
	2	0	1	1
	3	0	0	0
	4	0	0	1
	5	1	1	0
	6	1	1	1
	7	1	0	0
	8	1	0	1

Note:
1: 0=poorly equipped; 1=fully equipped.
2: 0=understaffed; 1=well-staffed.
3: 0=underdeveloped; 1=developed.

However, studies usually have more attributes and levels, resulting in more possible choices. Essentially, a study with 4 attributes at 4 levels would result in 256 possible scenarios (4^4=256). More generally, the number of possible scenarios is a^n where *a* is the number of levels and *n* is the number of attributes. If the number of possible levels varies across attributes, the number of possible hypothetical scenarios is $a^n \times b^m$ where *a* and *b* are the different attribute levels and *n* and *m* the different attributes.

Further, where choices are presented, with each choice set involving 2 options, the number of possible choices can become very large. For example, with 256 possible scenarios, and each choice including 2 options, there would be [256*255]/2 unique choice sets i.e. 32,640.

Presenting respondents with all possible choices is known as a full factorial design (as in figure 1.4). However, this is often not possible, since it generates too many choice sets, and fractional factorial designs are used to reduce the profiles for which preferences are elicited. As their name suggests, they are a fraction of the total number of possible choice sets, and are derived using experimental design methods. Use of experimental design methods allow preferences for all job profiles to be identified, not just those presented in the DCE.

When deciding on the fractional factorial design, the researcher will need to state if he or she wants interaction terms included in the model. This refers to the preferences of one attribute being determined by the levels of another attribute. For example, are the preferences for salary influenced by the level of education provided, or the preferences for housing influenced by location? Inclusion of interactions results in the need to present respondents with more choices. (A full factorial design, as in figure 1.4, allows all interaction terms to be estimated). It is common practice in the DCE literature to include only main effects, since it is argued that such effects explain most of the variation in preferences (de Bekker-Grob et al. 2012).

This practice is also common in the human resource literature. For example, Rockers et al. (2012) present only main effects estimates (see the Uganda case study). However, the study by Kolstad (2011) allowed for interaction effects between housing and wage, education and location, and workload and equipment (as these attributes to some extent appeared to be interrelated in the interviews). However, the interaction effects did not prove to be significant in the analysis, and were thus excluded for simplicity in the final analysis (see the Tanzania case study, section 2.1).

How to get an appropriate experimental design?

When using experimental design methods to reduce the full set of scenarios (full factorial design) down to a manageable level (fractional factorial design), orthogonal designs are often used. These are based on orthogonal arrays, which are readily available in design catalogues (for example, Hahn and Shapiro 1966); statistical programs (http://www.spss.com; Kuhfeld 2010; http://www.sawtoothsoftware.com), or websites (Sloane 2009; http://www.research.att.com/~njas/oadir). These arrays have the properties of orthogonality (attributes are statistically independent of one another) and level balance (levels of attributes appear an equal number of times). Such properties can seem appealing intuitively.

What is orthogonality?

One common interpretation is to look at the correlations between two attributes and, if this is 0 or low, define it as orthogonal. Thus, orthogonality can be seen as the opposite of multi-collinearity, where attributes move together, and it is not possible to identify the independent effect of attributes. If high levels of multi-collinearity exist between variables (as is often the case with revealed preference data), it is not possible to identify what attribute is driving preferences. In the worst case, the regression model will not run—it will not produce results. It is thus important to ensure that high levels of multi-collinearity do not exist.

Table 1.3 shows the correlations for the full factorial design in figure 1.4. These correlation coefficients were estimated using a range of tests, including Pearson's correlation coefficient (Pearson's Product of Moments, or PPM), Kendall's Tau-b, and Spearman's Rho (within SPSS). All showed zero correlation, that is, perfect orthogonality.

Table 1.3 Correlation matrix for the design in figure 1.4

	equipment	staffing	social amenities
equipment	1	0	0
staffing	0	1	0
social amenities	0	0	1

The PPM is the most commonly used correlation coefficient, but while popular, it is strictly valid for continuous variables. When this is not the case other correlation measures are more appropriate, such as the G-Index or Kendall's Tau-b for two dummy variables or the Spearman rank correlation for two ordinal variables. In such situations the PPM should be seen as an approximation (Hensher et al. 2005) (for more on appropriate tests when the attributes are not on a ratio scale, see Hensher et al. (2005) or any introductory statistics textbook). The popularity of the PPM has resulted in software packages often offering limited correlation measures.

What does level balance mean?

The second criterion, level balance, means that attribute levels appear an equal number of times, ensuring that all levels have an equal chance of being chosen. Table 1.4 reports the full factorial design in figure 1.4, showing that all attributes appear the same number of times and the perfect level balance.

Table 1.4 Level balance for design in figure 1.4

Attribute and levels	Number of appearances	Staffing
Equipment		
Fully equipped	4	50
Poorly equipped	4	50
Staffing		
Well staffed	4	50
Understaffed	4	50
Facility mix		
Good	4	50
Poor	4	50

Creating choice sets

When the researcher uses catalogues and websites, he or she derives a set of profiles (job descriptions) and has to create "choice sets." A number of approaches are available. One is to offer a binary choice, that is "Would you take up this post?", with responses being Yes or No (figure 1.5).

Figure 1.5 Binary choice questions

	Equipment	Staffing	Social amenities	Would you take this post up? Yes	No
1	Poorly equipped	Well staffed	Underdeveloped	☐	☐
2	Poorly equipped	Well staffed	Developed	☐	☐
3	Poorly equipped	Understaffed	Underdeveloped	☐	☐
4	Poorly equipped	Understaffed	Developed	☐	☐
5	Fully equipped	Well staffed	Underdeveloped	☐	☐
6	Fully equipped	Well staffed	Developed	☐	☐
7	Fully equipped	Understaffed	Underdeveloped	☐	☐
8	Fully equipped	Understaffed	Developed	☐	☐

However, it is more common to offer respondents a choice between two job profiles. One approach is to take the set of profiles derived from the orthogonal design, then take one of these profiles and compare all the other profiles to this one constant comparator. So, for example, using the eight job profiles above, one job profile would be taken out and compared to all other job profiles, thus creating seven choices. An example of this approach is provided in Hanson and Jack (2010). They generated a main-effects fractional factorial design with 16 profiles (that satisfied orthogonality and level balance, as above). They then chose a "middling" job profile, and compared it to all the other profiles. Thus, each respondent was presented with 15 pairwise choices. Mangham and Hanson (2008) adopted a similar approach.

While this approach is known to maintain the properties of the fractional factorial design, alternative methods exist (and the constant comparator approach is generally not recommended). Louviere et al. (2000) propose "foldover" or "shifting" methods to do this.

Originally foldover referred to a mirror image of the original design, namely, recoding such that 0=1 and 1=0 for 2 levels; 0=3, 1=2, 2=1, 3=0 for 4 levels (Louviere et al. 2000). However, the usefulness of a foldover design depends on the number of options in the choice sets, the number of levels for the attributes, and how the foldover design is defined. For example, if there are 2 options in the choice sets (job A or job B) and the attributes have 4 levels then the foldover design 0=3, 1=2, 2=1, 3=0 will result in 0% efficiency (because levels 0 and 3 only appear together, as do levels 1 and 2). However, a foldover design defined as a systematic level change or cyclical shifting of the original design (0=1, 1=2, 2=3, and 3=0) will result in a design that has a higher efficiency (Street et al. 2005).

An example of a foldover design, for the full factorial presented in figure 1.5, is shown in figure 1.6. Job B is a mirror image of Job A.

Figure 1.6 Foldover design

	Job A			Job B			Do you prefer Job A	Do you prefer Job B
	Equipment	Staffing	Social amenities	Equipment	Staffing	Social amenities		
1	Poorly equipped (0)	Well staffed (1)	Under-developed (0)	Fully equipped (1)	Under-staffed (0)	Developed (1)	☐	☐
2	Poorly equipped (0)	Well staffed (1)	Developed (1)	Fully equipped (1)	Under-staffed (0)	Under-developed (0)	☐	☐
3	Poorly equipped (0)	Under-staffed (0)	Under-developed (0)	Fully equipped (1)	Well staffed (1)	Developed (1)	☐	☐
4	Poorly equipped (0)	Under-staffed (0)	Developed (1)	Fully equipped (1)	Well staffed (1)	Under-developed (0)	☐	☐
5	Fully equipped (1)	Well staffed (1)	Under-developed (0)	Poorly equipped (0)	Under-staffed (0)	Developed (1)	☐	☐
6	Fully equipped (1)	Well staffed (1)	Developed (1)	Poorly equipped (0)	Under-staffed (0)	Under-developed (0)	☐	☐
7	Fully equipped (1)	Under-staffed (0)	Under-developed (0)	Poorly equipped (0)	Well staffed (1)	Developed (1)	☐	☐
8	Fully equipped (1)	Under-staffed (0)	Developed (1)	Poorly Equipped (0)	Well staffed (1)	Under-developed (0)	☐	☐

What is minimum overlap?

When creating choice sets, an additional property is minimum overlap—within any choice the attribute level is not repeated across options. This ensures that no choice set has the same level of a given attribute, drawing out the maximum information from respondents. This is clearly adhered to when a foldover or shifting process is used (as in figure 1.6).

Statistically efficient designs

When using statistically efficient designs the orthogonality criteria may be sacrificed (though some level of orthogonality will be required, but not perfect, for the model to be estimated). Statistical efficiency has been defined in terms of D-efficiency, which can be interpreted as minimizing the determinant of the covariance matrix. This ensures minimum variation around the parameter estimates by minimizing the estimated standard errors. SAS software can be used to generate such designs (Kuhfeld 2010; http://support.sas.com/documentation/index.html). Blaauw et al. (2010) used this approach to derive the choices in their study that was looking at policy interventions to attract nurses to rural areas in Kenya, South Africa, and Thailand (SAS software has a very good user manual that can be downloaded). Ngene (http://www.choice-metrics.com/) has been developed to create D-efficient designs. Ngene allows the researcher to force the design to maintain orthogonality while optimizing efficiency.

D-efficient fractional factorial designs can also be produced using the Sawtooth software package (http://www.sawtoothsoftware.com/). While Sawtooth has been increasingly used in private industry applications, it has not been used widely in health policy applications. However, Kruk et al. (2010) and Rockers et al. (2012) have used Sawtooth to conduct human resource DCEs in Ghana and Uganda, respectively. Sawtooth can be used to produce efficient designs and to administer DCEs on a computer. Computer-based administration eases logistical constraints and simplifies data entry (see the Uganda case study for an application of DCE with Sawtooth).

Utility balance and prior information (Bayesian designs)
When deriving efficient designs researchers commonly assume that parameters are zero, that is, they have no a priori information about the parameters to be estimated. However, a recent development is to use prior assumptions about parameters to improve statistical efficiency (de Bekker-Grob et al. 2012). This helps ensure that the choices presented to individuals are close in terms of utility (known as utility balance), and therefore really force respondents to think about their choices (as opposed to providing choices involving options that are wide apart in terms of utility space). This does, though, create the potential danger that the job profiles become too similar, making the choices very cognitively challenging. Such priors, which may be obtained from pilot work, can be incorporated using software such as SAS (Kuhfeld 2010; http://www.sawtoothsoftware.com) and Ngene (http://www.choice-metrics.com/). This approach has not yet been used in the human resource literature.

Deriving an experimental design for a labeled DCE
When designing a DCE, practitioners should consider whether they will use a generic or labeled design (Figures 1.1 and 1.2 are both examples of generic choices—jobs defined as "Job A" or "Job B," or "Job 1" or "Job 2"—and the method for developing their design has been described above). Figure 1.3 provides an example of a labeled choice (jobs defined as "rural location" or "urban location"). Within the human resources literature only Blaauw et al. (2010) has used a labeled design, though such designs are popular in the transport field (Hensher et al. 2005).

This labeled choice approach is one way to allow for interaction terms between attributes, that is, better housing might be more highly valued in a rural than urban location. It also allows for different attribute levels across urban and rural locations. For example (from table 1.2), the rural posts offers additional money per year over the amount offered for the urban position (R120,000 per year). Similarly, while the rural position offers "basic" or "superior" housing, the urban position offers only basic (or none). The time to wait before getting study leave to specialize offers an improved situation of 2 years (compared to 6) for the rural position but not the urban position, and a car allowance is offered for the rural position but not the urban position. Other attribute levels are the same across urban and rural positions. (Both SAS and Ngene allow for constraints when designing choice sets, that is, one level cannot go with another.)

For this labeled design, this combination of attributes and levels resulted in an experimental design with 10 attributes at 2 levels and 1 attribute at 4 levels (table 1.5).

Table 1.5 Deriving a design for a labeled choice

Attribute	Level	Summary (Levels^{Attributes}, L^A)
Type of facility	Rural = 2 Urban = 2	2^2
Annual salary	Rural = 4 Urban = constant	4^1
Provision of subsidized housing	Rural = 2 Urban = 2	2^2
Time to wait before getting study leave to specialize	Rural = 2 Urban = constant	2^1
Car allowance	Rural = 2 Urban = constant	2^1
Number of years to be spent in the facility until being eligible for promotion	Rural = 2 Urban = 2	2^2
Workplace management and culture	Rural = 2 Urban = 2	2^2
Summary of design		$2^{10} \times 4^1$

When developing such a design the practitioner would then look for a fractional factorial design for 11 attributes, 10 at 2 levels and 1 at 4 levels. Each profile with the design would therefore contain the urban scenario and the rural scenario.

Sometimes, the order of attributes in the design may not run successively for the rural and urban job (or vice versa). For example, assuming 8 attributes (4 for rural and 4 for urban), attributes 1, 2, 4, and 5 may be for the rural post and 2, 6, 7, and 8 for the urban post.

Challenges and concerns for the experimental design: statistical and respondent efficiency

Experimental design methods aim to maximize the efficiency of the design, while giving no consideration to the realism of the job choices posed—in other words, the choices generated may be unrealistic. Here the researcher must balance statistical efficiency against respondent efficiency, and a pragmatic approach is recommended. There is no point having a very efficient design (statistically) if the job choices do not seem realistic to the respondents/health workers (thereby lacking respondent efficiency).

The SAS software package allows the researcher to define constraints within the design (levels of attributes that cannot go together). If catalogues or websites are used to derive choices then it would be possible to manually alter some choices to ensure respondent efficiency. If this is done it would be important to check the design properties (such as orthogonality, level balance, and minimum overlap) once manual change has been done. It is also recommended that response data are simulated and that the model to be run can be estimated.

Number of choices and cognitive fatigue
Even after using experimental design methods, a large number of choices may remain for presentation to respondents. This raises the question of the number of choices subjects can respond to, before becoming tired, bored, or unmotivated. The number of choice sets that respondents are presented with in DCEs in health has increased, with the mean number at 14 (de Bekker-Grob et al. 2012). Within the applications summarized in the appendix the number of choices range from 12 in the study by Kruk et al. (2010) to 18 in the study by Chomitz et al. (1998). It is often argued in health that respondents may not understand or may lack familiarity in the applications (where individuals are not used to making choices). This is less likely the case for job choices, where individuals think about their future and possibilities for a better life, and so will be more familiar with such decisions.

The practitioner should address the issue of the feasible number of choices in the pilot work. When the experimental design produces too many choices to present to one respondent, it is possible to block the design into smaller sets. A design with 32 choices may be blocked into two groups of 16 choices. This may be done randomly—Kolstad (2011) gives an example—or the software package SAS generates blocks that still satisfy efficient design criteria.

To have or not to have an opt-out option
Once choices have been derived from the experimental design it is important to consider whether to have a forced choice or add an opt-out option. A forced choice, as the name suggests, forces individuals to choose one of the jobs on offer. (Figures 1.1. and 1.3 are examples of forced choices.) An opt-out gives respondents the option of not choosing any of the jobs on offer. (An example of this is shown in figure 1.2.) If an opt-out is included it is important to be very clear to respondents what this means. Where health workers are being surveyed, the opt-out would be to continue doing what they are doing; hence information on what they are doing must be collected (See Vujicic et al. 2010a for an example).

The inappropriate use of forced choices may result in biases with respect to parameter estimates. That is, individuals may be forced to take up a job when in reality they would choose not to. There are however a number of potential disadvantages to incorporating non-forced choices into a DCE. Respondents may select such an alternative not because it provides the highest benefit (utility) among the alternatives but to avoid making a difficult decision. Additionally, allowing respondents to select an opt-out option provides less information on respondents' relative preferences for the attributes in the hypothetical alternatives.

Given that the objective of many studies is to look at WTP estimates and probability of take-up, the practitioner is encouraged to consider the role of opt-out options in their DCE. When constructing choices the researcher simply adds an opt-out/current situation option to the choice set derived from the experimental design.

Designs are readily available (Street et al. 2008) and design experts will quite often provide advice and guidance, and sometimes even create the design for practitioners.

A final point is that it is good practice to indicate the experimental design method used in developing the choice set when reporting on the DCE (de Bekker-Grob et al. 2012).

1.2.3 Development of the questionnaire, pretesting, and data collection

SUMMARY OF 1.2.3

When researchers develop a DCE questionnaire (or survey), they should consider the following:

1. Rationality/internal consistency checks should be included in the final DCE instrument. These checks allow the researcher to ensure that respondents were engaging in the exercise and taking it seriously. The information collected for these checks should not be included in the final econometric analysis.
2. If the researcher wants to investigate how preferences differ according to the characteristics of respondents then information must be collected on such characteristics, for example, sex, age, location, indicators of socioeconomic status, previous work experiences, and attitude toward working in rural areas.
3. An introduction to the questionnaire is required, indicating the goals of the study, why the respondent has been chosen, who is carrying out the survey, and how the results will be used.
4. Sample size must be determined. Previous experience suggests that, for each predetermined subgroup of the main sample (such as sex and remote or rural background), a minimum sample size of 30 is required (based on econometric criteria).
5. How to collect the data? A number of data collection methods exist, including self-administered questionnaires (on paper or computer), completion in a classroom or examination setting, and trained fieldworkers interviewing respondents individually. The research team will have to make a judgment on this.
6. Pilot testing should be conducted to determine whether respondents understand the definitions of attributes and levels; whether they can cope with the number of attributes and number of choices; and whether they understand the choices. The pilot should be of a sufficient sample size to do econometric analysis on the data to test that coefficients are moving in the expected direction.
7. Translation of questionnaires to different languages, ensuring the true meaning of the attributes and associated levels has not been lost in translation.

Formulating the questionnaire (or survey)

An introduction to the questionnaire will be required, indicating the subject of the study, why the respondent has been chosen, who is carrying out the survey, and how the results will be used. The challenge for the researcher is to ensure the questionnaire is "incentive compatible," that is, that the respondent engages and responds honestly.

The choices generated from the experimental design form the basis of the DCE questionnaire (Mangham et al. 2009). The following may also be included:

- warm-up choices—to familiarize respondents with the question design; and
- rationality/internal consistency checks to allow the researcher to ensure that respondents were engaging in the exercise and taking it seriously.

Information collected from warm-up choices and internal consistency checks should not be included in the final econometric analysis (if they were added to the choices derived from the experimental design).

The most common internal consistency test is **nonsatiation** or **dominance**. Here choices are included where one option clearly dominates (or is superior). If no a priori assumption can be made re the preference ordering for an attribute, these should be constant across options within a choice set. Figure 1.7 shows an example of a dominance test for the attributes and levels provided in Kolstad (2011). Respondents would be expected to choose Job A over Job B. Where opt-out options are included, such a test will not be so suitable and the researcher should consider transitivity tests or Sen's Expansion and Contraction properties (See de Bekker-Grob et al. 2012 for more on these tests).

Figure 1.7 Example of a consistency test

Job A

Availability of equipment and drugs	Housing	Education opportunities/ possibility of upgrading qualifications	Workload	Infra-structure	Salary and allowances	Location
Sufficient	Decent house is provided	Education offered after 2 years of service	Normal: Nearly enough time to complete duties. One hour of extra work per day	The place has mobile coverage, electricity and water	T Sh 350,000 per month	Regional headquarters

Job B

Availability of equipment and drugs	Housing	Education opportunities/ possibility of upgrading qualifications	Workload	Infra-structure	Salary and allowances	Location
Insufficient	No house is provided	Education offered after 2 years of service	Heavy: Barely enough time to complete duties. Three hours of extra work per day	The place has unreliable mobile coverage, no electricity or water	T Sh 200,000 per month	A 3-hour or more bus ride from the district headquarters

Considering your current situation, which of the two jobs would you choose?

Job A: ☐ **Job B:** ☐

If respondents "fail" such tests, the question is what to do with them? Lancsar and Louviere (2006) argue that deletion of such respondents may be inappropriate since such responses may be valid. Supporting this concern, San Miguel et al. (2005) and Ryan et al. (2009), both using qualitative research techniques, and following up respondents after they responded to the choices (through either giving respondents an extra question to get feedback on the choices (San Miguel et al. 2005) or through interviews with respondents (Ryan et al. 2009), found that individuals who had been defined as failing rationality tests had "rational" reasons for doing so. Lancsar and Louviere (2006) also note that random utility models (see section 1.2.5 on random utility theory) are robust to errors made by individual in forming and revealing their preferences. The practitioner should investigate reasons for "failing" rationality tests in qualitative pilot work as this may provide some insight into the limitations of the questionnaire.

Warm-up choices and consistency checks, which are additional to the experimental design, should be dropped from the econometric analysis (since their inclusion will reduce the statistical efficiency of the data matrix).

How to present the questions

Once the choices have been decided (from the experimental design, warm-up questions, and rationality/internal consistency checks), consideration should be given regarding their presentation. Mangham et al. (2009) suggest that pictures are useful to explain attributes in a low- or middle-income country context where literacy cannot be assumed. However, given that DCEs concerned with human resource issues usually target students and health workers, this may not be such a problem. Still, visual elements may still help by reducing potential boredom and helping respondents engage.

What additional information can or should be collected during the survey

Having finalized choices, and how to present them, practitioners should consider the additional information to be collected in the questionnaire. Such information may include information on the socioeconomic characteristics of respondents (sex, age, location, indicators of socioeconomic status) as well as motivation-type questions (so as to gather insight into the intrinsic motivations of individuals).

Gathering additional information is useful for a number of reasons, including testing the representativeness of the respondents to the DCE, as well as conducting subgroup analysis to see how preferences differ by such factors. For example, in addition to looking at preferences for the whole sample, Kolstad (2011) investigated preferences by sex; rural, remote background; and "willingness to help people".

She found that women are less responsive to monetary incentives and are more concerned with factors that directly allow them to do a good job, while those with parents living in a remote rural area are generally less responsive to the proposed policies. She also found that when the willingness to help other people is a strong motivating force, policies that improve the conditions for helping people appear particularly effective. These are useful findings for policy.

Target sample

Who receives the questionnaire will depend on the group of potential or actual health workers in which the researcher is interested. For example, Chomitz et al. (1998) targeted final-year medical students (who would be shortly be of choosing locations for their compulsory medical service). Hanson and Jack (2010) targeted registered nurses, and Kolstad (2011) targeted clinical officers. Rockers et al. (2012) targeted students in the final year of their training program as well as in-service health workers from medical, nursing, pharmacy, and laboratory cadres.

The DCE objectives should be relevant to the health workers targeted. For example, it is likely that health workers near retirement or married would make a different choice on remote postings than younger or unmarried workers. This highlights the importance of having a representative sample and of collecting background information on the respondent. The sample may be purposely selected according to certain characteristics.

Sample size

Given a defined target sample, sample size must be determined. This is a very important issue because samples that are "too large" may waste time, resources, and money, while samples that are "too small" (less than 30) may lead to inaccurate results (imprecise estimates).

Various questions need to be answered before a suitable sample size can be determined. The first refers to the level of accuracy (precision) required. In general, the higher the level of accuracy required, the larger the sample size should be. Sometimes the sample size required is so close to the entire survey population that it makes more sense to simply survey everyone. More often "smart" designs are used to reduce the required sample size without reducing the accuracy.

A second issue is whether estimates for subgroups, as well as for the overall population, are required. The overall sample size needs to be large enough to ensure that an adequate level of accuracy for these subgroups can also be achieved.

Another important question affecting the sample size required is the level of variability between responses. Usually, the less variable the responses, the smaller the sample size required to achieve the same level of accuracy.
Finally, the burden placed on respondents needs to be evaluated. If people are surveyed too frequently, they are less likely to take the survey seriously, so the sample size should not be larger than necessary to obtain the accuracy needed.

Determining the "correct sample size" is not therefore a simple task. In fact, a large part of determining the sample size is not just "how many should be sampled", but how cleverly the sample is chosen. A smarter sample design can give more accurate estimates with a smaller sample size.

Sample-size calculations are available in most survey-sampling textbooks (Cochran 1977, for example). Louviere et al. (2000, p. 262) provide a formula to calculate the minimum sample size needed to measure choice probabilities (or proportions) with some desire level of accuracy using a random sample. The book by Ben-Akiva and Lerman (1985) includes a full chapter on sampling theory (Chapter 8, pp. 217–252). Hensher et al. (2005, pp. 193–196) also provide an overview of the reality (as opposed to the theory) of sampling practices within studies of choice.

Ultimately the selection of sample strategy and size largely depends on the budget and resources available. However, using econometric criteria, subgroups of smaller than 30 individuals would be too small to conduct meaningful statistical analysis (The appendix shows the sample size for studies applying DCEs to human resources for health issues). In a low- or middle-income country context it is often difficult to use random sampling to obtain sufficient sample sizes. Some studies combine different worker categories and others use stratified or cluster sampling strategies.

An important point is that while a DCE produces multiple responses from individuals, this does not imply that very small samples can be used. For example, 10 individuals responding to 16 choices would produce 160 observations. This is not the same as 160 independent observations (since the 160 observations are only from 10 people, and therefore unlikely to be representative of the target population). Regression techniques allow for the fact that individuals provide multiple observations (section 1.2.4 discusses this further).

How to collect the data
Consideration must then be given to how to collect the data. Several data collection methods exist, including self-administered questionnaires, completion in a classroom or examination setting (Chomitz et al. 1998; Blaauw et al. 2010; Kolstad 2011) and trained fieldworkers interviewing respondents individually (Mangham and Hanson 2008; Hanson and Jack 2010). Further, DCE instruments can be administered either on paper or with computer-based administration software (Kruk et al. 2010; Rockers et al., 2012). Respondents must know that for self-administered questionnaires they should not discuss questionnaire responses with each other.

Translation and piloting
In low- and middle-income countries, where different languages may be spoken, translation of the questionnaire may be required and issues in relation to the use of languages should be considered (Mangham et al. 2009).

Having developed the questionnaire, identified the target sample, and method for administration, it is important to pilot the questionnaire on the target population. Interviews and focus group discussions are an important part of the piloting. Important questions include:

- Do respondents understand the definitions of attributes and levels?
- Can they cope with the number of attributes and number of choices (and are they not adopting simple decision-making heuristics to respond to the questionnaire)?
- Do respondents understand the task, as indicated by responses to rationality/internal consistency checks and internal validity? This may be assessed by analyzing the pilot response data

and checking that attribute parameters are moving in the expected direction (section 1.2.5). It is thus important to have a sufficient sample size to do meaningful econometric analysis after the pilot (at least at the aggregate level). Hence pilot work will require a minimum of 30 individual responses.

1.2.4 Data input

SUMMARY OF 1.2.4

DCE data must be converted to a specific format for statistical analysis. Each respondent will have several rows of data in the final DCE dataset, depending on how the choices were presented (such as binary choice, forced choice, or multiple choice, including an opt-out).

- If an individual was presented with 8 binary choices (will you take up the job, yes or no), that individual will have 8 rows of data in the final dataset.
- If an individual was presented with 8 forced choices, each with 2 job scenarios, that individual will have 16 (8x2) rows of data in the final dataset.
- If an individual was presented with 8 choice sets, each with 2 job scenarios and an opt-out (3 options per choice set), that individual will have 24 (8x3) rows of data in the final dataset.

Practitioners must be familiar with both the software package they are using and the associated regression commands to know all necessary columns for their data matrix.

Once the data have been collected they must be entered into a computer. Data entry programs include CSPro (http://legacy.measuredhs.com/cspro/), EpiInfo (http://wwwn.cdc.gov/epiinfo/), Excel (http://office.microsoft.com/en-gb/excel/), and SPSS (http://www-01.ibm.com/software/uk/analytics/spss/). Double entry is advisable. At a minimum, a random sample of the data should be checked by a second person. Given the nature of the DCE data—each individual is provided with a number of choices and each choice involves a number of options—in the final data matrix each individual provides multiple rows of data.

For example, for a binary choice (as in figure 1.4), the final data matrix is as follows (table 1.6).

Table 1.6 Example of final data matrix for 1 individual—binary choice

personid	choiceset	equipment	staff	social amenities	choice	age	sex
1	1	0 (poor)	1 (well)	0 (under-developed)	0	25	1
1	2	0 (poor)	1 (well)	1 (developed)	0	25	1
1	3	0 (poor)	0 (under)	0 (under-developed)	1	25	1
1	4	0 (poor)	0 (under)	1 (developed)	0	25	1
1	5	1 (fully)	1 (well)	0 (developed)	1	25	1
1	6	1 (fully)	1 (well)	1 (under-developed)	0	25	1
1	7	1 (fully)	0 (under)	0 (developed)	1	25	1
1	8	1 (fully)	0 (under)	1 (under-developed)	1	25	1

personid: in this design the individual is presented with 8 binary choices, and hence there are 8 rows of data for the one individual who is identified by this variable.

choiceset: indicates which choice is presented (there are 8 choices).

equipment, staff, and social amenities: are the levels of the attributes within the choices. Each attribute has only two levels—0 or 1. For example, choiceset1 has poor equipment, is well staffed, and has underdeveloped social amenities. By contrast, choiceset2 has poor equipment, is well staffed, and has developed social amenities.

choice: indicates the respondent's choice, that is, whether the respondent chose the job (1) or not (0). So, for choiceset = 1 and 2 , the respondent chose not to take up the job (0) whereas for choiceset 3 they chose the job (1).

age and sex: given that all 8 rows show responses for the same individual, the personal characteristic information (such as age and sex are repeated for that person.

When individuals are presented with 2 or more options, possibly including an opt out, the data matrix becomes even larger.

- If an individual is asked 8 choices, and each choice involves 2 options (Job A and Job B), that individual will provide 16 rows of data (8 choices x 2 options). This means that information on the characteristics of respondents (such a sex r, age) is given 16 times in the data matrix for that individual. (For an example of such a data matrix see the Tanzania case study, table 2.3, for 16 choices, and the Uganda case study, table 3.2, for 12 choices.) When an individual is presented with a forced choice (2 options) it is also possible to define the attributes as differences, with each choice then taking up one row of data (as opposed to having the data stacked).
- If an individual is asked 8 choices, and each choice involves 3 options (Job A, Job B, and neither), that individual will provide 24 rows of data (8 choices x 3 options). Information on the characteristics of respondents (such as sex, age) is given 24 times in the data matrix for that individual.

The data may be input in a wide rather than a long format. For example, for the individual responses shown in table 1.6, the initial data matrix may be as follows:

Table 1.7 Wide-format initial data matrix

personid	choice1	choice2	choice3	choice4	choice5	choice6	choice7	choice8	age	sex
1	0	0	1	0	1	0	1	1	25	1

Where personid, age, and sex are as defined above and choice1–choice8 represent the responses to the 8 choice questions, the row of data for personid=1 above corresponds to the 6th column in table 1.6. It is possible to go between wide and long data within the Stata program using the reshape command:

Reshape long Choice, i(personid), j(choiceset)

where

long tells reshape to go from *wide* to *long* (it is also possible to go from long to wide)

Choice tells reshape that the *stem* of the variable to be converted from *wide* to *long* is *Choice* *i(personid)* tells reshape that personid is the unique identifier for records in their *wide* format

j(choiceset) tells reshape that the suffix of *Choice* (i.e., 1, 2, 3, 4, 5, 6, 7, 8) should be placed with a variable called choiceset.

Table 1.7 then becomes:

personid	choiceset	choice	age	sex
1	1	0	25	1
1	2	0	25	1
1	3	1	25	1
1	4	0	25	1
1	5	1	25	1
1	6	0	25	1
1	7	1	25	1
1	8	1	25	1

Using the reshape command, variables that do not vary by personid, that is, sex and age, will be input automatically.

To insert the attribute levels for each choice into the data matrix it would be possible to include syntax in the reshape command (if the variables were included in the initial data matrix) or use the generate command.

As other programs will allow the data matrix to be converted, and as different packages will require different columns of data for different regression commands, the practitioner needs to be familiar with the software package (and how the final data matrix should look) and associated regression commands.

Continuous and categorical attributes
The researcher will need to define how the attributes will be modeled when inputting the data. Continuous attributes are usually modeled as continuous variables, with the actual values input. The estimated parameters are then interpreted as the value of a unit change in that continuous variable. The unit will depend how the variable is measured; for salary it would be a unit of currency, for a time variable it would be a unit of time (year).

Categorical variables are commonly modeled as ordinal or dummy variables. For dummy variables preferences are modeled relative to some base case (coded as 0). Thus the interpretation of the estimated parameter is the marginal value of a movement from the base case to a defined level. An application of this can be seen in the study by Kruk et al. (2010), where all attributes other than salary were defined as dummies, the parameter "children's education" shows the value of moving from "no allowance for children's education" to "allowance for children's education". Similarly, the parameter "infrastructure, equipment, supplies" shows the value of moving from "basic (unreliable electricity, X-ray, intermittent drug supply)" to "advanced (reliable electricity, ultrasounds, constant drug supply)". The case studies on Tanzania and Uganda further discuss data input for dummy variables.

1.2.5 Data analysis and interpretation

SUMMARY OF 1.2.5

Statistical analysis of DCE data is based on the random utility model. Several models can be used to estimate respondent preferences for job attributes included in the DCE, including random effects binary probit and logit, conditional logit, and mixed logit.

Whatever method of analysis is used the results of DCE analyses can be used to determine:

- which attributes are important and how important one attribute is in comparison to another attribute;
- how individuals trade between attributes of a job (how much of one attribute they are willing to give up for improvements in another);
- how much salary an individual would be willing to give up for improvements in others attributes of a job; and
- the probability of individuals taking up a job with specified attributes.

DCEs thus provide additional information to qualitative data, providing quantitative information on strength of preference, monetary values and predictions of the likely take-up of defined jobs.

This type of information cannot be obtained from detailed focus groups or interviews, nor from existing longitudinal data when new policies are being implemented (since such data will not exist).

The random utility model provides the theoretical underpinning for analysis of the DCE data (so the practitioner should be familiar with this). In this framework individual *n* is assumed to choose between *J* alternative jobs, opting for the one associated with the highest utility (benefit or satisfaction). Thus, individual *n* will choose job *i* over *j* if and only if

$$U_{ni} > U_{nj} \qquad \forall i \neq j \in$$

where *U* is the utility for a given job.

The random utility model assumes that the utility (U) associated with a particular job is made up of two components. The deterministic component V_{ni} is a function of *m* job attributes ($x_1, \ldots, x_m$), which are observed and the random component, ε_{ni} , which is a function of unobserved job attributes and individual-level variation in tastes. The utility, U, to individual n associated with job *i* can be specified as:

$$U_n = V_n + \varepsilon_n = \alpha_1 + \beta_1 x_{1n} + \beta_2 x_{2n} + \ldots + \beta_m x_{mni} + \varepsilon_n \qquad (1)$$

where the betas, ß, provide quantitative information on the strength of preference for each attribute level, as well as trade-offs, monetary values, and predicted take-up of posts (see below).

However, the utility of any given job is not directly observable, and therefore the coefficients in equation (1) cannot be estimated directly. The DCE data are therefore modeled within a probabilistic framework. That is, when individual *n* is presented with a pair of jobs, the probability (P) individual n chooses job *i* over job *j* can be estimated as

$$P_{ni} = \Pr[U_{ni} > U_{nj}] \qquad \forall i \neq j \in$$

Using equation (1) this becomes

$$P_n = \Pr \ \varepsilon_n - \varepsilon_n > V_n - V_n \ ] \qquad (2)$$

To estimate equation (2) an assumption has to be made about the distribution of the error term ε_{ni} —a probit approach assumes a normal distribution (Chomitz et al. 1998; Hanson and Jack 2010; Mangham and Hanson 2008), and a logit model a logistic distribution (Kolstad 2011). Given the flexibility of the logit approach, it has been the preferred approach in the DCE literature (de Bekker-Grob et al. 2012).

Using the logit model, the probability of choosing job *i* is defined as:

$$P_i = \frac{\exp(V_i)}{\sum_{j=1}^{N} \exp(V_j)} \qquad (3)$$

For an application of this model see the Tanzania case study. Readily available software such as Stata (http://www.stata.com/), Limdep/nlogit (http://www.limdep.com/), SAS (http://www.sas.com/offices/europe/uk/), and Sawtooth (http://www.sawtoothsoftware.com/) can be used to estimate such models. Matlab (http://www.mathworks.co.uk/) and R (http://www.r-project.org/—the only free program for analysis, as well as design) are also available but require the researcher to do their own programming. The researcher analyzing the data should look for data analysis coding using terms such as logit, probit, and mixlogit. Software packages differ in the exact terms they use.

The coefficients (ßs) generated from the logit (or probit) model (equation 1) can be used for two main purposes:

- To determine whether the attributes are important (statistically significant, as shown by the significance level of the ß), the direction of importance (shown by the sign of the estimated ß) and relative importance (size of the estimated parameter). Lagarde and Blaauw (2009) note that studies applying DCEs to inform human resource policy interventions have shown that nonmonetary incentives are significant determinants of job choice, sometimes more so than financial ones.
- The direction of the coefficient signs also provides a check on the theoretical/internal validity of the DCE model—that is, whether the coefficients move as economic theory or a priori expectation would predict. For example, economic theory would predict the salary attribute to have a positive sign—that is, the higher salary, the more desirable the post.

Although the above information is very useful, the real value of DCEs is in using them to look at two things: the trade-offs that respondents are willing to make among attributes; and the probability of take-up of defined posts. This type of information cannot be obtained from detailed focus groups or interviews, ranking or rating exercises, nor existing longitudinal data when new policies are being implemented (since such data will not exist). Looking at each of these in turn:

- **trade-offs among attributes** can be estimated as long as a continuous variable is included. If this continuous variable is salary, the monetary value for other attributes can be estimated. For example, the ratio of any given coefficient divided by the negative of the price proxy (salary in this application) can be used to estimate health workers' WTP for various job attributes, for example, how much salary they are willing to give up for better working conditions. For example, Mangham and Hanson (2008) were interested in the extent to which nurses (in Malawi) were willing to trade between monetary and nonmonetary benefits, which the literature has referred to as the "compensating differential" (Chomitz et al. 1998), namely, the amount of money that is equivalent to having better working conditions. It is a commonly used output in the literature. Since WTP is derived as the ratio of two random variables, WTP is itself a random variable. Confidence intervals should therefore be estimated for WTP estimates (discussed further in the case studies).
- **the probability of individuals taking up a job** with specified attributes can be estimated, using equation 3 above. These predictions are very useful to policy makers as they show

the predicted impact on health worker decisions of alternative levels of job attributes, that is, alternative jobs offered. For example, Hanson and Jack (2010) found the following: that doubling wages in areas outside the capital would increase the share of doctors willing to work there from about 7% to 50%; providing high-quality housing would increase physician supply to about 27% (equivalent to paying a wage bonus of about 46%); doubling wages to nurses for working in rural areas outside cities would increase their labour supply from 4% to 27%; and that the most effective nonwage attribute in inducing them to relocate to rural areas is the quality of equipment and drugs. As with WTP estimates, it is useful to estimate confidence intervals for the probability estimates.

All the above information can be estimated for the total sample, or for subgroups of the population. Subgroup analysis is commonly carried out in analysis of DCE data. For example, Kolstad (2011) found that women were less responsive to monetary incentives and more concerned with factors directly allowing them to do a good job, while those with parents living in a remote rural area were generally less responsive to the proposed policies.

The mixed logit (MXL) model has been developed to allow for unobserved heterogeneity of preferences (sub-samples of the population that do not have to be identified by the researcher, the data will identify them). This modeling approach also allows for multiple observations being obtained from individuals and violation of the independence of irrelevant alternatives assumption of the conditional logit model (the assumption that the introduction or removal of a choice has no effect on the proportion of probability assigned to each of the other choices).

When adopting the MXL model selected parameters (or job attributes) are permitted to vary according to defined statistical distributions. Thus, preference heterogeneity in the sample is incorporated into the model by treating the coefficients as random rather than fixed (Kruk et al. 2010; Blaauw et al. 2010; Vujicic et al. 2010a; Rockers et al., 2012). While the econometric framework accommodates a number of parametric distributions for the coefficients (de Bekker-Grob et al. 2012), studies tend to assume normality (Kruk et al. 2010; Blaauw et al. 2010; Vujicic et al. 2010a) and the coefficient on the money attributes is assumed fixed. For an application of this model, and issues raised when estimating WTP using the MXL model, see the Uganda case study and Hole and Kolstad (2010).

The normality assumption implies there will be both positive and negative values across the population being sampled for a given attribute. For example, for a housing attribute taking on the values "poor" and "good", the normal assumption implies a proportion of the population prefer "good" housing and a proportion "poor" housing. It is often the case that this assumption may not be realistic, and it would be more intuitive to assume a log-normal distribution (where the attribute sign would always be positive). For example, for housing, all respondents prefer "good" housing. The assumption of normality is often made for ease of estimation rather than realism i.e. the model will not converge when different assumptions are made across random coefficients. Larger sample sizes and developments in computer software may help but the practitioner should be aware of the intuition behind the assumptions made regarding random coefficients that are normally distributed.

While the MXL model also allows for heterogeneity of preferences, there is a question of how useful such information is to policy makers since it is not possible to identify where preferences differ. A more useful approach may be to use conditional logit to gain better insight into observed variation which can potentially be acted upon by policy makers. The practitioner should be aware of the advantages and limitations of the approach adopted.

1.3 Logistical issues in conducting DCEs

This section looks at the logistical issues facing those conducting DCEs. The information comes from a small survey of researchers who have conducted DCEs for rural retention and recruitment in recent years. The survey, conducted to feed into this User Guide, was carried out by the World Health

Organization/Human Resources in Health, in April–May 2011. It adopted a question and answer format, and was administered via email.

The survey was answered by Margaret Kruk of Columbia University, United States; Nonglak Pagaiya of the International Health Policy Program, Thailand; Peter Rockers, Harvard University and Capacity*Plus*; Julie R. Kolstad, University of Bergen, Norway; and Lindsay Mangham, London School of Hygiene and Tropical Medicine, United Kingdom. Responses to the six questions included in the survey are summarized below.

1.3.1 Why conduct a DCE?

In most situations, DCE arises as a pure research interest, which can be partly explained by the fact that it was in fact the research community who made this method advance. However, more and more policy makers have become interested in the results of DCE studies, and in some situations it was the Ministry of Health who requested a DCE in order to inform broader human resources for health policy discussions.

This guide recommends that before starting a DCE, substantial discussions need to take place in order to make full use of the results in the policy decision process. The method is quite expensive and requires significant investment in time and other resources; therefore, it is critical that its results are effectively used for policy implementation.

1.3.2 How long does it take?

Conducting a DCE in low- and middle-income countries can take on average 8–12 months, ranging from 3 months to 1.5 years in case of a multi-country study. The various phases of conducting this work are:

- a planning phase: 1–2 months
- qualitative work: 1–3 months
- design of questionnaires, including piloting and testing: 2–3 months
- survey administration: 1–2 months
- data entry and analysis: 1–3 months
- report writing: 1–2 months.

1.3.3 How much does it cost?

It is not always easy to get accurate information on costs, as often complex research programs are funded from various sources. A rough assessment of costs for conducting DCEs in five low- and middle-income countries (Ghana, Thailand, Uganda, Tanzania and Malawi) in 2009–2011 showed that total costs varied from $20,000 to $150,000, including costs of international consultants, local research teams, in-country travel, and other elements (such as laptops).

1.3.4 What skills and competencies does the research team need?

The research team is usually composed of both international and national researchers, and requires multidisciplinary skills. This includes expertise in qualitative research for developing attributes; knowledge of experimental design methods for informing the selection of choices to be presented to health workers or students; and advanced analytical and statistical skills, in particular logistic regression, for analysis of the data. Other competencies include broad human resources for health policy, health economics, social sciences, and program management.

Many low- and middle-income countries will need much capacity building and support to conduct a DCE well.

1.3.5 What logistical challenges are likely?

Weather often delays or impedes the execution of a DCE, so this has to be taken into account in planning. Local transportation, infrastructure, and availability of cars, along with security issues, need to be considered. Absence of health workers or important shortages can impede finding enough health workers for an adequate sample size. Access to heads of local institutions is sometimes difficult, so good discussions for the introduction of the study to policy makers are needed. A significant challenge has been the uptake of analytical skills by local researchers, despite extended training courses.

1.3.6 How can policy makers use the results?

The results of a DCE can be useful for policy makers who want to understand better the relative importance of different job characteristics on choices of rural work. The DCE should not be conducted isolated from the policy debate, but should inform discussions around rural recruitment and retention issues. Once the national human resources for health policy and plan have been discussed by policy makers, adequate policies to respond to shortages, especially in remote and rural areas, can be designed, based on quantitative information coming from the DCE.

Such information as the stated preference of health workers for rural jobs, or the estimated proportion of health workers who will choose a rural job if certain aspects of the job will be changed, are of critical importance in designing appropriate recruitment and retention strategies. Once the results of a DCE are made available, real-life experiments can be set up to implement the strategies suggested by the DCE results, then monitor their implementation, and eventually evaluate their impact.

1.4 Conclusions

DCEs have been applied in the developed and developing world to assess how job attributes influence job choice. In a DCE respondents are presented with a number of hypothetical job choices which vary with respect to attributes and levels. The responses can be used to obtain the following information that is useful at the policy level:

- Which job attributes are important and how important one attribute is in comparison to another attribute.
- How much salary a health worker would be willing to give up for improvements in others attributes of a job.
- The probability of respondents taking up a job with specified attributes.

The DCE practitioner should be aware that DCEs:

- Rely on responses to hypothetical choices—it is therefore important to follow up implementation of policies to ensure the validity of responses
- Require skills in the area of qualitative work, experimental design and econometric analysis
- Restrict the number of attributes to prevent choices becoming too cognitively challenging
- Require the data to be linked up with costs of attributes, and this may be challenging in some contexts.

Despite these challenges, DCEs are very useful at the policy level. They offer the policy maker a practical instrument, alongside detailed qualitative work and investigation of secondary datasets, to better understand job choices, and ultimately develop policies that attract and retain health workers in underserved areas.

2. How to make rural jobs more attractive to health workers: a DCE case study from Tanzania

SUMMARY OF SECTION 2

This section provides a case study of a DCE application in Tanzania. It explored preferences of clinical officer students.

As with the Uganda case study (see section 3), it highlights the importance of the qualitative component of the DCE methodology. In particular, it details the steps to identify job-posting attributes and levels that are important to health workers in the local setting. It also provides examples of how demographic questions can be used to supplement DCE information.

This case study uses in-depth interviews to gather the necessary qualitative information for DCE. It uses personal programming in Stata (Mata) to generate a D-efficient design, and employs the conditional logit model for the econometric analysis of the data. It uses the regression equation estimate to look at the relative importance of job attributes, WTP, and uptake rates.

It details the technical aspects of conducting the DCE (touching on logistical matters), and sheds light on the challenges of collecting data. Finally, it describes the steps taken to analyze data, interpret results, and demonstrate the value of a DCE over qualitative data.

This section explores the stages of a DCE, using the points highlighted in the step-by-step guide. A case study eliciting preferences of clinical officer students for rural jobs in Tanzania is used to illustrate the points (Kolstad 2011). The study was conducted in autumn 2007.

First it presents the background to the study, then examines the stages of a DCE: identification of attributes and levels; experimental design and constructing choice sets, with the properties of the design assessed; questionnaire development; data input, with consideration to the formation of the data matrix; and analysis and interpretation of data, with a focus on the policy relevant uses of a DCE, such as what attributes are important, the monetary value of attributes (willingness to give up salary for an improvement in other attributes of a job), and the probability of take-up of defined jobs. The focus is "hands-on" advice—common in all DCEs.

2.1 Background

As in many low- and middle-income countries, the geographic imbalance of the health workforce in Tanzania is a serious problem for delivering crucial health services to a large share of the population. Clinical officers form the group of clinicians more likely to work in rural areas, but even for this group the distribution is very much in favor of urban areas. Consistent with studies applying DCEs to address workforce issues in low- and middle-income settings, the aim of this DCE was to:

- Examine the importance of different attributes when clinical officers make job choices
- Establish the trade-off between these attributes, that is, how much salary would a respondent be willing to give up for improvements in other aspects of the job?
- Investigate the probability of job take-up as attribute levels change.

The study thus aimed to provide valuable information for policy makers considering different incentive packages to recruit health workers to rural areas.

2.2 Identification of attributes and levels

DCEs are an attribute-based measure of value, as said. Thus, the first stage is to define the attributes and levels. The study began with a long list of possible attributes, based on available empirical literature on job choices, job satisfaction and location decisions, as well as other factors that economic theory predicts will be important for the choice of job and job location. The list is featured in box 2.1.

Box 2.1 Initial list of potential attributes

- Salary
- Location
- Housing
- Equipment and drugs at the facility
- Access to further education
- Promotion possibilities
- Extra income generating opportunities
- Proximity to family and friends
- Professional environment, defined as the number of colleagues of the same or higher level of education
- Workload
- Facility ownership (public, private for-profit, NGO, faith-based etc.)
- Recognition from supervisor or boss.

Having established the initial list of potential attributes, researchers carried out in-depth interviews to further investigate these attributes and their respective levels to include in the DCE. They interviewed clinical officer students to identify additional attributes specific to Tanzania and this cadre. As a result, infrastructure and social/cultural opportunities were added to the list of potential attributes.

As indicated in section 1, the theoretical background to a DCE assumes that when an individual completes a DCE, for each choice he or she considers all the attributes and levels, and makes trade-offs. It is therefore important not to include too many attributes in the final DCE, or individuals may resort to simple decision-making strategies (such as always choosing the post with the highest salary). The interviews also provided information on the relative importance of attributes indicating how to reduce the number to a manageable level. In addition the interviews were used to discuss levels for the attributes. Details of the interviews are given in box 2.2.

Seven attributes were identified as both important to interviewees and policy relevant (table 2.1). Evidence suggests this is a manageable number, though at the higher end (de Bekker-Grob et al. 2012). It is also consistent with applications of DCEs in the workforce area.

Box 2.2 Summary of interviews to derive attributes and levels

Sixteen semi-structured in-depth interviews were conducted at Kibaha and Sengerema Clinical Officer Training Centres in March 2007.

The schools were chosen for their location; they were in different parts of the country, one rural in the north and one more central relatively close to Dar es Salaam.

The aim was to interview different types of students with different experiences and preferences at the two locations. However, it became clear that the official policy was to randomly assign students to a school, so in principle there should be no systematic differences between the students at these two schools. This seemed to be the case, although at both schools a slightly higher presence of students from the nearby areas than from the rest of the country was observed. Students were recruited from both sexes with both rural and urban backgrounds given the evidence that these groups have different preferences.

The interviews were conducted in Swahili and English, depending on the respondents' preferences for language. Two research assistants fluent in both English and Swahili conducted the interviews, with the principal investigator as an observer. The interviews were semi-structured, following an interview guide. The guide started with broad questions about where the respondent planned to seek his or her first job, why, their thoughts about working in a rural area, and whether they had any advice to give to policy makers who wanted to make rural jobs more attractive.

The initiative was initially with the respondents, but after a while topics that had not been brought up were drawn to the attention of the respondents. In the last part of the interviews, possible attribute levels were explored by presentation of hypothetical examples and discussion of their relevance. After each interview, the guide was modified to ensure that new insights could be explored.

Table 2.1 Attributes and levels in Tanzania study

Attribute	Salary and allowances	Education opportunities/ possibility of upgrading qualifications	Location	Availability of equipment and drugs	Workload	Housing	Infrastructure
Level 1	T Sh 650,000 per month	Education offered after 2 years of service	Dar es Salaam	Sufficient	Normal: Nearly enough time to complete duties. One hour of extra work per day.	Decent house is provided.	The place has mobile coverage, electricity and water.
Level 2 Lowest level for 2 level attributes)	T Sh 500,000 per month	Education offered after 4 years of service.	Regional headquarters	Insufficient	Heavy: Barely enough time to complete duties. Three hours of extra work per day.	No house is provided.	The place has unreliable mobile coverage, no electricity or water.
Level 3	T Sh 350,000 per month	Education offered after 6 years of service.	District headquarters				
Level 4 (Lowest level for 4 level attributes)	T Sh 200,000 per month	No education offered.	A 3-hour or more bus ride from the district headquarters				

T Sh = Tanzania shilling.
100,000 T Sh could be exchanged for about $100 at the time of the study.

2.3 Experimental design and construction of choice sets

2.3.1 Design

Once agreeing on attributes and levels, the researcher defines the choice sets, which are hypothetical jobs (or job profiles) resulting from combining the attributes and levels. Often, the combinations derived from the full set of attributes and levels (full factorial) result in too many choice sets to present to individuals. So, for example, in this study, the full factorial is _43*24=1024_ possible job profiles (3 attributes at 4 levels and 4 attributes at 2 levels). This implies _(1024*1023)/2=523,775_ possible choice sets. As seen in section 1, experimental design methods are commonly used to reduce the choice set to a manageable level, while allowing the researcher to infer preferences for all profiles.

In addition, the researcher must consider the specification of the utility function to be estimated at the design stage, taking account of potential interaction terms and the choice between labeled and generic experiments. Interaction terms were explored but not found to be significant, so the main effects of generic design are discussed here. Table 2.2 shows that 6 of the 7 attributes are modeled as dummy variables, and salary is modeled as continuous in the regression analyses. It also shows the regression coding labels for the variables.

Table 2.2 Attributes, regression coding, levels, and modeling

Attribute	Regression label	Level	Modeling
Salary and allowances	salary	T Sh 650,000/month T Sh 500,000/month T Sh 350,000/month T Sh 200,000/month	Continuous
Education opportunities/ possibility of upgrading qualifications	edu	edu_2 edu_4 edu_6 edu_0	Dummy variable
Location	loc	loc_dsm loc_reg loc_dis loc_3hour	Dummy variable
Availability of equipment and drugs	equipdrugs	equipdrugs_s equipdrugs_i	Dummy variable
Workload	work	work_normal work_heavy	Dummy variable
Housing	housing	housing_yes housing_no	Dummy variable
Infrastructure	infra	infra_good infra_bad	Dummy variable

Having defined the functional form of the utility function to be estimated, the researcher must then employ experimental design methods to derive the choice set. As shown in section 1, both orthogonal and D-efficient designs have been employed to date. Here a D-efficient design was developed, with no a priori assumptions made about the parameters. The design was developed from a computer program written by one of the researchers involved. Not all researchers conducting a DCE have the skills to write the experimental design program, and catalogues, software, and experts can help them generate such designs (see section 1 and the Uganda case study).

The experimental design applied in this study generated 32 choices (table 2.3).

Table 2.3 Experimental design for Tanzania study

choice-set	alt	salary	edu_0	edu_6	edu_4	edu_2	loc _3ho	loc _dis	loc _reg	loc _dsm	housing _no	housing _yes	work _heavy	work _normal	equip drugs_i	equip drugs_s	infra _bad	infra _good
1	1	3	0	0	0	1	1	0	0	0	1	0	0	1	0	1	1	0
1	2	3	0	1	0	0	0	0	0	1	0	1	1	0	1	0	0	1
2	1	0	0	0	0	1	0	0	0	1	1	0	0	1	1	0	1	0
2	2	2	1	0	0	0	0	0	0	1	1	0	1	0	0	1	0	1
3	1	1	0	0	0	1	0	0	0	1	1	0	1	0	0	1	0	1
3	2	3	1	0	0	0	0	1	0	0	0	1	0	1	0	1	1	0
4	1	2	1	0	0	0	1	0	0	0	1	0	1	0	0	1	1	0
4	2	1	1	0	0	0	0	0	1	0	1	0	1	0	1	0	0	1
5	1	2	1	0	0	0	1	0	0	0	0	1	0	1	1	0	1	0
5	2	1	0	0	1	0	0	0	1	0	0	1	0	1	0	1	0	1
6	1	0	0	0	1	0	0	0	1	0	0	1	1	0	0	1	1	0
6	2	3	0	1	0	0	0	0	0	1	1	0	0	1	0	1	0	1
7	1	3	0	0	1	0	0	0	0	1	0	1	1	0	0	1	0	1
7	2	1	0	0	1	0	0	1	0	0	0	1	1	0	1	0	1	0
8	1	1	0	1	0	0	0	1	0	0	0	1	0	1	1	0	0	1
8	2	1	0	0	0	1	0	0	1	0	1	0	1	0	1	0	1	0
9	1	3	0	0	0	1	0	0	1	0	0	1	0	1	1	0	1	0
9	2	0	1	0	0	0	0	1	0	0	0	1	1	0	1	0	0	1
10	1	0	0	1	0	0	0	0	0	1	1	0	1	0	0	1	1	0
10	2	3	0	0	0	1	0	0	0	1	0	1	0	1	0	1	0	1
11	1	3	0	1	0	0	1	0	0	0	1	0	0	1	1	0	1	0

choice-set	alt	salary	edu_0	edu_6	edu_4	edu_2	loc _3ho	loc _dis	loc _reg	loc _dsm	housing _no	housing _yes	work _heavy	work _normal	equip drugs_i	equip drugs_s	infra _bad	infra _good
11	2	0	0	0	0	1	1	0	0	0	1	0	0	1	0	1	0	1
12	1	0	1	0	0	0	0	0	0	1	0	1	0	1	1	0	1	0
12	2	2	0	0	1	0	1	0	0	0	1	0	1	0	1	0	0	1
13	1	1	0	0	1	0	1	0	0	0	0	1	0	1	1	0	1	0
13	2	3	0	0	1	0	0	0	1	0	1	0	1	0	1	0	0	1
14	1	0	0	1	0	0	1	0	0	0	1	0	1	0	1	0	1	0
14	2	2	0	0	1	0	0	0	1	0	0	1	0	1	1	0	0	1
15	1	3	0	0	0	1	1	0	0	0	0	1	1	0	0	1	0	1
15	2	2	0	0	1	0	0	0	0	1	1	0	0	1	0	1	1	0
16	1	2	0	1	0	0	0	1	0	0	0	1	1	0	0	1	0	1
16	2	1	0	1	0	0	0	0	1	0	0	1	0	1	1	0	1	0
17	1	1	1	0	0	0	1	0	0	0	1	0	0	1	1	0	0	1
17	2	2	0	0	0	1	0	0	0	1	0	1	1	0	0	1	1	0
18	1	2	0	1	0	0	0	0	1	0	0	1	1	0	1	0	0	1
18	2	0	0	1	0	0	0	0	0	1	0	1	0	1	0	1	1	0
19	1	1	1	0	0	0	0	0	1	0	0	1	1	0	0	1	1	0
19	2	0	0	0	1	0	1	0	0	0	0	1	1	0	1	0	0	1
20	1	2	0	0	0	1	1	0	0	0	0	1	1	0	1	0	1	0
20	2	0	1	0	0	0	0	0	1	0	1	0	1	0	0	1	0	1
21	1	0	1	0	0	0	0	0	0	1	1	0	0	1	1	0	0	1
21	2	3	0	0	1	0	0	1	0	0	0	1	1	0	0	1	1	0
22	1	1	0	0	0	1	0	0	1	0	0	1	1	0	1	0	0	1

choice-set	alt	salary	edu_0	edu_6	edu_4	edu_2	loc _3ho	loc _dis	loc _reg	loc _dsm	housing _no	housing _yes	work _heavy	work _normal	equip drugs_i	equip drugs_s	infra _bad	infra _good
22	2	1	0	0	0	1	0	1	0	0	1	0	0	1	1	0	1	0
23	1	3	1	0	0	0	0	0	0	1	0	1	1	0	1	0	1	0
23	2	0	0	0	0	1	0	0	1	0	0	1	0	1	1	0	0	1
24	1	2	0	0	1	0	0	1	0	0	1	0	0	1	1	0	0	1
24	2	1	1	0	0	0	1	0	0	0	0	1	0	1	0	1	1	0
25	1	3	0	0	1	0	1	0	0	0	1	0	1	0	0	1	1	0
25	2	2	0	0	0	1	0	1	0	0	1	0	1	0	1	0	0	1
26	1	0	0	1	0	0	0	0	1	0	0	1	1	0	0	1	0	1
26	2	0	0	0	1	0	0	0	1	0	1	0	0	1	1	0	1	0
27	1	0	1	0	0	0	0	0	1	0	0	1	1	0	1	0	1	0
27	2	1	1	0	0	0	0	1	0	0	1	0	1	0	0	1	0	1
28	1	2	0	1	0	0	1	0	0	0	0	1	0	1	0	1	0	1
28	2	3	0	1	0	0	0	1	0	0	1	0	1	0	1	0	1	0
29	1	1	0	1	0	0	0	1	0	0	1	0	0	1	0	1	1	0
29	2	3	0	0	0	1	0	1	0	0	1	0	1	0	0	1	0	1
30	1	0	0	0	1	0	0	1	0	0	1	0	0	1	0	1	0	1
30	2	2	1	0	0	0	0	1	0	0	1	0	1	0	1	0	1	0
31	1	1	0	0	1	0	0	0	0	1	0	1	1	0	1	0	0	1
31	2	2	0	1	0	0	0	0	1	0	1	0	0	1	0	1	1	0
32	1	2	0	0	0	1	0	1	0	0	0	1	0	1	0	1	1	0
32	2	2	0	0	0	1	1	0	0	0	1	0	0	1	1	0	0	1

The variables are:

choiceset: indicates the choice set from the DCE questionnaire. There were 32 choice sets.

alt: indicates the alternative within each choice set. Given each option had 2 choices, alt takes on the value of 1 or 2.

salary: is the continuous variable salary, which takes the value in the design coding of 0 to 3 where 0 = T Sh 200,000/month, 1 = T Sh 350,000/month, 2 = T Sh 500,000/month and 3 = T Sh 650,000/month.

edu_0, edu_6, edu_4, edu_2: 4 dummy variables for the education opportunities attribute. Any given alternative will always take a value of 1 for 1 of the dummies and 0 for all others (since only 1 level of the education attribute will be provided).

loc_3hour, loc_dis, loc_reg, loc_dsm: dummy variable levels for the location attribute

housing_yes, housing_no: dummy variable levels for the housing attribute

work_heavy, work_normal: dummy variable levels for the workload attribute

equipdrugs_i, equipdrugs_s: dummy variable levels for availability of equipment and drugs

infra_bad, infra_good: dummy variable levels for infrastructure attribute

In table 2.3 choice_set 1 has two alternatives. Alternative 1 (Job A) is defined as: salary 3 (T Sh 650,000 /month); edu_2; loc_3hour; housing_no; work_normal; equip_s; and infra_bad. Alternative 2 (Job B) has salary 3 (T Sh 650,000 /month); edu_6; loc_dsm; housing_yes; work_heavy; equip_i; and infra_good. An example of the first choice set in the design matrix is in figure 2.1.

Figure 2.1 Example of a choice set

Job A						
Availability of equipment and drugs	**Housing**	**Education opportunities/ possibility of upgrading qualifications**	**Workload**	**Infra-structure**	**Salary and allowances**	**Location**
Sufficient	No house is provided	Education offered after 2 years of service	Normal: Nearly enough time to complete duties. One hour of extra work per day	The place has unreliable mobile coverage, no electricity or water	T Sh 650,000/ month	A 3-hour or more bus ride from the district headquarters

Job B						
Availability of equipment and drugs	**Housing**	**Education opportunities/ possibility of upgrading qualifications**	**Workload**	**Infra-structure**	**Salary and allowances**	**Location**
Insufficient	A decent house is provided	Education offered after 6 years of service	Heavy: Barely enough time to complete duties. Three hours of extra work per day	The place has mobile coverage, electricity and water	T Sh 650,000/ month	Dar es Salaam

Considering your current situation, which of the two jobs would you choose?

Job A: ☐ **Job B:** ☐

2.3.2 Checking properties

Section 1.2.2 outlined the characteristics of a good design. While a D-efficient design was developed, the properties of a good (but not perfect) design were still expected to be present. It is useful to check the design's properties. Below, the orthogonality, level balance, and minimum overlap for the design in table 2.3 are considered. Given there was no a priori information available on parameter estimates, utility balance was not considered.

Orthogonality

This criterion requires that the levels of each attribute vary independently of each other. The correlations for this design, approximated with the pwcorr command in Stata, which uses the PPM for estimating pairwise correlation coefficients and their significance, is shown in table 2.4. All are sufficiently low not to cause concern.

Table 2.4 Correlation matrix

	salary	edu_0	edu_6	edu_4	edu_2	loc_ 3hour	loc_dis	loc_reg	loc_
salary	dsm								
edu_0	-.155	1							
edu_6	.025	-.319***	1						
edu_4	.008	-.333***	-.319***	1					
edu_2	.121	-.347***	-.333***	-.347***	1				
loc_3hours	.106	0.000	-.064	0.000	.061	1			
loc_dis	.106	0.000	.021	0.000	-.020	-.333***	1		
loc_reg	-.232*	-.020**	.001	.061	-.041	-.347***	-.347***	1	
loc_dsm	.025	.021	.042	-.064	.001	-.319***	-.319***	-.333***	1
housing_no	-.001	.018	-.020	-.054	.054	.090	.090	-.158	-.020
housing_yes	.001	-.018	.020	.054	-.054	-.090	-.090	.158	.020
work_hard	.072	-.109	-.072	.036	-.073	-.109	.036	.069	.002
work_normal	-.072	.109	.072	-.036	.073	.109	-.036	-.069	-.002
equipdrug_n	-.127	.036	-.072	.036	-.002	.036	-.036	.140	-.146
equipdrug_s	.127	-.036	.072	-.036	.002	-.036	.036	-.140	.146
infra_bad	.014	-.072	.037	-.072	-.035	.072	0.000	-.035	-.037
infra_good	-.014	.072	-.037	.072	.035	-.072	0.000	.035	.037

*, **, and *** indicate significance at 10%, 5% and 1% level respectively.

Table 2.4 Correlation matrix (continued)

	housing_ no	housing_ yes	work_ hard	work_ normal	equip drug_n	equip drug_s	infra_ bad	infra _good
housing_no	1							
housing_yes	-1***	1						
work_hard	-.029	.029	1					
work_normal	.029	-.029	-1	1				
equipdrug_n	-.029	.029	-.004	.004	1			
equipdrug_s	.029	-.029	.004	-.004	-1	1		
infra_bad	-.031	.031	-.125	.125	0.000	0.000	1	
infra_good	.031	-.031	.125	-.125	0.000	0.000	-1***	1

*, **, and *** indicate significance at 10%, 5% and 1% level respectively.

Level balance

Level balance requires all levels of each attribute to appear with equal frequency across profiles. Thus for a 2-level attribute, each level should appear in 50% of the profiles, and for a 4-level attribute each level should appear in 25% of the profiles. In this case each level of the salary attribute should appear in 25% of the job profiles, the same applies to each level of the education and location attributes. For the remainder of the attributes, each level should appear in 50% of the job profiles. The level balance of the design applied in this study is in table 2.5, which shows that the design has a relatively good, if not perfect, level balance.

Table 2.5 Level balance

	Number of appearances	%
650,000 T Sh per month	16	25.0
500,000 T Sh per month	16	25.0
350,000 T Sh per month	17	26.6
200,000 T Sh per month	15	23.4
edu_0	16	25.0
edu_6	15	23.4
edu_4	16	25.0
edu_2	17	26.6
loc_3hour	16	25.0
loc_dis	16	25.0
loc_reg	17	26.6
loc_dsm	15	23.4
housing_no	31	48.4
housing_yes	33	51.6
work_normal	34	53.1
work_heavy	30	46.9

	Number of appearances	%
equipdrugs_i	34	53.1
equipdrugs_s	30	46.9
infra_good	32	50.0
infra_bad	32	50.0

Minimum overlap
This criterion requires that a repeated attribute level within a choice set be minimized. This ensures that the experiment provides maximum information on respondents' trade-offs. If an attribute takes the same level in each choice, no information is revealed about preferences. As can be seen from the design in table 2.5, there are a few overlaps of attributes within the choice sets in this study. For example, in the first choice set, presented in figure 2.1.1 the salary is the same in job A and job B.

2.3.3 Supplemental questions

Supplemental questions were asked about demographics, including sex and age as well as previous experience from rural areas; educational experience, including prior training programs; work experience, including previous experience of working in rural areas; and motivational issues, including reasons why respondents became health workers, what they expected of the future, etc. The questionnaire became relatively long—49 questions prior to the DCE exercise. The questionnaire (including the DCE) took on average 1.5 hours to complete; students were therefore given a modest snack and a soda in the middle of the session to keep their energy up. Information collected about sex, rural background, and willingness to help others was included in the subgroup analysis presented in Kolstad (2011). It is often interesting to collect these kinds of supplemental data because, as in Kolstad (2011), different types of health workers have different preferences.

2.4 Development of the questionnaire, pretesting, and data collection

2.4.1 Warm-up question

One warm-up exercise was included, where respondents were introduced to the choice situation. The questionnaire also included a one-page introduction to the task. The researcher explained this example on the blackboard, stressed the importance of considering all attributes, and answered questions before the students completed the choices.

2.4.2 Pilot questionnaire

The DCE was tested in a pilot with 30 students at the Kilosa Clinical Officer Training Centre (this training center did not participate in the main data collection). The respondents completed a relatively long questionnaire before the DCE exercise and were then interviewed about various issues concerning their participation. The focus was on three parts of the DCE:

- Formulation of attributes and levels: were the attributes and levels clear and did they have the right range? Were any important attributes lacking for the choices to be meaningful or were any included attributes perceived not relevant when making choices?
- Was the task understood? Were instructions good enough? Were all attributes traded off for each other?
- How did the students experience the exercise: were there too many choices to make? Was it fun, boring, etc?

As a result of the pilot, the formulation of some of the attribute levels was changed to make them clearer to the respondents and to get the salary levels right in particular. Moreover, some of the questions respondents were asked before they participated in the DCE were reformulated.

The pilot also indicated that respondents found 32 choices a large number to complete. The 32 were therefore divided into two blocks, with each respondent facing 16 choices (half the respondents were presented with the first 16 choice sets, the other half with the next 16 choice sets).

Choices in the pilot questionnaire included an opt-out option. This option was frequently chosen by respondents. When asked their reasons for choosing the opt-out in the interview, many respondents commented that they were comparing the option to their "dream job". Describing the opt-out was therefore difficult (since respondents had no current job, and it was difficult to identify a typical clinical officer job), and since all but one of the participants in the pilot said that they wanted a career in the health system in Tanzania after their studies, it was decided a forced choice would be the most relevant type of choice for this group. That is, respondents were forced to choose between Job A and Job B (figure 1.1).

2.4.3 Main data collection

The data were collected during the autumn of 2007 with the help of four local research assistants traveling in teams of two. As whole classes were visited at the time, it was practical to have two assistants at each location, but in other contexts (depending on the number of respondents, and safety issues, etc.), it may well be sufficient to have one researcher or research assistant. The principal investigator traveled between the teams. Some 320 clinical officer finalists (around 60% of all clinical officer finalists in Tanzania) from 10 randomly selected schools all over the country took part in the DCE. All finalists in these schools were invited to participate.

The data were mostly collected during school time, on the school premises. This largely explains the response rate of around 96%, which is high for a DCE. The finalists were divided into groups of around 15 and seated in their classrooms or another suitable room. They were given a plenary introduction to the study, signed consent forms, and were guided through a couple of examples of choice sets before completing a paper version of the DCE by themselves (often at their desks).

Participation was voluntary, and students were not compensated in any way, but were offered a soda and a small snack because of the long completion time. In addition to the DCE choices, the respondents answered a series of questions which covered, among other things, their background, motivation, beliefs, and attitudes. The questionnaire took on average 1.5 hours to complete, confirming that it was wise to let each respondent make only 16 choices instead of the total of 32 choices generated by the design. Data collectors spent between one and two days at each training center.

If the intention is to conduct a DCE on more experienced health workers already working in different areas of the country, logistical challenges on reaching the respondents will probably be greater, because more travel will be required, and so the DCE is likely to be more costly. However, if conducted correctly, they will provide important and valuable information on the preferences of the existing stock of health workers, and the gains may more than outweigh the costs.

Even though relatively thorough testing had been carried out in the pilot, 20 students were randomly held back, 2 at each location, after the completion of the DCE. These students were asked to explain how they made their choices. This was done to check that the task was understood, that real trade-offs were being made, and to get a better impression of the reasoning behind the choices made. The interviews were very reassuring in the sense that all students reported making trade-offs between the attribute levels and were able to reconstruct and demonstrate the trade-offs they had made.

2.5 Data input

This section is included because the data matrix generated from DCEs is quite different from that generated for most questionnaires. One feature common to all DCE datasets is that respondents answer more than one discrete choice question, resulting in multiple observations for each individual. Furthermore, choice sets presented to individuals contain two or more alternatives, giving multiple observations for each choice set.

The number of observations in a dataset depends on the number of respondents, the number of choice sets per respondent and the number of alternatives in each choice set. For instance, in the study covered here each choice set has two alternatives (Job A and Job B), so each choice set contributes two observations to the dataset. Moreover, each respondent is presented with 16 choices. As each choice contributes two observations and each respondent faces 16 choices, there are 32 observations per respondent (16 choices x 2 observations per choice). A sample of the final data matrix (an extract from the full dataset) for the case study is in table 2.6. Most variable names in this table refer to those in table 2.2.

As with any dataset it is useful to start by ordering the variables in some logical way. One suggestion—followed here—is to present all the variables in a sequence that first describes how the data are organized (such as respondent identifier, choice set identifier), then present the independent variables from the experimental design (attribute levels) followed by the dependent variable (what option respondents chose). Datasets also include other variables relating to the individual, such as socioeconomic characteristics.

The variables are:

personid: The first variable is an identification variable unique to each respondent. It will be the same for the first 32 rows, then for the next 32 rows etc.

obsid: Stata requires a variable indicating each unique choice made. This increases successively for each choice.

alt: the alternative within each choice (where alt=1 represents the first and alt=2 the second alternative in each choice set (Job A and B respectively).

cno: represents the choice number in the DCE questionnaire; as each respondent made 16 choices cno will range from 1 to 16.

choiceset: since the design consists of 32 different choice sets, there is a variable indicating which of the 32 choice sets are being observed.

(Two identical choice sets presented to different respondents would thus have the same choice-set value but different obsid values. For individual 1 the obsid values will range from 1 to 16, for individual 2 they will range from 17 to 32, for individual 3 they will range from 33 to 48. Thus, for choice 3 presented to respondent 1 obsid=3, for choice 3 presented to respondent 2 obsid=19, for choice 3 presented to respondent 3 obsid=35 etc.)

The attributes in this study are a mixture of continuous and categorical dummy variables.

salary: is the salary attribute taking the values in the dataset that correspond to the levels presented in the questionnaire. Salary is treated as a continuous variable in the regression analysis, and it has thus been given one column only (unlike the categorical dummy attributes).

All categorical attributes were entered as dummy-coded variables. Here the effect of a level of an attribute is estimated relative to a base comparator or reference point.

edu_2, edu_4, edu_6 and edu_0,: there are 4 levels for the education attribute (education offered after 2, 4, or 6 years, and no education offered). Dummy variables take the value of 1 if the level is present in the alternative and 0 otherwise. For instance, the first alternative in table 2.6 offers education after 2 years of service. Thus, *edu_2*=1, while *edu_4*=0, *edu_6*=0 and *edu_0*=0.

Table 2.6 Final data matrix

person id	obsid	alt	cno	choice-set	salary	edu _0	edu _6	edu _4	edu _2	loc_ 3hour	loc _dis	loc _reg	loc_ dsm	housing _no	housing _yes	work_ heavy	work_ normal	equip drugs_i	equip drugs_s	infra _bad	infra _good	const	choice	sex	rural back-ground
1	1	1	1	1	650,000	0	0	0	1	1	0	0	0	1	0	0	1	0	1	1	0	0	0	1	0
1	1	2	1	1	650,000	0	1	0	0	0	0	0	1	0	1	1	0	1	0	0	1	1	1	1	0
1	2	1	2	2	200,000	0	0	0	1	0	0	0	1	1	0	0	1	1	0	1	0	0	1	1	0
1	2	2	2	2	500,000	1	0	0	0	0	0	0	1	1	0	1	0	0	1	0	1	1	0	1	0
1	3	1	3	3	350,000	0	0	0	1	0	0	0	1	1	0	1	0	0	1	0	1	0	1	1	0
1	3	2	3	3	650,000	1	0	0	0	0	1	0	0	0	1	0	1	0	1	1	0	1	0	1	0
1	4	1	4	4	500,000	1	0	0	0	1	0	0	0	1	0	1	0	0	1	1	0	0	1	1	0
1	4	2	4	4	350,000	1	0	0	0	0	0	1	0	1	0	1	0	1	0	0	1	1	0	1	0
1	5	1	5	5	500,000	1	0	0	0	1	0	0	0	0	1	0	1	1	0	1	0	0	1	1	0
1	5	2	5	5	350,000	0	0	1	0	0	0	1	0	0	1	0	1	0	1	0	1	1	0	1	0
1	6	1	6	6	200,000	0	0	1	0	0	0	1	0	0	1	1	0	0	1	1	0	0	0	1	0
1	6	2	6	6	650,000	0	1	0	0	0	0	0	1	1	0	0	1	0	1	0	1	1	1	1	0
1	7	1	7	7	650,000	0	0	1	0	0	0	0	1	0	1	1	0	0	1	0	1	0	1	1	0
1	7	2	7	7	350,000	0	0	1	0	0	1	0	0	0	1	1	0	1	0	1	0	1	0	1	0
1	8	1	8	8	350,000	0	1	0	0	0	1	0	0	0	1	0	1	1	0	0	1	0	0	1	0
1	8	2	8	8	350,000	0	0	0	1	0	0	1	0	1	0	1	0	1	0	1	0	1	1	1	0
1	9	1	9	9	650,000	0	0	0	1	0	0	1	0	0	1	0	1	1	0	1	0	0	1	1	0
1	9	2	9	9	200,000	1	0	0	0	0	1	0	0	0	1	1	0	1	0	0	1	1	0	1	0
1	10	1	10	10	200,000	0	1	0	0	0	0	0	1	1	0	1	0	0	1	1	0	0	1	1	0
1	10	2	10	10	650,000	0	0	0	1	0	0	0	1	0	1	0	1	0	1	0	1	1	0	1	0
1	11	1	11	11	650,000	0	1	0	0	1	0	0	0	1	0	0	1	1	0	1	0	0	1	1	0
1	11	2	11	11	200,000	0	0	0	1	1	0	0	0	1	0	0	1	0	1	0	1	1	0	1	0
1	12	1	12	12	200,000	1	0	0	0	0	0	0	1	0	1	0	1	1	0	1	0	0	0	1	0

person id	obsid	alt	cno	choice-set	salary	edu _0	edu _6	edu _4	edu _2	loc_ 3hour	loc _dis	loc _reg	loc_ dsm	housing _no	housing _yes	work_ heavy	work_ normal	equip drugs_i	equip drugs_s	infra _bad	infra _good	const	choice	sex	rural back-ground
1	12	2	12	12	500,000	0	0	1	0	1	0	0	0	1	0	1	0	1	0	0	1	1	1	1	0
1	13	1	13	13	350,000	0	0	1	0	1	0	0	0	0	1	0	1	1	0	1	0	0	0	1	0
1	13	2	13	13	650,000	0	0	1	0	0	0	1	0	1	0	1	0	1	0	0	1	1	1	1	0
1	14	1	14	14	200,000	0	1	0	0	1	0	0	0	1	0	1	0	1	0	1	0	0	0	1	0
1	14	2	14	14	500,000	0	0	1	0	0	0	1	0	0	1	0	1	1	0	0	1	1	1	1	0
1	15	1	15	15	650,000	0	0	0	1	1	0	0	0	0	1	1	0	0	1	0	1	0	1	1	0
1	15	2	15	15	500,000	0	0	1	0	0	0	0	1	1	0	0	1	0	1	1	0	1	0	1	0
1	16	1	16	16	500,000	0	1	0	0	0	1	0	0	0	1	1	0	0	1	0	1	0	0	1	0
1	16	2	16	16	350,000	0	1	0	0	0	0	1	0	0	1	0	1	1	0	1	0	1	1	1	0
2	17	1	1	17	350,000	1	0	0	0	1	0	0	0	1	0	0	1	1	0	0	1	0	1	1	1
2	17	2	1	17	500,000	0	0	0	1	0	0	0	1	0	1	1	0	0	1	1	0	1	0	1	1
2	18	1	2	18	500,000	0	1	0	0	0	0	1	0	0	1	1	0	1	0	0	1	0	1	1	1
2	18	2	2	18	200,000	0	1	0	0	0	0	0	1	0	1	0	1	0	1	1	0	1	0	1	1
2	19	1	3	19	350,000	1	0	0	0	0	0	1	0	0	1	1	0	0	1	1	0	0	0	1	1
2	19	2	3	19	200,000	0	0	1	0	1	0	0	0	0	1	1	0	1	0	0	1	1	1	1	1
2	20	1	4	20	500,000	0	0	0	1	1	0	0	0	0	1	1	0	1	0	1	0	0	1	1	1
2	20	2	4	20	200,000	1	0	0	0	0	0	1	0	1	0	1	0	0	1	0	1	1	0	1	1
2	21	1	5	21	200,000	1	0	0	0	0	0	0	1	1	0	0	1	1	0	0	1	0	0	1	1
2	21	2	5	21	650,000	0	0	1	0	0	1	0	0	0	1	1	0	0	1	1	0	1	1	1	1
2	22	1	6	22	350,000	0	0	0	1	0	0	1	0	0	1	1	0	1	0	0	1	0	0	1	1
2	22	2	6	22	350,000	0	0	0	1	0	1	0	0	1	0	0	1	1	0	1	0	1	1	1	1
2	23	1	7	23	650,000	1	0	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	0	1	1
2	23	2	7	23	200,000	0	0	0	1	0	0	1	0	0	1	0	1	1	0	0	1	1	1	1	1
2	24	1	8	24	500,000	0	0	1	0	0	1	0	0	1	0	0	1	1	0	0	1	0	1	1	1

person id	obsid	alt	cno	choice-set	salary	edu _0	edu _6	edu _4	edu _2	loc_ 3hour	loc _dis	loc _reg	loc_ dsm	housing _no	housing _yes	work_ heavy	work_ normal	equip drugs_i	equip drugs_s	infra _bad	infra _good	const	choice	sex	rural back-ground
2	24	2	8	24	350,000	1	0	0	0	1	0	0	0	0	1	0	1	0	1	1	0	1	0	1	1
2	25	1	9	25	650,000	0	0	1	0	1	0	0	0	1	0	1	0	0	1	1	0	0	1	1	1
2	25	2	9	25	500,000	0	0	0	1	0	1	0	0	1	0	1	0	1	0	0	1	1	0	1	1
2	26	1	10	26	200,000	0	1	0	0	0	0	1	0	0	1	1	0	0	1	0	1	0	0	1	1
2	26	2	10	26	200,000	0	0	1	0	0	0	1	0	1	0	0	1	1	0	1	0	1	1	1	1
2	27	1	11	27	200,000	1	0	0	0	0	0	1	0	0	1	1	0	1	0	1	0	0	0	1	1
2	27	2	11	27	350,000	1	0	0	0	0	1	0	0	1	0	1	0	0	1	0	1	1	1	1	1
2	28	1	12	28	500,000	0	1	0	0	1	0	0	0	0	1	0	1	0	1	0	1	0	1	1	1
2	28	2	12	28	650,000	0	1	0	0	0	1	0	0	1	0	1	0	1	0	1	0	1	0	1	1
2	29	1	13	29	350,000	0	1	0	0	0	1	0	0	1	0	0	1	0	1	1	0	0	1	1	1
2	29	2	13	29	650,000	0	0	0	1	0	1	0	0	1	0	1	0	0	1	0	1	1	0	1	1
2	30	1	14	30	200,000	0	0	1	0	0	1	0	0	1	0	0	1	0	1	0	1	0	0	1	1
2	30	2	14	30	500,000	1	0	0	0	0	1	0	0	1	0	1	0	1	0	1	0	1	1	1	1
2	31	1	15	31	350,000	0	0	1	0	0	0	0	1	0	1	1	0	1	0	0	1	0	0	1	1
2	31	2	15	31	500,000	0	1	0	0	0	0	1	0	1	0	0	1	0	1	1	0	1	1	1	1
2	32	1	16	32	500,000	0	0	0	1	0	1	0	0	0	1	0	1	0	1	1	0	0	1	1	1
2	32	2	16	32	500,000	0	0	0	1	1	0	0	0	1	0	0	1	1	0	0	1	1	0	1	1
3	33	1	1	1	650,000	0	0	0	1	1	0	0	0	1	0	0	1	0	1	1	0	0	0	0	1
3	33	2	1	1	650,000	0	1	0	0	0	0	0	1	0	1	1	0	1	0	0	1	1	1	0	1
3	34	1	2	2	200,000	0	0	0	1	0	0	0	1	1	0	0	1	1	0	1	0	0	1	0	1
3	34	2	2	2	500,000	1	0	0	0	0	0	0	1	1	0	1	0	0	1	0	1	1	0	0	1
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…

For the same alternative the location is remote, that is, loc_3hour=1 while loc_dist=0, loc_reg=0 and loc_dsm=0. Similarly, no housing is offered so housing_no=1 while housing_yes=0 etc.

const: when using conditional logit the researcher has to include a constant term in the data matrix, indicating whether a row of data represents Job A or Job B in a choice set. It is designed as a dummy taking the values 0 and 1. The constant is often included in the model as a test for specification error (Scott 2001). Further, when dummy variables are included it soaks up the preference for the base comparator (Bech and Gyrd-Hansen 2005).

choice: is the dependent variable, indicating their choice of job (Job A or Job B). This is represented as a dichotomous variable taking the value of 1 for the chosen alternative and zero for the one not chosen. From table 2.6 it can be seen that the first respondent chose Job B in the first choice set (obsid=1, illustrated above in figure 2.1.1), and Job A in the next choice set (obsid=2).

Alongside DCE responses, information was collected about respondents' socioeconomic characteristics, such as sex and rural background. Given that each respondent has more than 1 row in the dataset, this information is copied on to each row related to an individual in the same manner as the id variable. From the example presented in table 2.6, respondent number 1 is a female (male coded 0 and female coded 1) with a nonrural background (nonrural background coded 0 and rural background coded 1), while respondent 2 is a female with a rural background and respondent 3 is a male with a rural background.

As socioeconomic characteristics do not vary within a choice, these cannot be added into the regression model directly. Including interaction terms between respondent characteristics and attributes allows slope coefficients to differ across subgroups. Such variables could be created by simply multiplying the variables of interest. For example, if the researcher is interested in whether preferences for salary vary according to the sex of the respondent, he or she can create a variable, "salary-sex", which is simply "salary*sex". This can then be entered into the regression model.

Most of the above can be set up before the data are collected. Often before administering a questionnaire to the sample, data are simulated for the response variable. The model the researcher intends to use for estimation is then applied to this simulated data as a check that the data are correctly coded and that the design allows the estimation of parameters of interest.

In this study the data collected on paper were entered twice (by two persons) in a software package called Epi data (http://www.epidata.dk/). This package is free for download and is well suited for this kind of data input as it allows the person in charge to make small programs in advance, which reduces incorrect data input substantially. However, it is unimportant which software is used for data entry as long as the data are entered correctly, since almost any format can be converted into data files for most software packages.

2.6 Model estimation and interpretation

2.6.1 Set-up of the basic regression model

Researchers should be aware of the requirements of the statistical software packages they are using to analyze the data. This section presents useful tips to prepare data for analysis in a commonly used software package, Stata.

The final sample used in the analysis comprised 296 respondents, each providing responses to 16 completed choices and resulting in 9472 observations (296 individuals x 16 choices x 2 options for each choice). Following on from section 1, the probability a respondent will select a specified job is modeled. The probability of choosing a given job is determined by the indirect utility. Here it is assumed that this is linear and additive and of the form:

$$V = \beta_1 salary + \beta_2 edu\_6 + \beta_3 edu\_4 + \beta_4 edu\_2 + \beta_5 loc\_dis+ + \beta_6 loc\_reg$$
$$+ \beta_7 loc\_dsm + \beta_8 housing\_yes + \beta_9 work\_normal + \beta_{10} equipdrugs\_s$$
$$+ \beta_{11} infra\_good + \beta_{12} const + \varepsilon$$

where V is the utility derived from a given job, ε refers to the error term as described in section 1, and all other variables are defined above.

Given the binary choices presented to individuals, the binary logit model and conditional logit model could be used to analyze the data. In Stata researchers can do a logit regression by using the logit command. However, when the data are presented as in table 2.6., a conditional logit should be used (since the data are stacked, with each option within a choice on a different row). This will yield exactly the same results as the binary logit, which requires the options to be on one row, and differenced, and therefore analyzed with binary logit (logit) or random effects binary logit (xtlogit, to allow for multiple observations).

The way this data were set up, the clogit command was used. The exact syntax in Stata is:
clogit choice salary edu_6 edu_4 edu_2 loc_dis loc_reg loc_dsm housing_yes work_normal equipdrugs_s infra_good const, group(obsid)

where all variables are defined above. The group(obsid) indicates which rows of data that came from the same choice set.

The regression results and the corresponding WTP measures are in table 2.7.

Table 2.7 Regression results and WTP

Attributes	Regression labeling	Betas	Coefficients[a]	WTP[b]
Salary	salary	β_1	.003***	
			(.0002)	
Education (relative to no education offered)				
Education after 6 years of service	edu_6	β_2	.354***	110.021
			(.0931)	(51.324462 - 168.71692)
Education after 4 years of service	edu_4	β_3	.707***	219.547
			(.0747)	(171.00109 - 268.09271)
Education after 2 years of service	edu_2	β_4	1.149***	356.758
			(.0687)	(306.88717 - 406.62958)
Location (relative to 3 miles + from district HQ)				
District headquarters	loc_dis	β_5	.216 ***	67.16
			(.0701)	(24.040591- 110.27902)
Regional headquarters	loc_reg	β_6	.021	6.566
			(.0650)	(-32.951269 - 46.082906)
Dar es Salaam	loc_dsm	β_7	-.308***	-95.657
			(.0771)	(-143.49899 -47.774036)

Attributes	Regression labeling	Betas	Coefficients[a]	WTP[b]
Location (relative to 3 miles + from district HQ)				
Decent housing offered	housing_yes	β_8	.216***	67.171
			(.0493)	(36.892623 - 97.449622)
Workload (relative to heavy workload)				
Normal workload	work_normal	β_9	-.603	-19.506
			(.0482)	(-49.291817 - 10.280629)
Equipment (relative to insufficient equipment)				
Sufficient equipment and drugs	equipdrugs_s	β_{10}	.413***	128.145
			(.0433)	(99.165972 - 157.12438)
Infrastructure (relative to poor infrastructure)				
Decent infrastructure	infra_good	β_{11}	.716***	222.36
			(.0381)	(195.32815 - 249.40972)
Constant	Const	β_{12}	-.017	
			(.0398)	
Number of respondents			296	
Number of observations			9472	
Log Likelihood			2424.2108	
Pseudo R2			0.2513	

a. Standard errors in parentheses. b. Confidence intervals in parentheses.
* significant at 10% level, ** significant at 5% level, *** significant at 1% level.

When looking at the output of a DCE the first thing the researcher should do is see whether the attributes are significant, and therefore have an impact on the probability of choosing an alternative. He or she should consider the sign of the coefficient, where significant. A positive sign implies that the attribute has a positive impact on the take-up of a given job; a negative coefficient the opposite.

β_2 for instance, shows that having education opportunities after six years of service, rather than none at all, increases the utility of the job by 0.354. Similarly, β_5 shows that if the job is in the district headquarters, the utility increases by 0.216. Most coefficients in table 2.7 have the expected signs. All else equal, the respondents prefer a job with higher salaries and the possibility of further education after 2, 4, and 6 years to no further education, the earlier the better. They prefer a job where sufficient equipment is provided to one without, and a job that offers decent housing and infrastructure to one that does not.

The respondents prefer to work in district headquarters rather than regional headquarters or in a location that is a 3-hour (or longer) bus ride from the district headquarters. The least popular location is the capital, Dar es Salaam. This may seem surprising but there are several plausible explanations. Living costs are very high in Dar es Salaam compared to other cities in Tanzania, but perhaps more important, the likelihood of being in charge of a health facility and to be able to practice as a clinician is smaller in Dar es Salaam, where most of the formally qualified medical doctors are based.

The coefficients for the workload attribute and for being located in regional headquarters are insignificant. There could be two reasons: either the researcher was unable to estimate the coefficients efficiently with the model used, or there is too much heterogeneity in the preferences for these attributes.

The attributes are measured in different ways—the continuous salary coefficient indicates how much the utility increases by having one extra shilling, while the other coefficients measure the change in utility from the references category. They are not, therefore, directly comparable.

2.6.2 Willingness to pay

Within the context of workforce issues, inclusion of a price proxy (such as salary) allows the researcher to estimate of the monetary value of attributes of a job, that is, how much salary a respondent would be willing to give up to have an improvement in other aspects of the job. As seen in section 1, this can be estimated as the ratio of the value of the coefficient of interest to the negative of the cost attribute—in this case, salaries. For example, how much monthly salary that respondents are willing to sacrifice to receive education after six years rather than no education can be estimated as:

$$WTP(edu_6) = -\frac{\partial U / \partial(edu_6)}{\partial U / \partial wage} = -\frac{\beta_2}{\beta_1} = -\frac{0.354}{0.003} = -11\mathbf{0.021}$$

Similarly, how much monthly salary respondents are willing to sacrifice for working in the district headquarters rather than a remote area is given by:

$$WTP(loc\_dis) = -\beta_5 / \beta_1 = 0.216/0.003 = 67.16$$

And how much monthly salary respondents are willing to sacrifice for working at a facility with sufficient equipment and drugs is given by:

$$WTP(equipdrugs\_s) = -\beta_{10} / \beta_1 = 0.413/0.003 = 128.145$$

The WTP values can be easily estimated by hand (with a calculator) as shown above, or in a software package such as Excel. The figures in table 2.7 are calculated within Stata and may deviate somewhat from the results obtained with a calculator, simply because of the number of decimals included in the coefficients above.

The advantage of estimating WTP within Stata is that the program will also estimate the confidence intervals (reported in parentheses under the WTP estimates).

Hole (2007) describes four approaches to estimating confidence intervals for WTP estimates within a DCE: the delta, Fieller, Krinsky Robb, and bootstrap methods. He also compares their accuracy using simulated data. Hole (2007) concludes that the four methods give very similar results when the model is correctly specified and the cost coefficient is relatively precisely estimated (t-stat > 10 in absolute value). He also found, more generally, that the methods tend to give similar results (personal communication).

Hole's wtp command for Stata implements the delta method, the Fieller method, and the Krinsky Robb (parametric bootstrap) method. Delta method confidence intervals can also be calculated using the nlcom command in Stata. Nonparametric bootstrap confidence intervals can be estimated using Stata's bootstrap command. The wtpcikr command in Stata can also be used to generate Krinsky Robb confidence intervals (this command was designed for use with contingent valuation data). Stata's wtp command implements the same method (Krinsky Robb) for models estimated using data from choice experiments. However, these commands only work for standard logit commands such as logit, random effects logit (xtlogit), and conditional logit (clogit).

To estimate WTP measures in Stata using either the nlcom command or the wtp command, both commands are entered immediately after the conditional logit command (clogit).

So, for example, to calculate the willingness to sacrifice salary for education after 6 years of service, rather than no education opportunities, the nlcom command will be the following:

nlcom (_b[edu_6])/- (_b[salary])

Alternatively, one can use the wtp command. The cost attribute, the salary in this case, will then have to be defined as a negative and placed in front of the other attributes, as specified below:

gen msalary=-salary
wtp msalary edu_6 edu_4 edu_2 loc_dis loc_reg loc_dsm housing_yes work_normal equipdrugs_s infra_good

In order to compress the information and make it more intuitive, WTP values may sometimes be presented graphically. An example of this is provided in the case study from Uganda.

2.6.3 Uptake rate

A useful output when using DCEs to look at recruitment and retention is how the probability of choosing a given post changes as levels of attributes are changed. One option is to consider the change in the probability of taking the baseline job (the reference category for all dummies) due to a change of the level in one of the job attributes. Then, the regression model appears as that outlined for the WTP model.

The logit probability of choosing alternative *i* rather than alternative *j* is given by:

$$P_i = \frac{e^{\beta' x_i}}{\sum e^{\beta' x_j}}$$

where *x* is a vector of attribute coefficients. Using this equation, the change in the probability of taking the baseline job because of a change in one of the job attributes—say, the salary is raised to 350,000 Tanzania shillings (T Sh) per month—is then (as long as all other attributes remain equal) given by:

$$P_{wage=350} - P_{wage=200} = \frac{e^{\beta_1 * 350}}{e^{\beta_1 * 200} + e^{\beta_1 * 350}} - \frac{e^{\beta_1 * 200}}{e^{\beta_1 * 200} + e^{\beta_1 * 350}}$$

$$= \frac{e^{0.003*350}}{e^{0.003*200} + e^{0.003*350}} - \frac{e^{0.003*200}}{e^{0.003*200} + e^{0.003*350}}$$

$$= 0.261$$

The syntax in Stata (straight after the conditional logit regression) for calculating such a change in the probability is the following:

nlcom exp(_b[salary]*350)/(exp(_b[salary]*200) + exp(_b[salary]*350)) - exp(_b[salary]*200)/(exp(_b[salary]*200) + exp(_b[salary]*350))

which can be reduced/simplified to:

```
nlcom (exp(_b[salary]*350) - exp(_b[salary]*200))
/(exp(_b[salary]*200) + exp(_b[salary]*350))
```

Then, as for the WTP calculations, Stata will also calculate confidence intervals.

Similarly, if one wanted to calculate the uptake rate of a policy of providing housing, the syntax would be:

```
nlcom exp(_b[salary]*200+_b[house_yes])/(exp(_b[salary]*200) +
exp(_b[salary]*200+_b[house_yes])) -
exp(_b[salary]*200)/(exp(_b[salary]*200) + exp(_b[salary]*200+_b[house_yes]))
```

which can be reduced/simplified to:

```
nlcom (exp(_b[salary]*200+_b[house_yes]) - exp(_b[salary]*200))
/(exp(_b[salary]*200) + exp(_b[salary]*200+_b[house_yes]))
```

In Kolstad (2011) this analysis was also done for three subgroups based on sex, rural background, and willingness to help other people. Background information from the survey was then used to divide the sample into subgroups and to carry out separate analyses on each subgroup in exactly the same manner as shown above. Alternatively, one can include interaction effects in the regression directly.

2.6.4 Comparison of results: DCE versus in-depth interviews

The in-depth interviews were conducted to inform the choice of attributes and levels for the DCE. It is therefore maybe unsurprising that the results from the DCE largely confirm the findings from those interviews. The DCE findings also concur well with findings from other qualitative studies and studies applying simpler quantitative methods. So, is it really necessary to go through with a challenging exercise like a DCE?

The answer is that it depends on what kind of information one wants to get out of the study. The aim of this study was to examine the importance of different attributes when clinical officers make job choices. Trade-offs among these attributes would be established, and predictions of uptake of different recruitment and retention packages produced.

This approach was chosen because policy makers will typically be interested in answers to questions like: How much salary would a respondent be willing to give up for improvements in other aspects of the job? Or, How would the probability of taking a rural job change if houses are provided as an incentive to go/remain there? Such questions cannot be answered with information from focus group and in-depth interviews or simple ranking exercises, but require either data stemming from a DCE or revealed preference data in combination with DCE data.

2.6.5 Presentation of results for policy

Tables like table 2.7 may be challenging to interpret and of little use to policy makers who have no formal training in quantitative methods. The good news is that results seem much more intuitive when shown graphically.

It is, for example, possible to play around with the results a bit and to use the probabilities of taking different jobs in a choice set (just use the above syntax in Stata) to explore several scenarios. Results from these types of exercises may be of particular interest to policy makers who often struggle with similar thought experiments when trying to find good recruitment and retention policies.

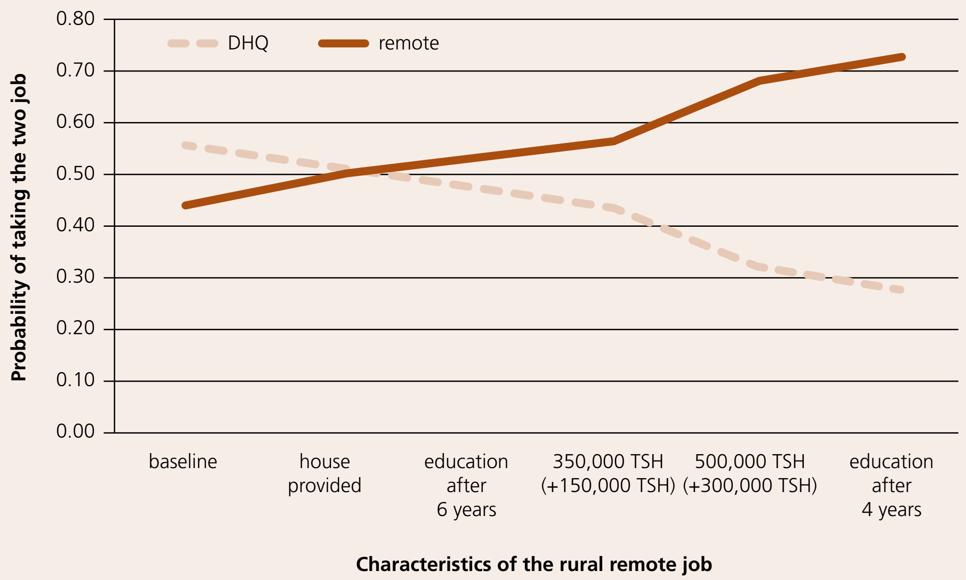
Figure 2.2 shows the varying probabilities of taking a rural remote job versus one in district headquarters, with various job conditions. It shows that the initial (baseline) probability of taking the rural remote job is 0.44, hence the probability of taking the job in the district headquarters is 0.56.

(In this baseline, both jobs have no education opportunities, bad infrastructure, a high workload, T Sh 200,000 per month salary, insufficient equipment and drugs, and no housing). The job in the district headquarters is thus preferred.

Different policies can make the rural remote job more attractive. If a house is provided with the remote job, the probability of taking the two jobs is equal (50:50). If opportunities for further education are provided after six years in the remote job, the probability of taking that job increases to 0.53 (so the remote job is preferred), and so on.

A simple graph can thus give a good impression of the impact of different policies, relative to each other. Similarly, it is also possible to experiment with bundles of policies.

Figure 2.2 Changes in probabilities of taking a job, rural remote versus district headquarters, as conditions in the rural remote job improve



DHQ = district headquarters.

2.6.6 Using DCE results for policy making

Ideally a DCE focusing on policy relevant issues should be developed with input from policy makers. The results should of course also be disseminated to policy makers so that they can make use of the information in developing their policies. However, the DCE presented in this case study was mainly developed as a PhD project, and there has been no formal cooperation with policy makers (unlike the Uganda case study).

Still, the main results have been written up and published in an academic journal, meaning that they at least reach academics and interested policy makers and practitioners. The main results were also presented and thoroughly discussed at a dissemination workshop in Tanzania in 2009, where policy makers and Tanzanian researchers participated. This year (2011) there will be a final dissemination workshop, for a larger project on motivation, availability and performance of human resources for health that this case study has been a part of. For that occasion short brochures summing up the main results and playing around with different kinds of graphical simulations will be developed and presented. These brochures will also be distributed to the training centers that took part in the study.

3. Retaining essential health personnel in underserved areas: a DCE case study in Uganda

SUMMARY OF SECTION 3

This section provides a case study of a DCE application in Uganda.

It explores preferences of in-service medical officers, nursing officers, and laboratory technicians practicing in rural health facilities.

As with the Tanzania case study (section 2), it highlights the importance of the qualitative component of the DCE methodology. In particular, it details the steps to identify job-posting attributes and levels that are important to health workers in the local setting. It also provides examples of how demographic questions can be used to supplement DCE information.

This case study uses focus group discussions to gather the necessary qualitative information for DCE. It uses Sawtooth software to generate a D-efficient design, and employs mixed logit for the econometric analysis of the data. It uses the regression equation estimate to look at the relative importance of job attributes, WTP, and uptake rates.

As with the earlier case study, it details important technical and logistical aspects of DCE, and sheds light on the challenges of collecting data. Finally, it describes the steps taken to analyze data, interpret results, and demonstrate the value of a DCE over qualitative data.

This section presents a case study of a DCE project that elicited the preferences for job postings among in-service health workers in medical, nursing, and laboratory cadres. The study was conducted during July and August 2010.

The section first presents background information and describes the motivation for the project. Next, it provides details for each of the stages of DCE described in the step-by-step guide (section 1): identification of attributes and levels; experimental design; data collection; data management and formation; and analysis and interpretation of data. Special attention is paid to aspects of the Uganda project that differed from the Tanzania project (section 2).

3.1 Background to study

The Uganda Ministry of Health (MOH) is facing challenges attracting and retaining health workers in rural facilities. In an effort to develop evidence-based policies to address this issue, the MOH took a leadership role on the DCE project described here. It was assisted by Capacity*Plus,* USAID's flagship global HRH project. As a first step, the MOH decided which health worker cadres were most important to target with attraction and retention policies: medical officers, nursing officers, pharmacists, and laboratory technicians. It selected these four cadres because they constitute the essential health worker team required to operate a health center in Uganda that can provide primary and secondary health care.

DCE data were collected in Uganda from both students in health worker training programs (to inform attraction policies) and in-service health workers (to inform retention policies). The process of collecting data from students and the results of subsequent analyses are described in Rockers et al. (2012). This case study focuses on DCEs conducted with in-service health workers. While DCE data were collected from students in all four cadres of interest, only in-service medical officers, nursing officers, and laboratory technicians were interviewed. The MOH decided that there were not enough in-service pharmacists to provide the necessary statistical power to include them in the study.

Many of the steps required to run a DCE are the same regardless of whether students or in-service workers are targeted. The primary difference is the fieldwork required to collect the DCE data, that is, students can often be interviewed at their school, simplifying the process. In-service health workers must often be interviewed at their place of work, which requires a process of finding and traveling to often-remote health facilities—a process that requires additional resources, including a team of trained interviewers, and travel logistics.

Different cadres of health workers have different preferences for job postings. For example, nurses may be more interested in working with supportive managers than are physicians, while physicians may be more interested in opportunities for professional advancement. Therefore, each cadre's DCE instrument should be developed separately and tailored to cadre-specific information collected in the local setting. This was done in the Uganda study.

However, many aspects of the process of designing and running a DCE are standard, regardless of the cadre under investigation, and so the following sections describe the process in general, and only refer to specific cadres when appropriate.

3.2 Identification of attributes and levels

The first step in a DCE focused on human resources for health is to identify the most important job posting attributes and the levels of those attributes appropriate for inclusion in the DCE instrument. For the Uganda study, this step was broken down into three parts: a review of the literature; discussions with policy makers, in this case key members of the Uganda MOH; and focus group discussions.

Literature review. WHO (2010) outlines financial and nonfinancial incentives that have previously been shown to be important to health workers when they are deciding where to work. Based on this report, and other recent research related to health worker attraction and retention (for example, Barnighausen and Bloom 2009; Blaauw et al. 2010; Kruk et al. 2010), an initial list of potentially important job attributes for retaining health workers in rural areas was compiled (box 3.1).

Box 3.1 Initial list of potentially important job attributes

- Salary
- Benefits, including allowances for travel, and children's education
- Housing accommodation
- Physical work environment, including quality of health facility infrastructure
- Reliability of equipment, supplies, and drugs
- Management environment, including support from managers
- Understaffing at facility and scope of responsibilities
- Scope of practice, that is, ability to practice a preferred type of medicine
- Tuition support for continued education
- Preferential entry to continued education programs
- Opportunities for in-service training
- Opportunities for dual practice, that is, working in the public and private sector
- Length of contract

Discussions with policy makers. The list in box 3.1 was presented by the project team (comprising staff from Capacity*Plus* and the MOH) to key decision makers within the MOH. Discussions were held with regard to which attributes were most relevant in the Uganda context, focusing on which attributes were most likely to reflect feasible policy options to address the issue of health worker attraction and retention. Based on these discussions, the project team decided that preferential entry to continued education programs was not a viable policy option, as the MOH did not have authority to grant preferential entry. The project team and MOH decision makers agreed that all other attributes on the list were appropriate and feasible policy options in Uganda.

Focus group discussions. The project team then brought the revised list of potential job attributes to focus group discussions (FGDs) to determine the final list of attributes and levels for inclusion in the DCE instruments. (This contrasts with the Tanzania study, which used key interviews to determine the final DCE attributes and levels.) In Uganda, FGDs were held with representatives from each cadre of interest. FGD participants were ineligible to participate in the DCE. All FGD participants gave written consent to participating. The project manager, trained in focus group methods, facilitated the FGDs. All FGDs were conducted in English, and lasted approximately one hour (box 3.2).

Box 3.2 Summary of focus group discussions

Participants in focus groups were first asked general questions about their perceptions of health worker job postings in Uganda, and their experiences living in rural areas and working in the country's health sector.

Next, participants discussed factors that they considered important when thinking about where they prefer to work. They were asked specific questions about job attributes that were previously identified as potentially important. Then participants were asked to individually rank these attributes according to which they felt was most important to them.

Finally, they nominated a final set of attributes. For each attribute, they were asked to identify levels that were realistic and appropriate in the local context. For example, when discussing salary, participants were asked to identify what they deemed to be a realistic and fair salary.

Based on information collected during FGDs, six attributes were included in the final DCE instrument for each of the cadres of interest. Five of six attributes were the same for each of the three DCEs (table 3.1): salary, quality of the health facility (with a focus on quality of equipment for laboratory technicians), housing, length of time committed to the job posting, and manager support. Levels for the salary attribute differed for each of the three instruments. DCE scenarios for medical officers and laboratory technicians had tuition for future schooling as the final attribute; for nursing officers they had health facility staffing.

Table 3.1 Attributes and levels for DCEs in Uganda study

	Health worker cadre		
	Medical	**Nursing**	**Laboratory**
Attribute 1	**Salary (per mo.)**	**Salary (per mo.)**	**Salary (per mo.)**
Level 1[1]	700,000	450,000	400,000
Level 2	1,000,000	550,000	500,000
Level 3	1,500,000	650,000	600,000
Level 4	2,000,000	750,000	700,000

	Health worker cadre		
	Medical	**Nursing**	**Laboratory**
Attribute 2	**Facility quality**	**Facility quality**	**Facility quality**
Level 1	Basic[2]	Basic[2]	Basic[4]
Level 2	Advanced[3]	Advanced[3]	Advanced[5]
Attribute 3	**Housing**	**Housing**	**Housing**
Level 1	None	None	None
Level 2	Allowance	Allowance	Allowance
Level 3	Provided	Provided	Provided
Attribute 4	**Length of commitment**	**Length of commitment**	**Length of commitment**
Level 1	2 years	2 years	2 years
Level 2	5 years	5 years	5 years
Attribute 5	**Support from manager**	**Support from manager**	**Support from manager**
Level 1	Not supportive	Not supportive	Not supportive
Level 2	Supportive	Supportive	Supportive
Attribute 6	**Future tuition**	**Facility staffing**	**Future tuition**
Level 1	No support[6]	50% understaffed	No support6
Level 2	Full support[7]	25% understaffed	Full support7
Level 3		Fully staffed	

All salary figures are in Ugandan shillings (U Sh). $1 = U Sh 2,350.
1 Level 1 salary figures represent base salary at the time of survey development for respective health worker cadres in Uganda
2 "Basic (e.g. unreliable electricity, equipment and drugs and supplies not always available)."
3 "Advanced (e.g. reliable electricity, equipment and drugs and supplies always available)."
4 "You have old equipment that often breaks down and an unreliable supply of drugs and reagents and gloves."
5 "You have equipment that always works and a reliable supply of drugs and reagents and gloves."
6 "The government will not provide any financial assistance for a study program after your commitment is over."
7 "The government will pay your full tuition for a study program (e.g. specialty training) after your commitment is over."

3.3 Experimental design and construction of choice sets

As discussed in sections 1 and 2, a full factorial design produces too many scenarios to fully investigate. For example, in Uganda the full factorial design for the medical officer DCE produced $4^1*3^1*2^4$ = 192 possible job profiles (1 attribute with 4 levels, 1 attribute with 3 levels, and 4 attributes with 2 levels), and (192*191)/2 = 18,336 possible choice sets. Experimental design methods were used to reduce the number of choice sets to a manageable level.

For this case study, a software package called Sawtooth Software (http://www.sawtoothsoftware.com/) was used to generate the DCE experimental design. It was also used to administer the DCE to health workers—that is, workers answered DCE questions on a computer. (In the Tanzania study, respondents answered DCE questions on paper. Sawtooth Software is a proprietary software package that, while not used widely for DCEs applications in the area of human resources for health research, can assist with certain aspects of a DCE project.) The primary advantages of computer administration are that data entry is simplified.

In Uganda, 10 blocks of 12 choice sets were administered to respondents from all cadres. The number of choice sets answered by each respondent, 12, was chosen based on the DCE literature and pilot data collected in Uganda. (As in Tanzania, all DCE instruments were designed to be D-efficient with no *a priori* assumptions about parameters, that is, betas were assumed to be zero.) The experimental designs for the cadres in Uganda were based on a utility function where salary was defined as a continuous variable, and all other attributes presented in table 3.1 were defined as dummy variables. All three DCEs in Uganda were generic (unlabeled). Further, an opt-out option was not included in any of the instruments, because it was assumed that all interviewed health workers would be in the market for a job posting within their current cadre. An example DCE choice set shown to medical officers is in figure 3.1.

Each block of 12 choice sets included one fixed choice set. That is, the experimental design was optimized to include 11 choice sets for each block, with each block also including one fixed choice set that was the same for all blocks. The fixed choice set was designed to test the internal validity of the instrument, and to test preferences for facility-level attributes (quality of facility infrastructure and support from facility management), because FGDs and previous research suggested a high preference for these attributes.

Figure 3.1 Example choice set for medical officers in Uganda 2010

Please tell us which of these job postings you prefer.

Choose by clicking one of the buttons below:

	Posting A	Posting B
Quality of the facility	Basic (e.g. unreliable electricity, equipment and drugs and supplies not always available)	Advanced (e.g. reliable electricity, equipment and drugs and supplies always available)
Housing	Housing allowance provided, enough to afford basic housing	No housing or allowance provided
Length of commitment	You are committed to this position for 2 years	You are committed to this position for 5 years
Study assistance	The government will not provide any financial assistance for a study program after your commitment is over	The government will pay your full tuition for a study program (e.g. specialty training) after your commitment is over
Salary	2,000,000 USh per month	1,000,000 USh per month
Management	2 years	2 years
	R120,000 per year	R120,000 per year
The workplace culture and style of management	The facility manager is not supportive and makes work more difficult	The facility manager is supportive and makes work easier
	☐	☐

In the fixed choice set, Job Posting A had poor facility quality and unsupportive management and good nonfinancial personal incentives (such as shorter time commitment and support for future tuition) while Job Posting B had good facility quality and supportive management and poor nonfinancial personal incentives (such as longer time commitment and no support for future tuition). Other attributes, including salary and housing, were held the same across the job postings in each choice set. Data from the fixed choice set were used to test the internal validity of the DCE instruments and are not included in the results of models presented in table 3.3.

3.3.1 Checking properties

Sawtooth provides options for the researcher to indicate level balance and minimum overlap, and in Uganda these options were selected to ensure an efficient design. However, it was still important to check that the designs produced by the software were indeed efficient in the manner that was expected, i.e. orthogonality, level balance, and minimum overlap. Because 120 different choices were used for each DCE in Uganda, this information is not easily included here in tables. (See section 2.3 above for a discussion of these checks in Tanzania.)

3.3.2 Supplemental questions

It is important to gather some non-DCE information from respondents, and the DCE methodology can easily be combined with a basic questionnaire that elicits personal information. In Uganda, information in three important domains was collected: demographics, including sex and age as well as previous experience living in a rural area; educational experiences, including prior training programs and perceptions of current training programs; and work experiences, including previous experiences working in rural and underserved areas.

Twenty supplemental questions were asked of respondents before they began the DCE. It is important to consider the length of time it may take respondents to answer these questions when designing the questionnaire, and to avoid respondent fatigue before they start the DCE questions.

3.4 Developing the questionnaire, pretesting, and data collection

3.4.1 Piloting questionnaire

Before data collection, all DCE instruments were pretested with health workers who were ineligible to be included in the main data collection. After pretest-participants finished the DCE, the project manager conducted debriefing interviews and took detailed notes. This debriefing information was then used to further refine the study instruments. Important information included the perceived realism of the framing and design of the DCE, the appropriateness of included attributes and levels, and any other difficulties that participants met when taking the DCE. After pretest-informed revisions were made, the DCE instruments for all cadres were finalized for formal data collection activities.

3.4.2 Sampling

As said, data were collected from in-service health workers from three cadres: medical officers, nursing officers, and laboratory technicians. These cadres, along with pharmacists, were identified by the Uganda MOH as the minimum-staffing requirement needed to effectively run a health facility in the country that provides primary and secondary services.

The primary sampling frame for in-service data was 10 rural districts in the Western region and 10 rural districts in the Northern region. Districts were chosen because they were identified as underserved (limited availability of health services and skilled health workers) by the MOH. Twenty districts were chosen to provide enough statistical power to accurately determine respondent preferences with DCE data.

In rural districts in Uganda, health center IV facilities are the primary employers of medical officers, nursing officers, and laboratory technicians. On average, each district has two health center IV facilities. For data collection activities, a team of interviewers visited all such facilities in sampled districts and invited all identified in-service health workers to participate in the study. Based on MOH health worker employment figures, it was calculated that 20 districts would yield data from approximately 50 medical officers, 150 nursing officers, and 100 laboratory technicians.

3.4.3 Training interviewers

A team of 12 interviewers was trained to administer the DCE to in-service health workers. The interviewers attended a one-day training session before the start of data collection activities.

The training focused on five areas: an introduction to the DCE project, to clarify the aims of the project so that interviewers were able explain those aims to respondents; an explanation of the data collection plan, including information on the logistics of team travel, survey administration, and data management; an introduction to the DCE instrument, to ensure that all interviewers were comfortable with the format of the DCE choice tasks as well as the supplemental questions so that they could answer any related questions respondents may have had; a review of the introductory script (a standard statement to be read to respondents just before beginning the survey); and an introduction to, and practice with, the computer technology that interviewers were to use in the DCE.

3.4.4 Travel to health facilities

Conducting a DCE with in-service health workers can include extensive travel. This issue was addressed in Uganda by splitting the team of interviewers into four groups of three interviewers, each assigned to cover a specific geographic area of five districts. Interviewer groups were each provided with a vehicle and assigned to visit all health facility IVs in their five districts. Again, these were the facilities identified as employing DCE-eligible health workers.

All team members had a rented laptop so they had the freedom to conduct interviewers with eligible participants at a pace independent from their teammates'. This also enabled each interviewer to be left at a facility alone to conduct interviews while others team members were taken to nearby facilities.

3.4.5 Survey administration

Data collection activities were conducted during August 2010. Surveys were administered on computers using Sawtooth Software's SSI Web CAPI program. Before beginning the DCE, all respondents were read an introductory script by project personnel. The purpose of the script was to acclimate respondents to the hypothetical nature of the DCE they were about to take. Respondents were asked to imagine making a real choice, take into account only the attributes described, and instructed that there were no right or wrong answers.

Respondents then completed the survey questionnaire and the DCE at their own pace on a laptop computer, taking about 30 minutes on average. All respondents provided consent before participating in the study. All questions and DCE scenarios were presented in English.

3.5 Data input

The DCE data collected was prepared for analysis. A section of the dataset for data collected from medical officers is in table 3.2.

Table 3.2 Example DCE dataset for Uganda data

personid	obsid	alt	cno	choiceset	salary	qual	house1	house2	house3	commit	tuition	mgmt	const	choice
1	1	1	1	1	1,000,000	0	0	1	0	1	1	1	1	0
1	1	2	1	1	1,500,000	1	1	0	0	0	0	0	0	1
1	2	1	2	2	1,500,000	1	0	0	1	1	1	0	1	1
1	2	2	2	2	2,000,000	0	0	1	0	0	0	1	0	0
1	3	1	3	3	1,000,000	1	0	1	0	0	0	0	1	0
1	3	2	3	3	700,000	0	1	0	0	1	1	1	0	1
1	4	1	4	4	2,000,000	1	1	0	0	0	1	1	1	0
1	4	2	4	4	700,000	0	0	0	1	1	0	0	0	1
1	5	1	5	5	1,500,000	0	0	1	0	0	1	0	1	0
1	5	2	5	5	1,000,000	1	0	0	1	1	0	1	0	1

personid	obsid	alt	cno	choiceset	salary	qual	house1	house2	house3	commit	tuition	mgmt	const	choice
1	6	1	6	6	700,000	1	0	1	0	1	1	0	1	1
1	6	2	6	6	1,000,000	0	1	0	0	0	0	1	0	0
1	7	1	7	7	2,000,000	0	0	0	1	0	1	0	1	1
1	7	2	7	7	1,500,000	1	0	1	0	1	0	1	0	0
1	8	1	8	8	700,000	1	0	0	1	0	1	1	1	0
1	8	2	8	8	2,000,000	0	1	0	0	1	0	0	0	1
1	9	1	9	9	1,000,000	1	0	0	1	0	0	0	1	1
1	9	2	9	9	1,500,000	0	1	0	0	1	1	1	0	0
1	10	1	10	10	2,000,000	1	0	1	0	0	0	1	1	1
1	10	2	10	10	1,000,000	0	0	0	1	1	1	0	0	0
1	11	1	11	11	700,000	0	0	1	0	0	0	0	1	1
1	11	2	11	11	1,500,000	1	1	0	0	1	1	1	0	0
1	12	1	12	12	700,000	1	1	0	0	0	1	1	1	1
1	12	2	12	12	1,500,000	0	0	0	1	1	0	0	0	0
2	13	1	1	13	700,000	0	0	1	0	0	0	1	1	0
2	13	2	1	13	2,000,000	1	0	0	1	1	1	0	0	1
2	14	1	2	14	1,000,000	1	0	0	1	0	0	0	1	0
2	14	2	2	14	1,500,000	0	1	0	0	1	1	1	0	1
2	15	1	3	15	700,000	0	0	0	1	1	1	1	1	1
2	15	2	3	15	1,500,000	1	0	1	0	0	0	0	0	0
2	16	1	4	16	1,000,000	0	0	1	0	0	1	0	1	1
2	16	2	4	16	2,000,000	1	1	0	0	1	0	1	0	0
2	17	1	5	17	1,500,000	1	0	0	1	0	1	1	1	0
2	17	2	5	17	700,000	0	1	0	0	1	0	0	0	1
2	18	1	6	18	2,000,000	0	1	0	0	0	1	0	1	0
2	18	2	6	18	1,000,000	1	0	1	0	1	0	1	0	1
2	19	1	7	19	700,000	1	0	1	0	1	1	0	1	1
2	19	2	7	19	1,000,000	0	0	0	1	0	0	1	0	0
2	20	1	8	20	1,500,000	1	1	0	0	0	0	0	1	1
2	20	2	8	20	2,000,000	0	0	1	0	1	1	1	0	0
2	21	1	9	21	700,000	0	0	0	1	1	0	0	1	0
2	21	2	9	21	1,000,000	1	1	0	0	0	1	1	0	1
2	22	1	10	22	2,000,000	0	0	1	0	0	0	1	1	0
2	22	2	10	22	1,500,000	1	0	0	1	1	1	0	0	1
2	23	1	11	23	1,500,000	1	1	0	0	0	0	0	1	1
2	23	2	11	23	700,000	0	0	1	0	1	1	1	0	0
2	24	1	12	24	2,000,000	1	0	1	0	1	1	0	1	1
2	24	2	12	24	1,000,000	0	0	0	1	0	0	1	0	0
3	25	1	1	25	700,000	1	1	0	0	0	1	1	1	1
…	…	…			…	…	…	…	…	…	…	…	…	…

The variables are:

personid: respondent identifier. Each respondent has 24 rows of data, because each respondent answered 12 choice sets with 2 job posting scenarios each, for a total of 24 scenarios.

obsid: identifier for each choice set. Each choice set has 2 rows of data, because each choice set is comprised of 2 scenarios.

alt: indicates the alternative within each choice set. Given each option had 2 choices, alt takes on the value of 1 or 2.

cno: represents the choice number in the DCE questionnaire. As each respondent made 12 choices, cno will range from 1 to 12.

choiceset: since the design consists of 120 different choice sets, there is a variable indicating which of the 120 choice sets are being observed.

salary: the level of the salary attribute represented in the scenario for a given row of data. For medical officers, the salary attribute took on 4 levels: U Sh 700,000 per month; U Sh 1,000,000 per month; U Sh 1,500,000 per month; and U Sh 2,000,000 per month (table 3.1, above).

qual: the level of the facility quality attribute represented in the scenario for a given row of data, converted to a dummy variable. If the level represented in the presented scenario was "Basic (e.g. unreliable electricity, equipment and drugs and supplies not always available)", then the value of the QUAL variable for that row of data will be "0", to represent that level. If, on the other hand, the level represented in the presented scenario was "Advanced (e.g. reliable electricity, equipment and drugs and supplies always available)", then the value of the QUAL variable for that row of data will be "1", to represent that level.

house1; house2; house3: dummy variables representing the 3 levels of the housing attribute. If the level represented in the presented scenario was "no housing provided", then the value of the house1 variable for that row of data will be "1", and the value of the house2 and house3 variables for that row of data will be "0". If the level represented in the presented scenario was "housing allowance provided enough to afford basic housing", then the value of the house2 variable for that row of data will be "1", and the value of the house1 and house3 variables for that row of data will be "0". If the level represented in the presented scenario was "basic housing provided for you", then the value of the house3 variable for that row of data will be "1", and the value of the house1 and house2 variables for that row of data will be "0".

commit: the level of the length of commitment attribute represented in the scenario for a given row of data, converted to a dummy variable. If the level represented in the presented scenario was "You are committed to this position for 2 years", then the value of the commit variable for that row of data will be "0", to represent that level. If, on the other hand, the level represented in the presented scenario was "You are committed to this position for 5 years", then the value of the commit variable for that row of data will be "1", to represent that level.

tuition: the level of the future tuition attribute represented in the scenario for a given row of data, converted to a dummy variable. If the level represented in the presented scenario was "The government will not provide any financial assistance for a study program after your commitment is over", then the value of the tuition variable for that row of data will be "0", to represent that level. If, on the other hand, the level represented in the presented scenario was "The government will pay your full tuition for a study program (e.g. specialty training) after your commitment is over", then the value of the tuition variable for that row of data will be "1", to represent that level.

mgmt: the level of the support from manager attribute represented in the scenario for a given row of data, converted to a dummy variable. If the level represented in the presented scenario was "The facility manager is not supportive and makes work more difficult", then the value of the mgmt variable for that row of data will be "0", to represent that level. If, on the other hand, the level represented in the presented scenario was "The facility manager is supportive and makes work easier", then the value of the mgmt variable for that row of data will be "1", to represent that level.

const: an alternative-specific constant variable. Within each choice set, for the row of data that represents the first job scenario, const = 1 and for the row of data that represents the second job scenario, const = 0.
choice: the respondent's choice of scenario for a given choice set. For a given pair of rows of data represented by a single choice set, choice = 1 for one row and choice = 0 for the other row. This reflects the fact that the respondent selected a single scenario from the choice set as being their most preferred.

As respondents completed the DCE on a computer, no data entry was necessary. (The .csv file produced by Sawtooth is readable by most statistical analysis software.)

3.6 Model estimation and interpretation

Stata software was used to analyze the DCE data. Although Sawtooth has data analysis capabilities, Stata provides a more flexible platform and allows the user more control over specifying modeling assumptions. MXL models were fitted to DCE data from each of the three health worker cadres that were investigated. MXL models allow attribute coefficients to vary across respondents, accounting for preference heterogeneity and improving the realism of model assumptions. Second, MXL models adjust the standard errors of utility estimates to account for repeated choices by the same individual.

All models were main effects—no interaction terms were included. All attribute variables were specified as having a random component except for salary, which was specified as fixed in all models. While random specifications of salary may improve model fit, a fixed coefficient ensures that the estimate of salary utility has the right sign and is preferred for calculation and interpretation of willingness to pay, as it avoids possible problems with dividing distributions on distributions. The constant variable was also specified as fixed. Further, all attribute variables were coded as dummy variables except for salary, which was specified as continuous in all models. All model coefficients were assumed to be normally distributed.

For the analysis of medical officers' data, the following Stata code was used:

mixlogit choice salary const, group(obsid) id(personid) rand(qual house2 house3 commit tuition mgmt) nrep(500)

where salary and const, as fixed variables, are located before the comma, and the other attribute variables, modeled as having a random component, are included within the rand() option. The nrep() command indicates the number of Halton draws the model will run; a value of 500 here ensures a robust output. MXL models employ simulation based estimation techniques, and the number of Halton draws indicates the number of unique times the MXL simulation is run. It is important to run the simulations enough times so that the model converges and model estimates stabilize, that is, simulation-induced variance is minimized.

The MXL command is not yet part of Stata, although it can be installed directly onto Stata using the following command:

ssc install mixlogit, replace

This will go to the internet, find the program and install it automatically.

The regression results and calculated WTP values are in table 3.3.

Table 3.3 Utility and willingness-to-pay estimates for job attributes, Uganda, 2010

	Health worker cadre					
	Medical officer		Nursing officer		Laboratory technician	
	ß (SE)	WTP[1] (95% CI)	ß (SE)	WTP (95% CI)	ß (SE)	WTP (95% CI)
salary (cont. x 500,000 U Sh/month)	0.61 (0.12)***	–	2.92 (0.38)***	-	1.66 (0.41)***	–
quality of facility: advanced (ref: basic)	1.00 (0.29)***	819,287 (356,587; 1,281,976)	1.20 (0.19)***	205,167 (136,834; 273,500)	1.70 (0.29)***	516,972 (254,949; 778,996)
housing: allowance (ref: none)	0.60 (0.25)**	488,841 (89,284; 888,397)	0.84 (0.18)***	144,232 (83,087; 205,376)	0.34 (0.21)	103,461 (-28,186; 235,108)
housing: provided (ref: none)	0.76 (0.25)***	618,488 (223,495; 1,013,482)	0.74 (0.17)***	127,507 (67,920; 187,095)	0.30 (0.21)	88,151 (-37,405; 213,707)
commitment: 2 years (ref: 5-years)	0.27 (0.18)	223,572 (-64,893; 512,038)	0.17 (0.13)	28,419 (-14,662; 71,500)	0.15 (0.16)	47,561 (-43,882; 139,004)
manager: supportive (ref: not supportive)	0.48 (0.19)***	393,089 (80,005; 706,172)	1.64 (0.23)***	281,172 (200,909; 361,436)	1.01 (0.20)***	299,633 (139,781; 459,485)
full tuition support for training (ref: no support)	0.65 (0.25)***	532,752 (107,601; 957,903)	–	–	1.12 (0.26)***	324,669 (134,117; 515,221)
staffing: 25% understaffed (ref: 50% understaffed)	–	–	0.05 (0.16)	9,025 (-43,236; 61,286)	–	–
staffing: Fully staffed (ref: 50% understaffed)	–	–	0.09 (0.16)	16,122 (-36,630; 68,874)	–	–
constant	0.22 (0.17)	–	0.11 (0.12)	–	0.17 (0.15)	–
Number of respondents	39		74		45	
Number of observations	858		1,628		990	
Log likelihood	227.2835		384.0762		228.9870	
Pseudo R2	0.2357		0.3193		0.3326	

[1] WTP presented in Ugandan shillings per month; $1 = U Sh 2,350.
*p < 0.10, **p < 0.05, ***p < 0.01

In choosing health sector job postings, respondents in all cadres expressed strong preference for increasing salary levels. This is evidenced by the ß estimates for the salary attribute in table 3.3. Better quality health facilities and supportive facility managers were important to all cadres in determining where they would prefer to work. While housing support was important to medical officers and nursing officers—there is a significant increase in utility for both a housing allowance and provided housing compared with the reference of no housing or allowance—respondents in these cadres did not express a clear preference for provided housing compared with receiving a housing allowance. In other words, the utility increase from no housing to a housing allowance is the same as the utility increase from no housing to provided housing. Finally, medical and laboratory cadres both expressed a strong preference for full tuition support for a future training program.

3.6.1 Willingness to pay

Willingness to pay estimates and 95% confidence intervals are presented in table 3.3 above. The confidence intervals were calculated with Stata's nlcom command (section 2.6 above). For example, to calculate the 95% confidence intervals for WTP for an advanced quality facility, the code used was:

```
nlcom (_b[quality])/(-_b[salary])
```

There is currently some debate regarding the appropriateness of calculating WTP estimates from MXL models of DCE data. Of particular concern are the assumptions that MXL models require regarding the distribution of the price variable (in the model above, the salary variable). Namely, by specifying the price variable as fixed, as done above, here it is assumed that all individuals have the same preference for salary, which may be unreasonable. However, it may be equally unreasonable to assume that the distribution of preferences for salary is normally distributed. Several alternative techniques have been suggested to address this issue (Hole and Kolstad 2010). However, no "gold standard" has been established. These concerns should be considered when using WTP estimates derived from MXL models of DCE data.

One can use the delta method employed by nlcom with the example MXL model presented above because the salary variable has been modeled as fixed. However, if the salary variable is modeled as having a random component, i.e. within rand() in the mixlogit statement, it would be inappropriate to estimate WTP confidence intervals with parametric procedures such as nlcom.

Rather, a non-parametric bootstrapping procedure would be required. In Stata, one would need to use the bootstrap command, with the syntax:

```
bootstrap wtp_qual=(_b[qual]/(-_b[salary))
wtp_commit=(_b[commit]/(-_b[salary))
wtp_tuition=(_b[tuition]/(-_b[salary))
wtp_house2=(_b[house2]/(-_b[salary]))
wtp_house3=(_b[house3]/(-_b[salary]))
wtp_mgmt=(_b[mgmt]/(-_b[salary]))
, cluster(personid) idcluster(newid) group(obsid) seed (123) reps(500)
: mixlogit choice salary const, group(obsid) id(newid) rand(qual
commit tuition house2 house3 mgmt) nrep(500)
```

where most of the variables are defined above. cluster() is the adjusted standard errors for repeated observations on the individual. So this is a panel identifier.

idcluster(newid), creates a unique identifier for each of the selected clusters. Thus if some panels were selected more than once, the temporary variable newid would assign a different ID number to each resampled panel.

reps() is the number of replications in the bootstrap, here the panel is sampled from 500 times. The seed command would allow the results to be replicated.

The mixlogit command has changed when bootstrapping, with the id(variable) changing from personid to newid. This bootstrapping procedure can be time intensive. A run with ~1,000 observations can take up to 12 hours to complete.

3.6.2 Proportion of population with positive effect

Assuming a normal distribution for random parameters, MXL models provide output that can be used to calculate the proportion of respondents for whom a job attribute has a positive or negative effect on preference for a job scenario. Table 3.4 provides the same ß values as were presented in table 3.3, and presents estimates of the standard deviation of the ß estimates.

The standard deviation output from a mixed logit model is different from the standard error estimates in table 3.4. Standard deviations indicate preference heterogeneity, while standard errors indicate estimate uncertainty. The proportion of the respondent population that has a positive preference for a job attribute can be calculated with the following equation:

Proportion positive = $\Phi(ß/SD)$

where Φ is the standard normal cumulative distribution function and ß is positive. If ß is negative, the above equation will estimate the proportion of the respondent population that has a negative preference for the job attribute.

Table 3.4 Mixed logit estimates and standard deviations with calculated proportions of positive effect for job attributes

	Health worker cadre								
	Medical officer			Nursing officer			Laboratory technician		
	ß	SD	% Pos.[1]	ß	SD	% Pos.	ß	SD	% Pos.
salary (cont. x 500,000 U Sh/month)	0.56	–	–	2.89	–	–	0.83	–	–
quality of facility: advanced (ref: basic)	0.92	1.19	78.0	1.23	1.06	87.7	1.70	1.32	90.1
housing: allowance (ref: none)	0.63	0.61	84.9	0.83	0.50	95.2	0.34	0.19	96.3
housing: provided (ref: none)	0.68	0.52	90.5	0.75	0.38	97.6	0.30	0.15	97.7
commitment: 2-years (ref: 5-years)	0.26	0.39	74.8	0.19	0.52	64.3	0.15	0.45	63.1
manager: supportive (ref: not supportive)	0.49	0.56	80.9	1.51	1.22	89.2	1.01	0.62	94.8
full tuition support for training (ref: no support)	0.70	1.10	73.8	-	-	-	1.12	1.08	85.0
staffing: 25% understaffed (ref: 50% understaffed)	–	–	–	0.04	0.05	78.8	–	–	–
staffing: Fully staffed (ref: 50% understaffed)	–	–	–	0.07	0.17	66.0	–	–	–
Number of respondents	39			74			45		
Number of observations	858			1,628			990		

[1] The proportion of the respondent population that has a positive preference for the job attribute.

3.6.3 Uptake rate

Potential job-posting uptake was simulated, given specific health system reforms, using preferences estimated by the main effects mixed logit models (table 3.3). (An explanation of the formulas used to calculate simulated preferences is in section 2.6, above.)

In the simulations, job posting was compared with improved attributes (as a result, for example, of reform efforts) to a baseline job posting with attributes found in rural Uganda today. The results of selected simulations are in table 3.5.

Table 3.5 Simulated preferences for job posting under various potential policy scenarios (confidence intervals in parentheses)

Improved job posting[1]			
Simulated scenario	**Medical**	**Nursing**	**Laboratory**
30% increase in salary	55% (53%–56%)	71% (66%–75%)	62% (56%–68%)
50% increase in salary	58% (55%–60%)	81% (75%–86%)	70% (61%–77%)
Advanced facility quality	73% (61%–83%)	77% (70%–83%)	85% (76%–91%)
Supportive facility manager	62% (53%–70%)	84% (77%–89%)	73% (65%–80%
Full tuition support	66% (54%–76%)	–	75% (65%–84%)
Full staffing	–	52% (44%–60)	–

[1] Compared with a baseline job posting defined as: current salary (medical: U Sh 700,000; nursing: U Sh 450,000; laboratory: U Sh 400,000); basic facility quality; housing allowance provided; length of commitment 5 years; facility manager not supportive; no future tuition support (for medical and labouratory); 50% understaffed (for nurses).

These results confirm the importance of health facility quality and manager support in retaining health workers to job postings in underserved areas.

3.6.4 Comparison of results: DCE versus FGDs

The DCE results share important similarities with information collected during preliminary FGDs. In both the DCE and the FGDs, health workers from all three cadres showed a strong preference for job postings with advanced health facility quality and a supportive manager. This suggests that the DCE results have strong criterion validity. However, the DCE results provide additional information that FGDs cannot provide.

Namely, with DCE data it was possible to estimate respondents' strength of preference and WTP for each job attribute (though it was not possible to determine how important each attribute was relative to other attributes in FGDs). The relative preference and WTP information that DCEs provide are very helpful for policy makers trying to determine the most effective policy.

Further, with DCE estimates it was possible to simulate job uptake, given various policy actions (table 3.5). This is not possible with FGD information. Estimates of job uptake can provide decision makers with a clear measure of the potential benefits of alternative policies under consideration during policy making.

3.6.5 Presentation of results for policy

The results in table 3.3 may be difficult to understand for policy makers. Figure 3.2 is a graphic representation of the WTP results for medical officers. Figure 3.3 and 3.4 are similar representations for nursing officers and laboratory technicians.

Figure 3.2 WTP estimates and 95% confidence intervals for job attributes for medical officers, Uganda

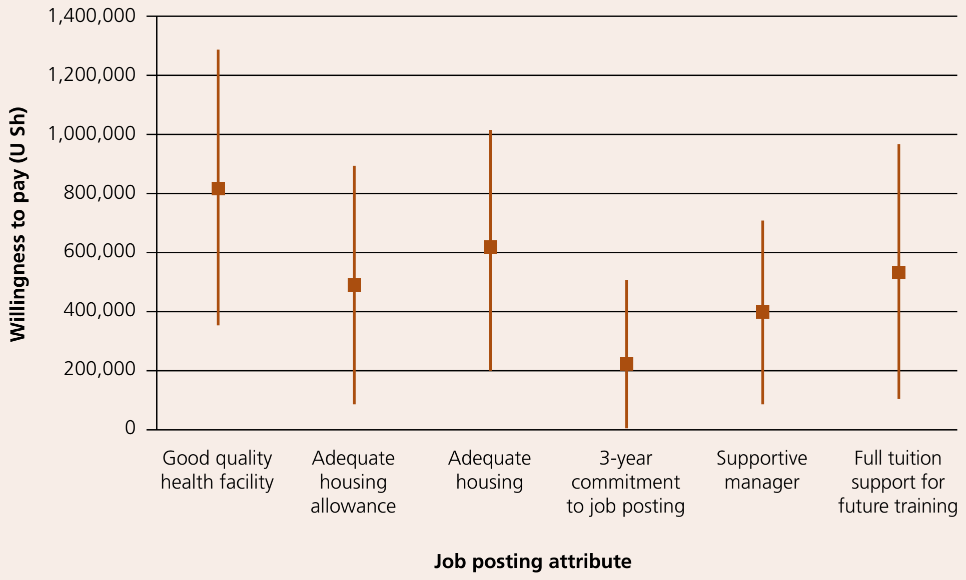


Figure 3.2 shows that medical officers were willing to pay more than U Sh 800,000 in monthly salary for a good-quality health facility. In other words, these respondents were willing to accept U Sh 800,000 per month less in salary for a posting at a facility with good quality compared with a posting at a facility with poor quality. Similarly, medical officers were willing to accept nearly U Sh 450,000 per month less in salary for a posting at a facility with a supportive manager compared with a posting at a facility with an unsupportive manager.

Figure 3.3 WTP estimates and 95% confidence intervals for job attributes for nursing officers, Uganda

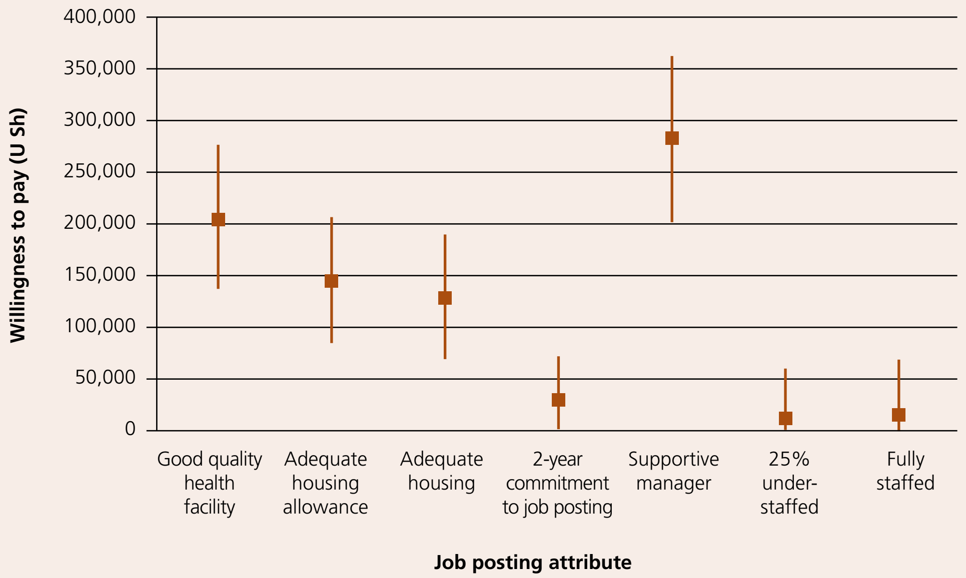


Figure 3.3 shows that nursing officers were willing to accept U Sh 200,000 per month less in salary for a posting at a facility of good quality compared with a posting at a facility of poor quality. Nursing officers most preferred health facilities with a supportive manager. Indeed, nursing officers were willing to accept U Sh 250,000 per month less in salary for a posting at a facility with a supportive manager compared with a posting at a facility with an unsupportive manager.

Figure 3.4 WTP estimates and 95% confidence intervals for job attributes for laboratory technicians, Uganda

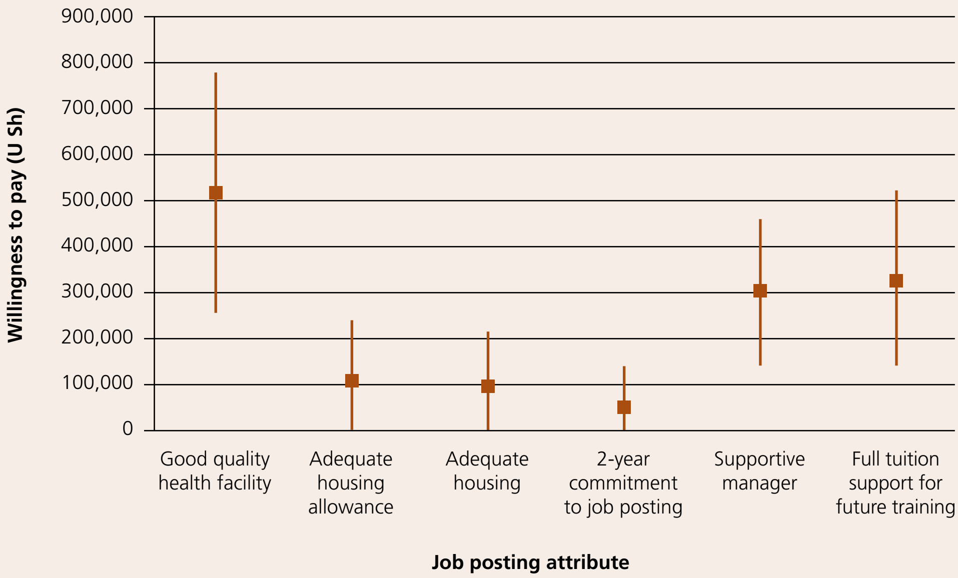


Figure 3.4 shows that laboratory technicians demonstrated a substantial preference for job postings at good-quality health facilities. They were willing to accept U Sh 500,000 per month less in salary for a posting at a facility of good quality compared with a posting at a facility of poor quality. Laboratory technicians were also willing to accept U Sh 300,000 per month less in salary for a posting at a facility with a supportive manager compared with a posting at a facility with an unsupportive manager.

3.6.6 Using DCE results for policy making

The DCE results presented here were used to inform human resources for health policy making in Uganda. Results were presented to the country's HRH Technical Working Group (TWG), comprising representatives from the MOH, Ministry of Finance, international donor community, local faith-based organizations, universities, and other nongovernmental organizations.

The DCE results were translated from a purely statistical presentation into one that could be more easily interpreted by laypersons. This is important because many policy makers are not researchers by background. The data were presented to the stakeholders in tables showing various packages of retention strategies, each in combinations of three or four retention interventions (DCE attributes) with corresponding estimates of uptake rates. The information derived from the DCE helped the TWG to decide which strategies would potentially be most effective in attracting and retaining health workers in rural areas.

Following the DCE the Ministry of Health, with technical assistance from Capacity*Plus*, conducted a costing exercise to understand the financial implications of each preferred package of retention interventions and to more clearly gauge feasibility within available fiscal space. For the TWG, the strategy packages were categorized according to potential uptake and costs of implementation.

As of this writing, it appears that the DCE results and the costing exercise helped in MOH policy making and that the MOH could be interested in pursuing one of the retention strategy packages preferred by health workers.

References

General

WHO (World Health Organization). 2006. *World Health Report: Working Together For Health*. WHO, Geneva, Switzerland.

WHO (World Health Organization). 2010. *Increasing Access to Health Workers in Remote and Rural Areas through Improved Retention: Global Policy Recommendations*. Geneva: WHO, Geneva, Switzerland. http://www.who.int/hrh/retention/guidelines/en/index.html. Accessed 18 April 2011.

General DCE

Bennett J., and E. Birol. 2010. *Choice Experiments in Developing Countries*. Edward Elgar, Cheltenham.

Blaauw, D., and M. Lagarde, 2010. "Overview of the use of discrete choice experiments in HRH research". Presentation made at the meeting:"Tools for implementing Rural Retention Strategies: Towards a 'How To Guide for 'Discrete Choice Experiments'– A Methods Workshop", World Health Organization, Geneva, Switzerland, October 2010. Meeting report available online at: http://www.who.int/hrh/resources/DCE_report.pdf

Bridges, J., B. Hauber, D. Marshall A. Lloyd L. Prosser D. Regier R. Johnson, and J. Mauskopf. Forthcoming. 2011 "Conjoint Analysis Applications in Health—A Checklist: A Report of the ISPOR Good Research Practices for Conjoint Analysis Task Force." *Value in Health*, 14:403-413.

de Bekker-Grob, EW., Ryan, M. & Gerard, K. 2012. "Discrete choice experiments in health economics: a review of the literature". *Health Economics*. 21: 145-172.

Hanley, H., S. Mourato, and R. Wright. 2001. "Choice Modelling Approaches: A Superior Alternative for Environmental Valuation?" *Journal of Economic Surveys* 15(3): 435-462.

Hensher, D.A., J.M. Rose, and W.H. Greene. 2005. *Applied Choice Analysis: A Primer*. Cambridge, United Kingdom: Cambridge University Press.

Hole, A. 2007. "A Comparison of Approaches to Estimating Confidence Intervals for Willingness to Pay Measures." *Health Economics* 16: 827-840.

Hole, R.A. and J.R. Kolstad. 2010. "Mixed Logit Estimation of Willingness to Pay Distributions: A Comparison of Models in Preference and WTP Space using Data from a Health-related Choice Experiment." University of Bergen Working Paper in Economics. http://www.uib.no/filearchive/wp-03.10_2.pdf

Lancsar, E., and J. Louviere. 2006. "Deleting 'Irrational' Responses from Discrete Choice Experiments: A Case of Investigating or Imposing Preferences?" *Health Economics* 15: 797-811.

Lancsar, E., and J. Louviere. 2008. "Conducting Discrete Choice Experiments to Inform Healthcare Decision Making. A User's Guide." *Pharmacoeconomics* 26 (8): 661-677.

Louviere, J., D. Hensher, and J. Swait. 2000. *Stated Choice Methods: Analysis and Application*. First edition. Cambridge, United Kingdom: Cambridge University Press.

Ryan, M., K. Gerard, and M. Amaya-Amaya (eds.) 2008. *Using Discrete Choice Experiments to Value Health and Health Care*. Dordrecht, The Netherlands: Springer.

Train, K. E. 2009. *Discrete Choice Methods with Simulation*. Second edition. Cambridge, United Kingdom: Cambridge University Press.

Applications of DCEs to human resources in health: low- and middle-income countries

Blaauw, D, E. Erasmus, N. Pagaiya, V. Tangcharoensathein, K. Mullei, S. Mudhune, C. Goodman, M. English, and M. Lagarde. 2010. "Policy Interventions that Attract Nurses to Rural Areas: A Multicountry Discrete Choice Experiment." *Bulletin of the World Health Organization* 88(5) 350-356.

Chomitz, K.M., G. Setiadi, A. Azwar, and N. Widiyarti. 1998. "What do Doctors Want? Developing Incentives for Doctors to Serve in Indonesia's Rural and Remote Areas." Policy Research Working Paper 1888. World Bank, Washington, DC; 1998.

Hanson K, Jack W. 2010. Incentives could induce Ethiopian doctors and nurses to work in rural settings. *Health Affairs*; 29; 1452-1460.

Jaskiewicz W, Phathammavong O, Vangkonevilay P, Paphassarang C, Phachanh IT, and Wurts L. 2012. "Toward Development of a Rural Retention Strategy in Lao People's Democratic Republic: Understanding Health Worker Preferences". Washington, DC: Capacity*Plus.*

Kolstad, J.R. 2011. "How to Make Rural Jobs More Attractive to Health Workers: Findings from a Discrete Choice Experiment in Tanzania." *Health Economics* 20(2): 196-211.

Kruk, M, J. Johnson, M. Gyakobo, P. Agyie-Baffour, K. Asabir, S.R. Kotha, J. Kwansah, E. Nakua, R.C. Snow, and M. Dzodzomenyo. 2010. "Rural Practice Preferences among Medical Students in Ghana: A Discrete Choice Experiment." *Bulletin of the World Health Organization* 88(5) 333-341.

Lagarde M., and D. Blaauw. 2009. "A Review of the Application and Contribution of Discrete Choice Experiments to Inform Human Resources Policy Interventions." *Human Resources for Health* 7(62).

Mangham, L.J., and K. Hanson. 2008. "Employment Preferences of Public Sector Nurses in Malawi: Results from a Discrete Choice Experiment." *Tropical Medicine and International Health* 13: 1433-1441.

Mangham, L.J., K. Hanson, and B. McPake. 2009. "How to Do (or Not to Do)… Designing a Discrete Choice Experiment for application in a low income country." *Health Policy and Planning* 24: 151-158

McAuliffe E, Revill P, Kamwendo F, Sidat M, Masanja H, Pinho H, and Araújo E. (forthcoming). "Factors influencing Job Preferences of Mid-Level Healthcare Providers in Sub-Saharan Africa: Results from Large Sample Discrete Choice Experiments in Malawi, Mozambique and Tanzania". Centre for Global Health, University of Dublin, Ireland, Dublin.

Penn-Kekana, L., D. Blaauw, K. San Tint, D. Monareng, and J. Chege. 2005. "Nursing Staff Dynamics and Implications for Maternal Health Provision in Public Health Facilities in the Context of HIV/AIDS." Centre for Health Policy (South Africa) and Frontiers in Reproductive Health (South Africa), Population Council.

Rao KD, Shroff Z, Ramani S, Khandpur N, Murthy S, Hazarika I, Choksi M, Ryan M, Berman P, and Vujicic M. (forthcoming). "How to Attract Health Workers to Rural Areas in India? Evidence from a Discrete Choice Experiment". Health, Nutrition and Population Discussion paper, World Bank, Washington/DC.

Rockers PC, Jaskiewicz W, Wurts L, Kruk ME, Mgomella G, Ntalazi F, Tulenko K. Preferences for working in rural clinics among trainee health professionals in Uganda: a discrete choice experiment. (2012). *BMC Health Services Research* 2012, **12**: 212.

Vujicic, M., M. Alfano, M. Ryan, C. Sanford Wesseh, and J. Brown-Annan. 2010a. "Policy Options to Attract Nurses to Rural Liberia. Evidence from a Discrete Choice Experiment." Health, Nutrition and Population Discussion paper, World Bank, Washington DC.

Vujicic M, Alfano M, Shengelia B, and Witter S. 2010b. "Attracting doctors and medical students to rural Vietnam: Insights from a discrete choice experiment." Health, Nutrition and Population Discussion paper, World Bank, Washington/DC.

Applications of DCEs to human resources in health: developed countries

Gosden, T., I. Bowler, and M. Sutton. 2000. "How do General Practitioners Choose their Practice? Preferences for Practice and Job Characteristics." *Journal for Health Service Research Policy* 5: 208-213.

Scott, A. 2001. "Eliciting GP's Preferences for Pecuniary and Non-pecuniary Job Characteristics." *Journal of Health Economics* 20: 329-347.

Scott, A, C. Bond, J. Inch, and A. Grant. 2007. "Preferences of Community Pharmacists for Extended Roles in Primary Care: A Survey and Discrete Choice Experiment." *Pharmacoeconomics* 25(9): 783-792.

Sivey, P.M., A. Scott, J. Witt, J. Humphreys, and C. Joyce. 2010. "Why Junior Doctors don't Want to Become General Practitioners: A Discrete Choice Experiment from the MABEL Longitudinal Study of Doctors (October 6). Melbourne Institute Working Paper 17/10. SSRN: http://ssm.com/abstracts=1688667.

Ubach, C., A. Scott, F. French, M. Awramenko, and G. Needham. 2003. "What do Hospital Consultants Value about their Jobs? A Discrete Choice Experiment." *British Medical Journal* 326: 1432-1437.

Wordsworth, S., D. Skatun, A. Scott, and F. French. 2004. "Preferences for General Practice Jobs: A Survey of Principals and Sessional GPs." *British Journal of General Practitioners* 54: 740-746.

Qualitative work to derive attributes and investigate responses

Coast, J., and S. Horrocks. 2007. "Developing Attributes and Levels for Discrete Choice Experiments using Qualitative Methods." *Journal of Health Services Research Policy* 12: 25-30.

Lievens T., M. Lindelow, and P. Serneels. 2009. "Understanding Health Workforce Issues: A Selective Guide to the Use of Qualitative Methods," in M. DalPoz and A. Soucat. *Handbook on Monitoring and Evaluation of Human Resources for Health*. Geneva: World Health Organization and Washington, DC: The World Bank.

Rao, K., S. Ramani, S. Murthy, I. Hazarik, N. Khandpur, M. Choskshi, S. Khanna, M. Vujicic, P. Berman, and M. Ryan. 2010. "Health Worker Attitudes Towards Rural Services in India: Results from Qualitative Research." Health, Nutrition and Population Discussion paper, World Bank, Washington DC, November.

Ryan, M., V. Watson, and V. Entwistle. 2009. "Rationalising the 'Irrational': A Think Aloud Study of Discrete Choice Experiment Responses." *Health Economics* 18: 321-336.

San Miguel, F., M. Ryan, and M. Amaya-Amaya. 2005. "Irrational Stated Preferences: A Quantitative and Qualitative Investigation." *Health Economics* 14: 307-322.

Experimental design

Hahn, G.J., and S.S. Shapiro. 1966. "A Catalogue and Computer Program for the Design and Analysis of Orthogonal Symmetric and Asymmetric Fractional Factorial Experiments." Schenectady, New York: General Electric Research and Development Center.

Kuhfeld, W. 2010. "Marketing Research Methods in the SAS, Version 9.2." SAS Institute Inc., Cary, NC, USA

Sloane, N.J.A. 2009. "A library of orthogonal arrays." http://www.research.att.com/~njas/oadir/. Accessed April 2009.

Street, D.J., L. Burgess, and J.J. Louviere. 2005. "Quick and Easy Choice Sets: Constructing Optimal and Nearly Optimal Stated Choice Experiments." *International Journal of research Marketing* 22: 459-470.

Street, D.J., L. Burgess, R. Viney, and J. Louviere. 2008. "Designing Discrete Choice Experiments for Health Care." In M. Ryan, K. Gerard, and M. Amaya-Amaya (eds.) *Using Discrete Choice Experiments to Value Health and Health Care* (pp. 47-72). Dordrecht, The Netherlands: Springer.

Sample size

Ben-Akiva, M. and S. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand.* Cambridge, MA: MIT Press. http://elsa.berkeley.edu/reprints/misc/multinomial.pdf Accessed November 2006.

Cochran, W.G. 1977. *Sampling Techniques*. John Wiley and Sons. New York.

Hensher, D.A., J.M. Rose, and W.H. Greene. 2005. *Applied Choice Analysis: A Primer*. Cambridge, United Kingdom: Cambridge University Press.

Louviere, J., D. Hensher, and J. Swait. 2000. *Stated Choice Methods: Analysis and Application*. First edition. Cambridge, United Kingdom: Cambridge University Press.

Tanzania case study

Bech, M. and D. Gyrd-Hansen. 2005. "Effects Coding in Discrete Choice Experiments." *Health Economics* 14: 1079-1083.

Kolstad, J.R. 2011. "How to Make Rural Jobs More Attractive to Health Workers: Findings from a Discrete Choice Experiment in Tanzania." *Health Economics* 20(2): 196-211.

Scott, A. 2001. "Eliciting GP's Preferences for Pecuniary and Non-pecuniary Job Characteristics." *Journal of Health Economics* 20: 329-347.

Uganda case study

Barnighausen, T., and D.E. Bloom. 2009. "Financial Incentives for Return of Service in Underserved Areas: A Systematic Review." *BMC Health Serv Res* 9: 86.

Rockers PC, Jaskiewicz W, Wurts L, Kruk ME, Mgomella G, Ntalazi F, Tulenko K. Preferences for working in rural clinics among trainee health professionals in Uganda: a discrete choice experiment. (2012). *BMC Health Services Research* 2012, **12**: 212.

WHO. 2010. *Increasing Access to Health Workers in Remote and Rural Areas through Improved Retention: Global Policy Recommendations*. Geneva: WHO. http://www.who.int/hrh/retention/guidelines/en/index.html. Accessed 18 April 2011.

Websites referenced

http://www.research.att.com/~njas/oadir (Sloane's array)
http://www.choice-metrics.com/
http://www.sawtoothsoftware.com/
http://support.sas.com/documentation/index.html
http://www.spss.com
http://www.epidata.dk/
http://www.limdep.com/
http://www.stata.com/
http://www.sas.com/offices/europe/uk/
http://www.mathworks.co.uk/
http://www.r-project.org/
http://www-01.ibm.com/software/uk/analytics/spss/
http://legacy.measuredhs.com/cspro/
http://wwwn.cdc.gov/epiinfo/
http://office.microsoft.com/en-gb/excel/

Appendix:
Application of DCEs to address recruitment and retention in low- and middle-income countries

Study	Country	Cadre(s)	Respondents	Sample Size	Attributes
Blaauw et al (2010)	Kenya, South Africa and Thailand	Nurses	Final year students	345 (KEN), 377 (SAF), and 342 (THA)	Facility type* Salary Training Housing Promotion Additional Benefit Workplace culture * Some attribute levels differ across countries.
Chomitz et al (1998)	Indonesia	Physician	Final year students	585	Province Remoteness Total monthly income Length of contract Probability of subsequent appointment to the civil service Probability of subsequent specialist training
Hanson and Jack (2010)	Ethiopia	Physicians	In-service workers	219	Location Net monthly pay Housing Equipment and Drugs Time Commitment Private Sector (for Physicians)/ Supervision (for Nurses)
		Nurses	In-service workers	642	
Jaskiewicz et al. (2012)	Laos	Medical Doctors	Students and workers in rural provinces	329 Students, 105 In-service	Quality of Facility Career Promotion Housing Salary Continued Education Transport
		Medical Assistants	Students and workers in rural provinces	280 Students, 90 In-service	Quality of Facility Career Promotion Housing Salary Continued Education Children's Education
		Nurses*/ Midwives * includes mid and low level training nurses	Students and workers in rural provinces	361 Students, 289 In-service	Career Promotion Housing Salary Continued Education Transport Award
Kolstad (2011)	Tanzania	Clinical Officer	Final year students	320	Salary and allowances Education opportunities/ possibility of upgrading qualifications Availability of Equipment and Drugs Location Workload Housing Infrastructure
Kruk et al (2010)	Ghana	Physicians	Final year students	302	Salary Children's education Infrastructure, equipment, supplies Management style Years of work before study leave Housing Transportation

Study	Country	Cadre(s)	Respondents	Sample Size	Attributes
Mangham and Hanson (2008)	Malawi	Nurses	In-service workers	107	Place of work Net monthly pay Availability of material resources (equipment, drugs and other supplies) Typical workload Provision of government housing Opportunity to upgrade qualifications
McAuliffe et al. (forthcoming)	Malawi, Mozam-bique and Tanzania	Mid-level providers	In service (maternity staff)	631 Malawi, 587 Mozambique, 854 Tanzania	Location Net Monthly Pay Housing Equipment and Drugs Continuing Professional Development Human Resources Management
Penn-Kekana et al. (2005)	South Africa	Nurses	In-service workers	147	Salary Social Amenities Equipment Staffing Facility Mix
Rao et al. (forthcoming)	India	Physicians	In-service workers and medical students	222 in-service, 163 students	Type of health Center Area Health center Infrastructure Staff Salary (including allowances, Rupees/ month) Change in location to city/town Professional development Job location
		Nurses	In-service workers and nurses students	235 in-service, 145 Students	
Rockers et al. (2012)	Uganda	Physicians, Nurses, Pharmacists, Laboratory	Students and In-service workers	158 In-service, 485 Students	Salary Housing Facility infrastructure and equipment Length of contract Manager support Tuition support Staffing level Opportunity for dual practice
Vujicic et al. (2010a)	Liberia	Nurses	In-service workers	197	Location Equipment Total Pay Transportation Housing Workload
Vujicic et al. (2010b)	Vietnam	Physicians	Final year students and In-service workers	292 in-service and 105 students	Location Equipment Official Income Skills development (short-term training) Long term Education (specialist training) Housing

ISBN 978 92 4 150480 5