ECONOMIC STUDIES

DEPARTMENT OF ECONOMICS
SCHOOL OF ECONOMICS AND COMMERCIAL LAW
GÖTEBORG UNIVERSITY
145

MEASURING TRUST AND THE VALUE OF STATISTICAL LIVES: EVIDENCE FROM BANGLADESH

Minhaj Mahmud

ISBN 91-85169-02-1
ISSN 1651-4289 print
ISSN 1651-4297 online

Economic Studies No. 145
(PhD Thesis)
Department of Economics
School of Economics and Commercial Law
Göteborg University

MEASURING TRUST AND THE VALUE OF STATISTICAL LIVES: EVIDENCE FROM BANGLADESH

Minhaj Mahmud

Address: Department of Economics, Göteborg University
Vasagatan 1, Box 640, SE-40530, Göteborg, Sweden
Phone: +46 31 773 4669; Fax: +46 31 773 1043
Email: Minhaj.Mahmud@economics.gu.se

To Mahin

Contents

Abstract

This thesis includes five self-contained essays. The first three essays relate to the measurement of trust using both an experimental and a survey approach, and the other two essays relate to the measurement of the value of (statistical) life using stated preference methods.

Essay 1: The proportion of money sent, which is typically assumed to reflect trust, decreased significantly as the stake size was increased in a trust game conducted in rural Bangladesh. Nevertheless, even with very large stakes, most senders and receivers sent substantial fractions.

Essay 2: Trust is measured using both survey questions and a trust experiment using random sample of individuals in rural Bangladesh. We found no significant effect of the social distance between Hindus and Muslims in the trust experiment in terms of fractions sent or returned, but the responses to the survey questions indicate significant differences: Hindus, the minority, trust other people less in general, and Hindus trust Muslims more than the other way around.

Essay 3: Levels of trust are measured by asking standard survey questions on trust and by observing the behavior in a trust game using a random sample in rural Bangladesh. Follow-up questions and correlations between the sent amount in the trust game and stated expectations reveal that the amount sent in the trust game is a weak measure of trust. The fear of future punishment, either within or after this life, for not being sufficiently generous to others, was the most frequently stated motive behind the respondents' behavior, highlighting the potential importance of motives that cannot be inferred directly from people's behavior.

Essay 4: Using the contingent valuation method in developing countries to value mortality risk reduction is particularly challenging because of the low level of education of the respondents. In this paper, we examine the effect of training the respondents regarding probabilities and risk reductions, in addition to using visual aids to communicate risk and risk reductions, in a contingent valuation survey. Our results indicate a significantly higher willingness to pay (WTP) for the trained sub-sample, and WTP is sensitive to the magnitude of risk reduction both with and without the training.

Essay 5: By assuming that an individual has preferences concerning different states of the world and these preferences can be described by an individual social welfare function, we explore the relative value of statistical life using survey data from Bangladesh. We apply a pair-wise choice experiment on life-saving programs to elicit individuals' preferences regarding differences in the values of statistical lives related to age. We find that the relative value decreases strongly with age and that people have strong preferences for saving more life-years, rather than lives *per se*. Moreover, in specific follow-up questions, it is again elicited that a majority of the respondents believe that it is better, from a social point of view, to save younger individuals.

Keywords: Social capital; trust; social distance; religion; trust game; stake size; field experiment; value of statistical life; contingent valuation; risk reduction; effect of training; willingness to pay; sensitivity to scope; social preference; choice experiment; life-saving programs; relative value of life; Bangladesh.

Preface

First of all, I would like to express my deepest gratitude to my thesis supervisors Olof Johansson-Stenman and Peter Martinsson. I am forever grateful to Olof and Peter for accepting my research ideas, and for providing continuous intellectual support and encouragement. They have not only provided me with excellent supervision, they have also shared with me their intellectual talents by co-authoring part of the thesis; they have made crucial contributions to this thesis and to my research skills. Olof and Peter, I cannot thank you enough. I look forward to working with you in the future.

I would like to thank Professor Karl Göran Mäler, who encouraged me to apply to the PhD program at Göteborg University when I met him in a Beijer Workshop in 1998. At that time I was at the University of Manitoba doing a Masters in Economics, and Professor Wayne Simpson and Dr. Betsy Troutt wrote two very nice letters of recommendation that made my way to Göteborg. I am thankful to them.

I would like to thank Professor Thomas Sterner and Gunnar Köhlin for their support and encouragement throughout my PhD studies and for facilitating a very generous funding for the completion of my thesis work. I look forward to working with them in the future.

I greatly appreciate the financial assistance from the Swedish International Development Agency (SIDA).

The thesis has benefited greatly from the comments of Fredrik Carlsson and Håkan Holm. I cannot thank them enough. I would like to thank Jean Robert-Tyran, David Reiley, and an anonymous referee for very useful comments on the first essay.

In the earlier stages of my work, I benefited from comments made by Fredrik Carlsson, David Layton, Martin Dufwenberg, Katarina Nordbolm and Dale Whittington, in addition to the suggestions of my supervisors. I thank them all. I also thank seminar participants at Göteborg University, Umeå University, participants at the ESA meeting in Tuscon in 2003, the 2nd Budapest Workshop in Behavioral Economics in 2004, the Nordic Conference in Development Economics in Göteborg in 2004, and the 5th European Conference on Health Economics in London in 2004, for useful comments that improved the thesis.

This thesis would not have been finished unless I had been granted a prolonged period of study leave from my home university, Jahangirnagar University, Bangladesh. I greatly appreciate the university administration's support of my work. I would like to thank all my teachers and colleagues in the Department of Economics, especially Professors Khandker Mustahidur Rahman, Anu Muhammad, and Nurul Hoque for their help and encouragement during my visits in Bangladesh for data collection. I would like to thank Dr. Rowshan Jahan (my older sister) of Jahangirnagar University for her encouragement and support during the data collection. I wish that my friend and colleague Abdullah Al Mamun was alive to see this work. Sadly, he died from a sudden heart attack soon after finishing his PhD in economics at Kansas University in November 2004. I would like to thank my friends who greatly encouraged my work: Bappi-apa, Naimul Islam, Mafruhi Sattar, Laek Sazzad, Shamsad Mortuza, and Rubana Mahjabeen.

I would like to thank my colleagues in the PhD programs, whom I have enjoyed working with during our PhD course works: Nasima Chowdhury, Astrid Nunez, Daniel Deng, Erik Liden, Marcus Eliason, Mahmud Yesuf, Razack Bakari, Wilfred Nyangena,

and Sandra Lerda (Beijer Institute). Special thanks to Nasima for finding some errors in the manuscript! My thanks go to all other colleagues in the unit including: Hala Abou Ali, Håkan Eggert, Anders Ekbom, Henrik Hammar, and Bjorn Olsson, who were always very helpful whenever I turned to them for support.

I would like to thank Elizabeth Földi, Gerd Georgsson, Eva Jonasson, Ulla Mellgren, Eva Lena Neth, Katarina Renström and Anna Karin Ågren for providing excellent administrative support, facilitating my work greatly. Special thanks to Elizabeth for her loving attention to make our stay in Göteborg smooth and easy.

I would like to thank my brother and sisters, and all family members for their encouragements throughout my work. I am very much grateful to my brother Shihabuddin Mahmud for his continuous support of my work. I would also like to thank my parent- in- laws for their encouragement, and especially to my mother- in- law who came all the way to Göteborg to take care of my family while I was doing my field work.

Finally, I would like to express my deepest gratitude to my parents for their love, encouragement and prayers that made me accomplish my work. Last but not least, I would like to express my heartfelt thanks and love to my wife Doreen for her great patience, love, and support of my work, and for taking care of our daughter Mahin almost on her own; Doreen- *vishon osombhobe tomakey chai*. I would like to dedicate this thesis to my Mahin babu.

Göteborg
February 2005 *Minhaj Mahmud*

Introduction

This thesis includes five self-contained essays. The first three essays relate to the measurement of trust using both an experimental and a survey approach, and the other two essays relate to the measurement of the value of (statistical) life using stated preference methods. The essays focus mostly on the methodologies. Although they are largely empirical and based on data collected in the field in rural Bangladesh, the essays address important methodological issues when measuring trust and the value of (statistical) life.

To begin with, a brief note on the sample used in this thesis (all essays) is provided. The survey and the experiments were conducted among a random sample of rural household heads in the following five districts of the Dhaka Division in Bangladesh: Gazipur, Narayanganj, Netrokona, Manikganj, and Mymensingh, during October-November of 2003. The sample is, however, not representative of the Bangladesh population, which consists of 64 districts. Moreover, the villages are chosen so that the respondents of the Hindu religion are over-represented in order to facilitate religious comparisons in the trust experiment. The detailed survey methodology and surrounding issues are discussed in each essay.

It is pertinent to provide a brief background on Bangladesh, before presenting an overview of the thesis. Bangladesh became independent from Pakistan in 1971. Unlike Pakistan, which was created on the basis of Muslim nationalism and where religion and the state were inseparable during the 1947-1971 period, Bangladesh drew inspiration from Bengali nationalism and its constitution made a categorical declaration that Bangladesh would be a secular state having no institutional relation with religion

(Banglapedia, 2003). However, the country failed to build a secular state and eventually Islam was declared the state religion in 1988. In Bangladesh, there are two main religions: Islam, accounting for 88% of the population, and Hinduism accounting for about 11% of the population; the remaining 1% mainly consists of Christians and Buddhists[1] (Banglapedia, 2003).

Since independence, the country has gradually improved the lives of its people, although many development pundits termed it "a test case for development" and questioned its long-run economic prospects (see Faaland and Parkinson, 1976). The country's per capita income grew annually at a rate of 1.9% during the 1975-2002 periods and at a rate of 3% during the 1990-2002 periods (HDR, 2004). During the 1990s the country's population growth rate has been reduced to 1.5 %, from 2.4% in the seventies (World Bank, 2004). In the same period there was a 36% rise in real per capita income, which helped reduce poverty by 9% (World Bank, 2004). The country also sustained an average 5% annual GDP growth during the recent years; however, almost half of the population still live below the poverty line (World Bank, 2004) and Bangladesh needs to achieve much higher economic growth for reducing its poverty by half by 2015, as envisaged in the millennium development goals.

Most notable, Bangladesh made substantial progress in terms of Human Development. For example, life expectancy at birth has increased from 42.2 years in 1970 to 61.1 years in 2002, and the infant mortality rate has been reduced from 145 per 1,000 live births in 1970 to 51 per 1,000 live births in 2002 (see HDR, 2004). Adult literacy is currently 41% (50% for males and 31% for females), and now Bangladesh is

[1] However, 98% of the population is Bengali and the remaining ethnic groups consist of tribal groups as well as non-Bengali Muslims.

termed "medium human development country" (HDR, 2004). In the appendix, however, we present figures (diagrams) relating some international statistics in order to compare Bangladesh's situation with some developing as well as developed countries. As depicted in those figures, both infant mortality rates and child mortality rates are still higher, than most included countries, and life expectancy is lower compared to even many developing countries (HDR, 2004). Moreover, Bangladesh is facing major environmental problems including air pollution, water contamination posing serious public health risks and increasing pre-mature mortality. This calls for the application of economic valuation to compare the benefits of available programs reducing public health risks with the costs of the implementation.

On the other hand, the problems of economic growth, poverty alleviation, social and human development, are all linked to a lack of good governance and corruption. The country has been top ranked in the Transparency International's corruption perception index implying the highest measured extent of corruption of public officials four years in a row, 2001-2004 (Transparency International 2003, 2004)[2]. For example, in 2001 corruption represented a loss to government revenue of 2 billion USD, or 4.7% of the 1999-2000 GDP (TI- Bangladesh, 2002). The issue of corruption makes it particularly relevant to study and measure trust in the context of Bangladesh. Given that public officials and politicians are perceived to be corrupt, this may also influence lower levels of society. As argued, for example, by Alesina and Ferrara (2002), trust in existing institutions may therefore affect trust in other people. Rothstein and Stolle

[2] Corruption is defined as the abuse of public office for private gain, e.g. bribe taking by public officials in public procurement. Corruption perception index is a composite index, which reflects the views of business people, and analysts from around the world, including the experts who are residents of the respective country (Transparency International 2003, 2004)

(2001) hypothesized that the development of institutional characteristics, such as corruption, is the most important factor for the spread of distrust and general suspicion in a society, in contrast to Putnam (1993) who argued that trust largely develops through people's interactions in local voluntary organizations.

The importance of trust and social capital is emphasized in recent literature on economic and social development. There is now much theoretical and empirical evidence that trust among people fosters cooperation and economic activities, and hence it is important for economic and social development (Putnam, 1993, Fukuyama, 1995; Knack and Keefer, 1997; La Porta et al., 1997; Zak and Knack 2001). Economists emphasize trust as a way of reducing the transaction cost, lubricating the economy (Arrow, 1972). As observed by Douglas North (1990, p.54), "The inability of societies to develop effective, low cost enforcement of contracts is the most important source of both historical stagnation and contemporary underdevelopment in the Third World."

The use of trust questions to explain differences in functioning between societies is now more and more frequently applied in social science. The most frequently used questions for measuring attitudinal trust is framed as "Generally speaking, would you say that most people can be trusted, or that you cannot be too careful in dealing with people?" Zak and Knack (2001) report, on cross-country observations, that higher trust is associated with stronger formal institutions for enforcing contracts and reducing corruption. They show that formal institutions appear to affect growth rates partly through their effect on trust. Knack and Keefer (1997) show that a one standard deviation increase in a measure of country level trust increases economic growth by more than one half of a standard deviation. Evidence from household and village studies

suggests that social capital in the form of trust can play an important role in alleviating poverty (Grootaertt, 1999; Narayan and Pritchett, 1999).

However, the use of survey questions to measure trust has frequently been criticized since the choices made are non-consequential for the respondents (e.g. Gleaser et al., 2000). On the contrary, the decisions made in a trust game (Berg et al., 1995) are consequential since the decisions will have a monetary effect for the participants, and amount gained depends on trust. Briefly, in a trust game, both a sender (first mover) and the receiver (second mover), anonymous to each other, are given a certain amount of money and the sender has to decide first on how much of this to send to the receiver, and how much to keep for herself. Any positive amount sent by the sender is normally tripled before it is given to the receiver. Then the receiver has to decide on how much of the tripled amount of money received to return to the sender, and after that the game is over. The theoretical sub-game perfect prediction of this game is that the sender should send nothing to the receivers, since she would realize, by using backward induction, that the receiver has no incentive to send anything back. This is a "social dilemma situation," however, an improvement for both parties is possible if money is sent, given that the receiver will send back at least one third of the tripled amount received. Thus, the structure of the game allows the sender to use trust in order to achieve an improvement over the predicted outcome. The amount sent by the sender is typically regarded as an indication of trust, and the amount returned by the receiver as an indication of trustworthiness.

Generally, we find that survey responses reflect rather low levels of general trust, implying support for the hypothesis that corruption creates distrust among people at the grass-root level. On the other hand, the fraction sent in the trust game is of an

order of magnitude similar to most previous studies, implying higher levels of trust, which does not provide any support for this hypothesis. In the first three essays, we attempt to address the following questions: i) in Essay 1, we investigate if stake size matters in a trust game i.e. whether the senders' behavior in a trust game depends on her initial endowment; ii) in Essay 2, we investigate whether social distance based on religious belief matters for trust and trustworthiness, i.e. we test whether individuals are less inclined to trust others who are different from themselves in terms of the two main religions of Bangladesh; iii) in Essay 3, we finally contribute to the discussion about what trust games and trust surveys actually measure.

Essay 4 and Essay 5 relate to the measurement of the value of statistical lives using stated preference approach, i.e. the contingent valuation method, as well as a choice experiment method.

The value of statistical life (VSL) is based on individuals' rate of trade-off between own money and a small risk change, i.e. the marginal rates of substitution between them (see Weinstein et al., 1980; Johanssson, 2002). The individual's willingness to pay (WTP) for a reduction in the mortality risk is converted to VSL by dividing the WTP by the risk change in question.[3] The absolute VSL has been studied in monetary terms for a long time using both revealed and stated preference methods, mostly in developed countries (Viscusi and Eldy, 2003; Hammitt and Graham, 1999). In developed countries, VSL estimates are of great interest as a basis for efficient risk regulation and are used in cost benefit analyses of governmental projects relating to

[3] For example, if an individual is willing to pay 10USD for reducing the risk of dying by 1 in 10,000 during a given year, then the individual's VSL is 100,000 USD, which is similar to say that 10,000 similar people would together be willing to pay 100,000 USD to reduce the same risk that would randomly kill one of them during that period

environment, health or safety regulations (Krupnick, 2004). Like many other developing countries, Bangladesh is facing major environmental problems such as air pollution and water contamination, requiring the government to undertake programs, on its own or with international assistance, to reduce the risks to public health. To compare the health benefit of programs reducing public health risks with the costs of implementation, it is useful to have the monetary values of the risk reductions, i.e. the values that individuals place on own risk reductions.

In Essay 4, we discuss the results from a contingent valuation (CV) study eliciting people's VSL from a developing country perspective. The CV method involves eliciting willingness to pay for hypothetical risk reduction. Past studies, however, have discussed that it is problematic to consistently measure the VSL using the CV method, largely due to the cognitive burden that the respondents face when comparing expected welfare effects of a small reduction in the risk to those of small monetary changes (Beattie et al., 1998; Hammitt and Graham, 1999). In particular, we measure the effect of training, regarding probability and risk concepts, on the WTP responses and on the sensitivity of WTP to the size of risk reduction, in the CV survey. We use substantial risk changes to be valued, corresponding to the respondents' stated subjective mortality risks during the next five years, which in turn are based on the age-related objective risks of dying presented in the survey. Over all, the results suggest that estimates of absolute VSL can be obtained using the CV method in a developing country like Bangladesh. However, there are remaining problems, most of which are related to the CV methodology *per se*, and not to the contexts of developing countries.

The idea of using relative VSL in decision making is more recent, and has been frequently discussed since the publication of the 1993 World Development Report

(World Bank, 1993). The authors of the report observe that "Most societies attach more importance to a year of life lived by a young or middle-aged adult than to a year of life lived by a child or an elderly person." There might be several reasons for prioritizing to save the younger people rather than the older. For example, as argued by Murray and Lopez (1994, p. 8), "In all societies social role varies with age. The young, and often the elderly, depend on the rest of society for physical, emotional and financial support. Given different roles and changing levels of dependency with age, it may be appropriate to consider valuing time lived at a particular age unequally." There is now evidence from developed countries, where a substantial amount of public revenue is being spent on social security systems, that people place more weight on saving younger people's lives than on saving the lives of older people when choosing between alternative life-saving programs (USA: Cropper et al., 1994; SWEDEN: Johannesson and Johansson, 1997; Johansson-Stenman and Martinsson, 2004).

In Essay 5, from a developing country perspective, we elicit individuals' preferences regarding differences in the values of statistical lives related to age by applying a pair-wise choice experiment on life-saving programs. Briefly, in a choice experiment, respondents make repeated choices between different alternative goods or projects described by their attributes (see Louviere et al., 2000; Alpizar et al., 2003). Hence, one could easily estimate the marginal impact of different attributes on the decision. In our case, the respondents were presented with six different pairs of life-saving programs which differed with respect to the number of lives saved and the age group of the saved persons; however, the programs were similar in other aspects, including the costs. The results indicate that relative value decreases strongly with the age of the individuals saved and people have stronger preferences for saving more life-

years, rather than lives. While our results are informative for priority setting in public health projects in the context of developing countries, it is emphasized that more research is needed focusing on the methodology in eliciting people's preferences on such issues.

References

Alesina, A., and La Ferrara, E. (2002), Who trusts others? Journal of Public Economics 85, 207-234.

Alpizar, F., Carlsson, F., and Martinsson, P. (2003) Using Choice Experiments for Non-Market Valuation, Economic Issues, 8:83-110.

Arrow, K. (1972), Gifts and Exchanges, Philosophy and Public Affairs 1, 343-362

Banglapedia (2003), National Encyclopedia of Bangladesh, Asiatic Society of Bangladesh, Dhaka.

Beattie, J., Covey, P. Dolan, L. Hopkins, M, Jones-Lee, G., Pidgeon, A. Robinson, and Spencer, A. (1998), On the Contingent Valuation of Safety and the Safety of Contingent Valuation: Part 1- Caveat Investigator, Journal of Risk and Uncertainty 17(3): 5-25

Berg, J., Dickhaut, J., McCabe, K.(1995), Trust, Reciprocity, and Social History, Games and Economic Behavior 10, 122-42.

Cropper, M. L., Aydede, S.K., Portney, P.R. (1994), Preferences for Life-saving Programs: How the Public Discounts Time and Age, Journal of Risk and Uncertainty 8, 243-265.

Faaland, J and Parkinson, J.R. (1976), Bangladesh: The Test Case for Development, London.

Fukuyama, F. (1995), Trust, New York: Free Press.

Glaeser, E. L., D.I. Laibson, J.A. Scheinkman, C.L. Soutter (2000), Measuring Trust, Quarterly Journal of Economics 115, 811-846.

Grootaerrtt (1999), Social Capital, Household Welfare and Poverty in Indonesia, World Bank Policy Research Working Paper 2148, Washington, DC.

Hammitt, J. K. and Graham, J.D. (1999), Willingness to Pay for Health Protection: Inadequate Sensitivity to Probability? Journal of Risk and Uncertainty 18, 33-62.

HDR (2004), Human Development Reports, UNDP, http://hdr.undp.org/statistics/data.

Johannesson, M and Johansson, P.O. (1997), Is the Valuation of a QALY gained independent of age? Some empirical evidence, Journal of Health Economics 16, 589-599.

Johansson, P.O. (2002), The Value of a Statistical Life: Theoretical and Empirical Evidence, Applied Health Economics and Health Policy, 1(1), 33-41.

Johansson-Stenman, O. and Martinsson, P. (2004), Are Some Lives More Valuable? Working Paper, Department of Economics, Göteborg University.

Knack, S. and Keefer, P. (1997), Does Social Capital Have an Economic Payoff? A Cross-Country Investigation, Quarterly Journal of Economics 112, 1251-1288.

Krupnick, A.J. (2004), Valuing Health Outcomes: Policy and Technical Issues, RFF Report, Resource for the Future, Washington, March.

La Porta, R., Lopez-De-Silanes, F., Shleifer, A., and Vishny, R.W. (1997), Trust in Large Organizations, American Economic Review Papers and Proceedings, May.

Louviere, J.J., Hensher, D.A., Swait, J.D (2000), Stated Choice Methods, Analysis and Application, Cambridge University Press.

Murray, C.J.L., and Lopez, A.D. (1994.), Global Comparative Assessments in the Health Sector: Disease Burden, Expenditures and Intervention Packages, World Health Organization, Geneva.

Narayan, D. and Pritchett, L. (1999), Cents and Sociability: Household Income and Social Capital in Rural Tanzania, Economic Development and Cultural Change 47(4), 871-897.

North, D.C. (1990), Institutions, Institutional Change and Economic Performance, Cambridge, Cambridge University Press.

Putnam, R.D. (1993), Making Democracy Work- Civic Tradition in Italy, Princeton University Press.

Rothstein, B. and Stolle, D. (2001), Social Capital and Street –Level Bureaucracy: An Institutional Theory of Generalized Trust, Prepared for the ESF conference Social Capital: Interdisciplinary Perspectives, in Exeter, United Kingdom, September 15-20.

TI–Bangladesh (2002), Corruption in Bangladesh: A Household Survey 2002, Executive Summary, http://www.ti-bangladesh.org.

Transparency International (2003), Transparency International Corruption Perception Index 2003, Press Release, http://www.transparency.org.

Transparency International (2004), Transparency International Corruption Perception Index 2004 Press Release, http://www.transparency.org

Viscusi, W. K., and Aldy, J.E.(2003), The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World, Journal of Risk and Uncertainty, 27(1), 5-76.

Weinstein, M.C., Shepard, D.S., and Pliskin, J.S. (1980), The Economic Value of Changing Mortality Probabilities: A Decision Theoretic Approach, Quarterly Journal of Economics 94(2), 373-396.

World Bank (1993), World Development Report 1993: Investing in Health, London, Oxford University Press.

World Bank (2004), Poverty in Bangladesh: Building on Progress, http://www.worldbank.org.

World Health Report (2004), The World Health Report 2004 -Changing History, The World Health Organization, Geneva.

Zak, P.J. and Knack, S. (2001), Trust and growth, Economic Journal 111, 295-321.

Appendix

Figure 1. Infant and Child Mortality Rate in Selected Countries (Region), 2002

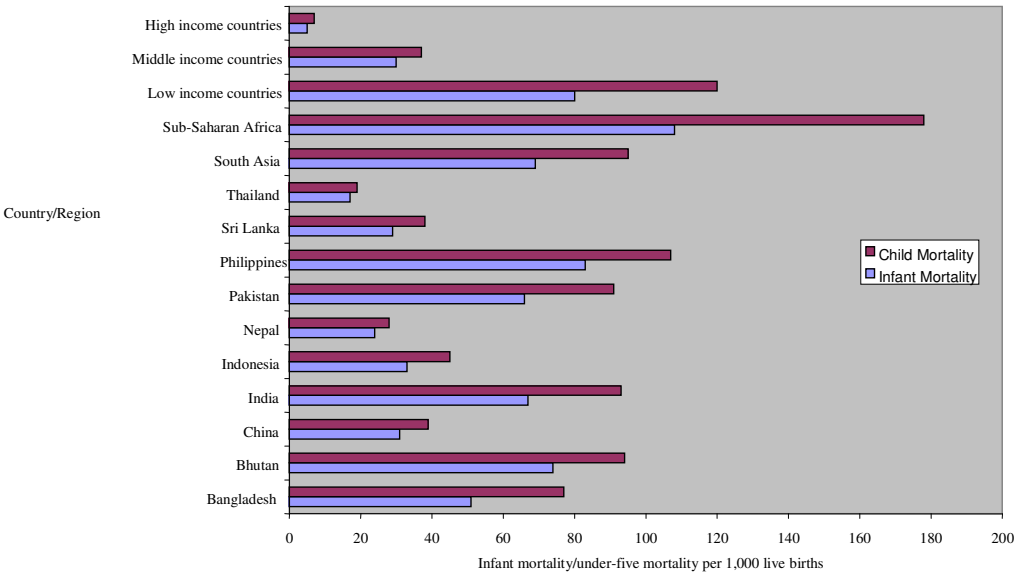
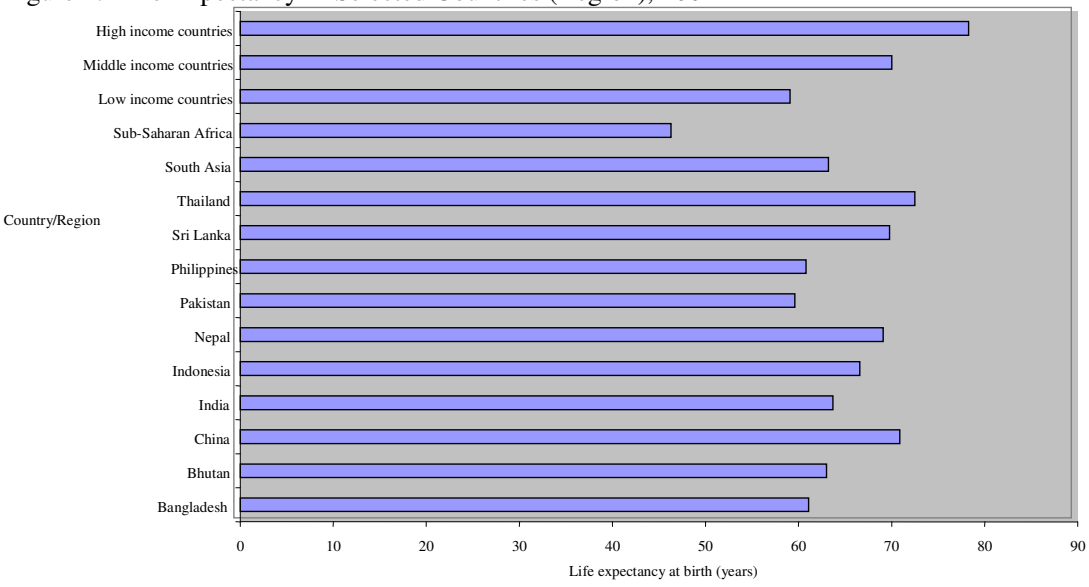


Figure 2. Life Expectancy in Selected Countries (Region), 2002



Does stake size matter in trust games?

Olof Johansson-Stenman, Minhaj Mahmud, and Peter Martinsson
Department of Economics, Göteborg University
Vasagatan 1, Box 640, SE-405 30, Göteborg, Sweden[*]

Abstract

The proportion of money sent, which is typically assumed to reflect trust, decreased significantly as the stake size was increased in a trust game conducted in rural Bangladesh. Nevertheless, even with very large stakes, most senders and receivers sent substantial fractions.

Key words: Trust; trust game; stake size; social capital; field experiment

JEL classification: C93, Z13

[*] The authors thank Håkan Holm, David Reiley, Jean-Robert Tyran and participants at the ESA Meeting in Tucson 2003 for useful comments.

1. Introduction

The importance of trust and social capital is emphasised in recent literature on economic and social development; see e.g. Knack and Keefer (1997) and Zak and Knack (2001). However, the question of how to measure trust accurately and reliably is debated. In this paper we test whether the stake size matters in a trust game (Berg et al. 1995), which has become the most frequently used measure of individual trust other than using survey questions.

There is considerable evidence that stake size does not, in general, significantly affect the offers made in ultimatum games (Hoffman et al., 1996; Slonim and Roth, 1998; Cameron, 1999; Munier and Zaharia, 2002). Perhaps more surprisingly, Forsythe et al. (1994) and Carpenter et al. (forthcoming) find that the stake size gives no significant effect in dictator games either. One might therefore suppose that the same would hold true for trust games as well. However, in this study, which, as far as we know, is the first study to test the effect of stake sizes in trust games, we found that the amount sent decreased significantly when the stake size was increased. This calls into question the validity of comparing point estimates from different trust games. Moreover, the results support evidence from other recent papers that the amount sent by the first player may be a poor measure of trust.

2. The trust game

Participants in the game are unknown to each other, divided into two groups and asked to act either as a 'sender' or as a 'receiver'. The sender is given a certain amount of money and decides how much of it to send to the receiver and how much to keep. Any positive amount sent is tripled before it reaches the receiver, who then decides how

much of the tripled amount of money received is to be returned to the sender. With perfect information, the conventional (albeit Pareto-inefficient) sub-game perfect prediction of this game is that rational senders should send nothing, since they would realise that a purely self-interested receiver has no incentive to send anything back. The fractions sent and returned are typically regarded as measures of trust and trustworthiness, respectively. Empirically, virtually all studies so far (including this one) have strongly rejected the conventional theoretical prediction.

This trust game was conducted among rural household heads in the districts Manikganj, Mymensingh and Netrokona of the Dhaka division in Bangladesh, at the end of a household survey. We applied ex-ante matching of senders and receivers from different villages, and a random sample strategy based on every fourth household in the villages. The sample was divided into three groups with different initial endowments for the senders. We ended up with 64, 59 and 62 pairs for low (40 Taka), medium (200 Taka) and high stakes (1000 Taka), respectively. The highest amount is substantial and equals 4.8% of the GNI per capita; the same fraction in the US would correspond to 1683 USD. Each respondent also gets a participation fee of Taka 100 for completing the whole survey including the trust game.

The enumerators provided a private interview environment, free from any possible interruptions, and read the instructions to the sender. They also presented the outcomes of different examples of decisions, both related to the amount sent by the sender and the amount returned to them by the receiver, and took great care to make sure that the respondents understood the mechanisms involved.

The senders were given two thick envelopes, one containing their original endowment and one empty. The enumerator ensured confidentiality by turning his back to the sender while the sender put the chosen amount of money into the initially empty envelope. The sender was then asked to close the envelope and seal it with a stamp before returning it to the enumerator. The sender was instructed to do so even if he/she had decided to send nothing, and was informed that he/she would be paid within three days. At the end of the day, the enumerator handed the envelopes to the principal investigator, who opened them and put the tripled amount of money into new envelopes with a pre-matched household code.

The following day the enumerators were given these new envelopes ready to be delivered to the assigned receiver, and the enumerators followed similar procedures to those followed in the senders' households. The receiver put the chosen amount to be returned in the previously empty envelope while the enumerator turned his back. The enumerator then returned the envelope from the receiver to the principal investigator, who checked and wrote down the amount to be transferred back by another enumerator the following day.[1]

3 Results

Table 1 reveals that the average proportion sent clearly decreases as the stake size increases. There is no equally clear pattern for receivers, although high-stake receivers on average sent back less.

[1] The complete questionnaire is available from the authors upon request.

Table 1. Average proportion sent and average proportion returned in the trust game

	Low stake	Medium stake	High stake	Total
Proportion sent	0.55	0.46	0.38	0.46
Proportion returned	0.46	0.46	0.38	0.43
Proportion of zero sent	0.06	0.06	0.03	0.05
Proportion of zero returned	0.07	0.03	0.09	0.06

To test whether the differences obtained are statistically significant or not we conduct pair-wise comparisons by stake size, using the non-parametric Wilcoxon-Mann-Whitney test. That the amounts sent in the high and the low stake treatments come from the same underlying statistical distribution is rejected at 0.1% significance level. The corresponding hypotheses are rejected for the amounts sent in the low and the medium stake treatments at 10% significance level, and the medium and the high stake treatments at 5% significance level. Similarly, based on the non-parametric Kruskal-Wallis test, we can reject at 0.1% significance level the hypothesis that the proportion sent for all stakes comes from populations with the same distribution. However, we cannot reject at 5% significance level that the proportions sent back in the different sub-samples come from populations with the same distribution, using either repeated pair-wise Wilcoxon-Mann-Whitney tests or the joint Kruskal-Wallis test (p-value = 0.22). The average proportions sent and returned are quite similar to many other trust games, such as Berg et al. (1995). Since the average return ratio is higher than one third, it is on average profitable for senders to send money to the receivers.

The regression results in Table 2 show again a significant effect of the size of the stake on the proportion sent, i.e. also after correcting for other variables.[2] The amount sent also strongly increases with equivalence-scaled household income.[3]

Table 2. Trust and trustworthiness: OLS estimates

Dependent variable	Proportion sent	Proportion returned	Proportion Returned
Medium stake endowment (200 taka)	-0.076	-0.012	-0.075
	(0.054)	(0.064)	(0.138)
High stake endowment (1000 taka)	-0.178***	-0.100	-0.318**
	(0.056)	(0.067)	(0.131)
Most people can be trusted	-0.003	0.013	0.012
	(0.016)	(0.020)	(0.020)
Confidence in public institution index	-0.014**	0.010	0.010
	(0.006)	(0.010)	(0.010)
Has experienced a recent misfortune	-0.104**	-0.080	-0.100
	(0.051)	(0.062)	(0.061)
Member of voluntary association	-0.030	0.010	0.010
	(0.052)	(0.060)	(0.100)
Age	0.002	0.003	0.003
	(0.002)	(0.002)	(0.002)
Illiterate	0.012	-0.020	-0.010
	(0.052)	(0.063)	(0.062)
Educated above high school level	-0.038	-0.018	-0.020
	(0.078)	(0.075)	(0.074)
Equivalence-scaled income per capita	0.270**	0.167	0.120
	(0.120)	(0.111)	(0.112)
Proportion sent by the senders		-0.022	-0.230
		(0.092)	(0.170)
Proportion sent × medium stake			0.010
			(0.230)
Proportion sent × high stake			0.480**
			(0.230)
Constant	0.660***	0.190	0.330*
	(0.136)	(0.160)	(0.178)
R squared	0.127	0.100	0.100
Number of observations	182	172	172

Standard errors are in parentheses. Superscripts *, **, *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

[2] Since several enumerators were used to run the experiment, we tested for enumerator effects and we cannot reject the hypothesis of 'no enumerator effects' in all three regressions in Table 2 (p-value 0.21, 0. 60, and 0.13 respectively for Model 1, Model 2, and Model 3).

[3] Calculated as [Household income/ (adults+ 0.5*children) $^{0.75}$]. Sample mean = 21,100 Taka/year.

Twenty-six percent of the sample had experienced at least one recent misfortune in terms of robbery/theft, mugging, personal assault, home attacked, land fraud, false accusation of a criminal offence or political harassment during the previous year. Such an experience significantly decreases the amount sent. However, it is possible that this effect largely reflects a reduction in wealth resulting from the crime.

Stated trust was measured on a six-point scale as the level of agreement with the statement "most people can be trusted," where "strongly disagree" is quantified as 0 and "strongly agree" is quantified as 5. The average score equals 2.3, indicating a low level of average stated trust. As in Glaeser et al. (2000), there is no significant effect of stated trust on the amount sent. Confidence in public institutions was measured as a summation index for the following institutions: banks, NGOs, the military, the police, the judiciary, the government (executive branch), the local government, educational institutions, political parties, and rural power elites. "Hardly any confidence at all" is quantified as 0, "only some confidence" is quantified as 1, and "a great deal of confidence" is quantified as 2. The sample mean equals 14.1 (out of 20). Such confidence does not increase, but rather reduces, the amount sent. Neither illiteracy (35% of the sample) nor an education above high school (13% of the sample) affect the amount sent. Being a member of any voluntary association (30% of the sample) does not affect the amount sent.

For the proportion sent back, none of the parameters associated with the explanatory variables, including the stake-size dummy variables, are statistically significant at conventional levels. The latter is consistent with the results of Forsythe et al. (1994) and Carpenter et al. (forthcoming) mentioned earlier, since this part of the trust game can be seen as a conditional dictator game. The result of the interaction effect

in the last reported model implies that when lower levels are sent, relatively more is returned in the low and medium stake cases, and vice versa. Thus, senders seem to be rewarded for sending a large proportion in the high-stake case. However, since the model explains very little of the observed variation, these results should be treated cautiously.

4. Conclusion

One possible explanation for our finding that the amount sent significantly decreases as the stake size increases is linked to the suggestion made in Karlan (forthcoming) that the first part of the trust game largely measures risk preferences, rather than trust. That higher stakes induce people to send lower amounts in trust games is then consistent with Binswanger (1980) and Holt and Laury (2002), who found that people tend to become more risk averse with higher stakes. Moreover, we found that the amount sent increases significantly with the respondent's income, implying that the amount sent decreases when the stake size as a fraction of income increases.

Finally, using very large financial incentives offers no rescue for the conventional economic predictions. The majority of both senders and receivers sent substantial fractions also with very high stakes.

References

Berg, J., Dickhaut, J., McCabe, K., 1995, Trust, Reciprocity, and Social History, Games and Economic Behavior 10, 122-42.

Binswanger, H. P., 1980, Attitudes toward risk Experimental measurement in rural India, American Journal of Agricultural Economics 62, 395-407.

Cameron, L. A., 1999, Raising the stakes in the ultimatum game: Experimental evidence from Indonesia, Economic Inquiry 37, 47-59.

Carpenter, J., Verhoogen, E., Burks, S., forthcoming, The effect of stakes in distribution experiment, Economics Letters.

Forsythe, R., Horowitz, J., Savin, N.E., Sefton, M., 1994, Fairness in Simple Bargaining Experiments, Games and Economic Behavior 6, 347-369.

Glaeser, E. L., Laibson, D., Scheinkman, J.A., Soutter, C.L., 2000, Measuring trust, Quarterly Journal of Economics 115, 811-846.

Hoffman, E., McCabe, K., Smith, V., 1996, On expectations and the monetary stakes in ultimatum games, International Journal of Game Theory 25, 289-301.

Holt, C. A. and Laury, S.K., 2002, Risk aversion and incentive effects, American Economic Review 92, 1644-55.

Karlan, D. S., forthcoming, Using Experimental Economics to Measure Social Capital and Predict Financial Decisions, American Economic Review.

Knack, S., Keefer, P., 1997, Does Social Capital Have an Economic Payoff? A Cross-Country Investigation, Quarterly Journal of Economics 112, 1251-1288.

Munier, B. and Zaharia, C., 2002, High stakes and acceptance behavior in ultimatum bargaining: A contribution from an international experiment, Theory and Decision 53, 187-207.

Slonim, R., and Roth, A.E., 1998, Learning in High Stakes Ultimatum Games: An Experiment in the Slovac Repuplic, Econometrica 66, 569-596.

Zak, P.J. and Knack, S., 2001, Trust and growth, Economic Journal 111, 295-321.

Trust and Religion: Experimental Evidence from Bangladesh

Olof Johansson-Stenman, Minhaj Mahmud, and Peter Martinsson
Department of Economics, Göteborg University
Vasagatan 1, Box 640, SE- 405 30, Göteborg, Sweden[*]

Abstract

Trust is measured using both survey questions and a standard trust experiment using a random sample of individuals in rural Bangladesh. We found no significant effect of the social distance between Hindus and Muslims in the trust experiment in terms of fractions sent or returned, but the responses to the survey questions indicate significant differences: Hindus, the minority, trust other people less in general, and Hindus trust Muslims more than the other way around.

Key words: social capital; trust; social distance; religion; trust game; field experiment; Bangladesh

JEL classification: C93, Z13

[*] The authors wish to thank Fredrik Carlsson and Håkan Holm, and seminar participants at Göteborg University and Umeå University and participants at the Nordic Conference in Development Economics 2004, for useful comments. Financial assistance from Adlerbertska Jubileumsfonden, VINNOVA and the Swedish International Development Agency (SIDA) are gratefully acknowledged.

1. Introduction

Whom shall I trust? This is a question that most of us ask ourselves almost on a daily basis. Trust in this sense refers to our expectation regarding the consequences of making ourselves vulnerable to subsequent actions and potential exploitation by someone else. At the social level there is much evidence that trust between people reduces transaction costs, fosters cooperation, and is hence important for economic and social development; see e.g. Fukuyama (1995), Knack and Keefer (1997), Zack and Knack (2001) Beugelsdijk et al. (2004) and Bohnet et al. (2005).

Fukuyama (1995) argues that in each culture or society there is a boundary of trust, such that people in relationships within that boundary are trusted, and thus considered to be trustworthy, to a much larger extent than people outside that boundary. Easterly and Levine (1997) showed that the degree of ethnic diversity, in terms of an ethnolinguistic fractionalization index, can explain much of the observed cross-country differences in pro-growth policies as well as political stability.

Both hypothetical trust questions related to social distance as well as to the trust experiment have been applied to analyse this issue with mixed results (e.g. Buchan and Croson, 2004, Buchan et al., 2004, and Glaeser et al., 2000). A slightly different approach is applied in an empirical analysis by Alesina and Ferrara (2002), who found that belonging to a minority, which is often a group that has historically been discriminated against, is associated with having low trust, whereas religious beliefs and ethnic origins *per se* do not significantly affect levels of trust. At the beginning of the

20th century, Max Weber argued that religion can have both a positive and a negative impact on economic growth, through its impact on social organisation.[1]

The use of survey-based measures of trust in order to explain differences in social functioning are being used increasingly frequently in the social sciences. However, economists have historically preferred to rely exclusively on observed revealed behavior, and hence they have been reluctant to use self-reported survey questions whose validity has been questioned (see e.g. Bertrand and Mullainathan, 2001). However, all methods have their problems and how best to measure trust in a society is still a debated question. In this paper, we combine standard trust survey questions and a trust game (see e.g. Berg et al., 1995) in rural Bangladesh, where we target the general population. It investigates the effect of social distance[2] (based on religious belonging) on trust and trustworthiness both within and between Muslims and Hindus. Thus, there are four different combinations that will be investigated.

In Bangladesh, there are two main religions; Islam, which is the dominating state religion to which 88% of the population belong, Hinduism which accounts for about 11% of the population, whereas the remaining 1% mainly consists of Christians and Buddhists. However, 98% of the population is Bengali, i.e. most of the population comes from the same ethnic group.[3] The Bangladeshi society of today is fairly mixed with Muslims and Hindus living together in many villages. However, since Bangladesh's independence from Pakistan in 1971, socio-economic distress as well as the lack of democratic governance has been contributing to an assault on minorities,

[1] Weber attributes the emphasis on good work and salvation in Protestant Ethics as the source for development.

[2] We use the term "social distance" broadly in the same way as it is defined by the Encyclopedia of Psychology (2000): "the perceived distance between individuals and groups".

[3] The remaining ethnic groups consist of tribal groups as well as Biharis, who are non-Bengali Muslims.

from time to time, because it has created, to varying extents, apprehension and alienation among the various communities, particularly in rural areas. More recently, there has been growing evidence of attacks on Hindu communities in the aftermath of the 2001 general election that brought a coalition of a nationalist and a pro-Islamic party into power (Guhathakurta, 2002, 2004), creating further tension and possible distrust. This sense of insecurity and uncertainty may make people mistrustful and hesitant about long term-exchange relations that are often emphasised as being important for social development. Similar tensions between groups of people are found in many, perhaps even most, developing countries.

Moreover, Bangladesh is a particularly interesting country in its own right for the study of trust because it has been ranked as the most corrupt country in the world for the last 4 years (2001 to 2004) in Transparency International's corruption perception index, which is based on several polls and surveys measuring the degree to which corruption is perceived to exist among public officials and politicians.[4] Given that public officials and politicians are perceived as being corrupt, this may also influence lower levels of society. As argued, for example, by Alesina and Ferrara (2002), trust in existing institutions may therefore affect trust in other people. Rothstein and Stolle (2001) hypothesised that the development of institutional characteristics such as corruption is the most important factor of the spread of distrust and general suspicion in a society, in contrast to Putnam (1993) who argued that trust largely develops through people's interactions in local voluntary organisations.

[4] This is a composite index, which reflects the views of business people and analysts from around the world, including experts who are resident in the respective countries. Corruption is defined as the abuse of public office for private gain, e.g. bribe-taking by public officials in public procurement. (Transparency International 2003, 2004).

The objective of this paper is to test whether individuals are less inclined to trust others who are different from themselves in terms of the two main religions of Bangladesh, i.e. Islam and Hinduism. Perhaps somewhat surprisingly, our results from the trust game do not indicate any statistically significant differences based on religion in the trust experiment. However, we do find that Hindus, i.e. people from the minority religion, trust significantly less according to the survey responses, which is consistent with the finding of Alesina and Ferrara (2002). From these responses we also found that Hindus trust Muslims more than Muslims trust Hindus. Moreover, we also find a low level of stated trust in general, consistent with a hypothesis that corruption creates distrust between people. However, the fractions sent in the trust game are quite similar to most previous studies in developed (and much less corrupt) countries, and therefore do not provide any support for this hypothesis.

The paper is organised as follows: Section 2 presents a brief review of the literature where the effects of social distance on trust are measured using trust games. Section 3 presents our survey and experimental design, Section 4 presents the main results from both the survey and the trust game and Section 5 provides the corresponding econometric analysis. Section 6 summarises and concludes the paper.

2. Trust games and social distance

Participants in a typical trust experiment are anonymous and unknown to each other as well as being divided into two groups. These two groups contain participants who are either defined as "senders" or as "receivers" respectively. The sender is assigned a certain amount of money and must decide how much of the given money to send to the

anonymous receiver, and how much to keep.[5] Any positive amount sent by the sender is, in general as well as in our experiment, tripled before it is given to the receiver. The receiver then decides how much of the total amount of money received, i.e. of the tripled amount of money sent by the sender, to transfer back to the sender. With perfect information, the theoretical sub-game perfect prediction of this game is that the sender should send zero to the receivers, since one should realise that the receiver has no incentive in sending anything back. However, a Pareto improvement is possible by sending some or all of the money, if the receiver returns at least one third of the tripled amount received. The sub-game perfect prediction has not been found at the average level in conducted trust experiments, although single participants do send and return zero amounts (see e.g. Cardenas and Carpenter, 2004 and Camerer, 2003). Thus, the structure of the experiment allows the sender to use trust in order to achieve an improvement over the sub-game perfect outcome. The amount sent by the sender is typically regarded as an indication of trust, and the amount returned by the receiver is typically regarded as an indication of trustworthiness.

Previous results for religion and ethnicity tests in a trust experiment setting are mixed. Glaeser et al. (2000) measure social distance by demographic similarities and they found no significant differences in the levels of trust, as measured by the amount sent, with a partner of different race or nationality. However, they did find that people were less trustworthy with such partners, choosing to send back less.[6] Based on Jewish

[5] In the original trust game by Berg et al. (1995), the receiver was also given the same initial amount of money. The procedure adopted in this paper follows e.g. Glaeser et al. (2000), where the receiver was not given any initial money in the trust game.

[6] Another approach to measuring social distance in an experiment is by focusing on the degree of anonymity between the experimenter and the subject (e.g. Hoffman et al, 1996), or between the subjects (e.g. Dufwenberg and Muren, 2005).

Israeli students, Fershtman and Gneezy (2001) found, a mistrust of men of Eastern origin. Holm (2001), however, following a similar design to that of Fershtman and Gneezy, did not find any statistically significant discrimination effects in trust among students with different ethnic backgrounds in Sweden. Fersthman, Gneezy and Verboven (2002) found that Flemish and Valloon students in Belgium trusted each other less than they trusted students of their own group, but that they were no less trustworthy to students of the other group. They also found that students at an ultra-orthodox institution in Israel trusted students from a secular institution less than students from another ultra-orthodox institution, and vice versa. Burns (2004) found in a South-African student-based trust experiment that black receivers were considerably less well trusted than white receivers. Moreover, Willinger et al. (2003) conducted a cross-country trust experiment between French and German students. They found that neither the French nor the German students sent significantly different amounts according to whether the receiver was of their own nationality or not; the return ratios were also the same. Moreover, they also found that German students sent significantly more than French students, irrespective of the type of receiver. Buchan and Croson (2004) asked students in the USA and China to act as senders in hypothetical trust experiment questions with different imaginary receivers with varying degrees of social distance to the respondent within their own country. As expected, both the USA and the Chinese students answered that they would send much more to close relatives or students they knew well in comparison to unknown students or strangers. They also found that Chinese students stated that they would send more, across all contexts, than the USA students did. Buchan et al. (2004) found, by using a real-money trust experiment that Chinese students sent more to other students than USA students did, which supports the

results in the hypothetical trust game in Buchan and Croson (2004). As Buchan and Croson (2004) point out, this response pattern is very different from the one obtained from the World Value Survey. There the Chinese trust less in general, and they trust people from other countries much less than Americans do, on average.

The non-significant effect of religion and ethnicity that is found in trust experiments when using a subject pool of university students might be a result of the fact that students from different backgrounds are studying at the same university. Moreover, during the last 10 years, the number of exchange students has increased substantially around the world, especially within Europe, which might be another reason for the results found. Few experiments have used a non-student subject pool. Barr (2004) investigated potential effects of kinship in Zimbabwe. After independence, a large number of Zimbabwean households were resettled into new villages on land previously owned and farmed by commercial farmers. As a consequence, most households in these new villages are unrelated, contrary to traditional villages. She found that senders in resettled villages sent significantly less to a receiver in their own village, than did senders in traditional villages. However, she found no significant effects on the fraction returned in the trust game. Bouckaert and Dhaene (2004), on the other hand, investigated trust among Turkish and Belgian businessmen in the Belgian town of Ghent without finding any significant differences due to ethnic origin.

3. The survey and the trust game

As far as we know this is the first study using a trust experiment to study religious discrimination based on a non-student sample. Although student samples are appropriate to analyse many tasks experimentally, on issues such as religious and ethnic

discrimination, one can question the degree to which one can generalise the results from a student sample to the general population. In this paper we therefore use a sample from the general population in rural Bangladesh to test for differences in trust and trustworthiness based on religious beliefs among senders and receivers.

The experiment was conducted in villages of five districts of the Dhaka division[7]; Netrokona, Mymensingh, Gazipur, Manikganj and Narayanganj. The trust game was conducted at the end of a household survey and it was run among household heads in these selected areas. The choice of household head as respondents in Bangladesh is due to the fact that financial decisions made within a household are generally made, or at least approved, by the household head. We matched each sender from one village with a receiver from a nearby village, where typically 55%-75% of the households are Muslim. In the trust experiment, the participants were clearly informed about the religious identity of the person with whom they had been matched.

As discussed in Holm and Danielsson (2005), there is a risk of self-selection into a trust experiment if participation is voluntary (e.g. recruitment by posters) and this might result in an over-representation of relatively trusting and trustworthy participants in the experiment. Therefore we chose a random sample strategy. Upon arrival at a village, the enumerators were allocated to different parts of it. They were then asked to conduct the household survey and to run the trust experiment in every fourth household with the household head.[8]

[7] Bangladesh consists of six divisions, with each division being made up of several districts. In total, Bangladesh has 64 districts, 16 of which are located in the Dhaka division.

[8] If the household head was not around, the enumerators were instructed to go back later. If a selected respondent was not at home during the second visit, the enumerator moved to the next neighbouring household. In the villages, people from the same family-chain normally live in a cluster of say 4-5 households and therefore a replacement from the next household, or in some cases the next but one

In the sender's household, the enumerator requested to conduct the trust experiment in private and free from any interruptions and they ensured the confidentiality of the responses. Then the enumerator began to read the instructions to the sender. The instructions included examples of the experiment presented as the outcomes of different decisions made by the sender and the receiver as well as the religion of the receiver. The senders were also informed, within the instructions that they would be paid within three days. The senders were then given two envelopes. One of them contained their original endowment of 200 Bangladeshi Taka and the other one was empty.[9] The senders used the empty envelope to send the amount of money that they decided upon to the receiver. In the experiment we used thick envelopes in order to prevent enumerators guessing by eye how much had been sent to the receiver, and thus implicitly knowing the amount kept. The enumerator ensured that the decision was made in private by turning his back to the sender while the money was being put into the envelope. The enumerator waited until the sender was ready with the envelopes. The sender was then asked to close the envelope that was to be sent to the receiver and seal it with a stamp that had been provided before returning it to the enumerator. The sender was instructed to do so even if he/she had decided to send nothing. The senders were assured that the enumerator would not know their decisions, as they would not open the envelope. At the end of the day, the enumerator handed the envelopes to the principal researcher, who opened them and put the tripled amount into new envelopes with pre-matched household codes.

---

household, should not bias the results. Replacement households form approximately 23 percent of our sample. Only 2% of available householders did not want to participate in the survey. Two days at most were spent in a single village due to the risk that people would start talking about the experiment.

[9] 57.8 Taka=1 USD at the time of the experiment.

The following day the enumerators were given these new envelopes to be delivered to the assigned receivers. After the instructions for the experiment had been read to them, including the same examples and information that were given to the senders, they received the envelope containing the tripled amount from the sender as well as an empty envelope. The enumerator then turned his back to the receiver who had been instructed to put the amount of money that the receiver wanted to return to the sender into the empty envelope. After putting the chosen amount in the envelope, the receiver closed and sealed it with a stamp that had been provided by the enumerator. At the end of the day, the enumerator returned the envelopes to the principal researcher, who checked and wrote down the amount to be transferred back. During the following day, i.e. on the third day, the envelopes were delivered back to the senders. Finally, all respondents were explicitly asked not to discuss the sums of money that they had earned from the experiment with anyone else.[10] There is, of course, always a potential risk of distrust towards the people and organization running the experiment. In order to minimize this, university students were used as enumerators, since the university is generally considered to be a trustworthy institution in Bangladesh (compared to many NGOs for example). Moreover, it was specifically mentioned that this was a joint research project being run by a local and a Swedish university.

4. Results

In Table 1 we show the average fractions sent and transferred back as well as the proportions of zeros for the whole sample as well as for each sub-sample of the

[10] The complete instructions are available upon request from the corresponding author.

experiment. In Appendix 1, we present the histograms on proportions sent and proportions returned by different sub-samples.

Table 1. Average proportion sent and returned by sub-samples based on religion.

	Total	Muslim sender Muslim receiver	Muslim sender Hindu receiver	Hindu sender Muslim receiver	Hindu sender Hindu receiver
Average proportion sent	0.46	0.46	0.46	0.43	0.50
Average proportion returned	0.46	0.46	0.51	0.42	0.44
Proportion of zero sent	0.06	0.06	0.05	0.10	0.08
Proportion of zero returned	0.06	0.03	0.03	0.05	0.07

The average amount sent, 92.2 Bangladeshi Taka, is about 46% of their initial endowment of 200 Taka. The average amount returned was 134.1 Bangladeshi Taka, which equals a return ratio of 46%. The amount returned is, on average, in excess of the amount sent and thus it is, on average, profitable for the sender to send money. In our case, 38% of the senders gained from sending money to the receivers. This magnitude is in-line with the findings of Cardenas and Carpenter (2004), who summarize trust experiments conducted in developing countries and countries undergoing transition. They found that senders on average have benefited from trusting (i.e. sending money) in 17 out of the 25 trust games reported. Moreover, out of the 256 senders who participated in the experiment, 18 senders (7%) sent nothing while 46 senders (18%) sent everything. Out of 237 receivers[11], 11 receivers (5%) sent back nothing while 9 receivers (4%) sent back everything to the senders. As shown in Table 1, there are generally fairly small differences in both the fractions sent and those returned with

[11] One receiver refused to take part in the experiment.

respect to the religion-based sub-groups. Based on non-parametric tests, we cannot reject the null hypothesis that the proportions sent in all sub-groups come from the same population using a joint Kruskal-Wallis test. Furthermore, we cannot reject the null hypothesis that two sub-samples come from the same underlying population at conventional levels using Wilcoxon-Mann-Whitney tests for all possible pair-wise combinations.[12] The corresponding null hypothesis for the fractions returned among the receivers cannot be rejected for any pair either (see Appendix for details).

Thus, we find no differences in trust, as measured by the fractions sent in a trust game, due to religious allegiance. Moreover, the fact that we find no difference in trustworthiness either indicates that people, on average, are correct in their judgement that there are no systematic differences in trustworthiness based on religious allegiance.

Using the most frequently used standard GSS trust question, "Generally speaking would you say that most people can be trusted or that you cannot be too careful in dealing with people?", we find that only about 3 percent choose the alternative that most people can be trusted, which seems to indicate a very low level of trust with the corresponding figures for Muslims and Hindus being 2% and 4% respectively. Since this measure, for obvious reasons, is very crude, we also asked a 6 level question on the degree to which they agreed to the statement "most people can be trusted" where they were informed that 1 corresponds to strongly disagree and 6 corresponds to strongly agree. The results are shown in Table 2. Even though these responses are less extreme, they still reflect rather low levels of general trust. We also asked the same question concerning people from their own religions as well as from

[12] Both tests are non-parametric, i.e. distribution-free tests, and are thus appropriate here since it is difficult to make any a priori assumption about the distributions; see e.g. Siegel and Castellan (2000) for a description of these tests.

others and find that, on average, people trust people from other religions less than they trust people from their own religion. This holds both for Muslims and Hindus although the differences appear to be somewhat smaller for Hindus. Based on nonparametric Wilcoxon-Mann-Whitney test, we find no statistical differences between Hindus and Muslims for their trust of people of the same religion (p-value=0.27), while general trust as well as trust of people of another religion are statistically significant at 1% level between the religious groups.

Table 2. Stated trust as the level of agreement with the following statements

Level of agreement with the statement:	Most people can be trusted	Most people from own religion can be trusted	Most people from other religions can be trusted
		Whole sample (N=512)	
Strongly disagree	13%	1%	13%
Disagree	25%	10%	23%
Partly disagree	31%	19%	28%
Partly agree	14%	19%	15%
Agree	8%	23%	8%
Strongly agree	9%	27%	13%
		Muslims (N=256)	
Strongly disagree	10%	2%	20%
Disagree	22%	9%	28%
Partly disagree	34%	20%	24%
Partly agree	13%	15%	13%
Agree	10%	25%	6%
Strongly agree	11%	29%	9%
		Hindus (N=256)	
Strongly disagree	16%	1%	6%
Disagree	29%	10%	19%
Partly disagree	29%	19%	32%
Partly agree	14%	24%	16%
Agree	6%	22%	10%
Strongly agree	6%	24%	16%

5. Econometric analysis

Table 3 defines the explanatory variables used in the econometric analysis and presents their mean values.

Table 3. Sample statistics

Variable	Definition	Mean	Min	Max	N
Muslim sender-Muslim receiver	Muslim sender is matched with Muslim receiver	0.252	0	1	512
Muslim sender-Hindu receiver	Muslim sender is matched with Hindu receiver	0.246	0	1	512
Hindu sender-Muslim receiver	Hindu sender is matched with Muslim receiver	0.252	0	1	512
Hindu sender - Hindu receiver	Hindu sender is matched with Hindu receiver	0.250	0	1	512
Hindu religion	The religion of the respondent is Hinduism	0.50	0	1	512
Muslim religion	The religion of the respondent is Muslim	0.50	0	1	512
Age	Age of the respondent in years	44.7	19	87	512
Illiterate	Cannot read and write	0.28	0	1	512
Low education	Literate or education up to high school level	0.57	0	1	512
High education	Education above high school level	0.15	0	1	512
Income per-capita	Annual household income (in Taka) adjusted with equivalence and economies of scale. Total yearly household income was divided by [(number of adults + 0.5× number of children)$^{0.75}$×100000]	0.23932	0.009	3.64	511
Stated trust	Level of agreement with the statement that most people can be trusted (1= strongly disagree, 2= disagree, 3=partly disagree, 4=partly agree, 5=agree, and 6 = strongly agree).	3.05	1	6	512
Trusting behavior	Frequency of lending money to friends and neighbours: 1=once a year or less, 2= about once every other month, 3= about once a month, 4= about once a week, 5= more than once a week.	1.81	1	5	512
Confidence index	Arithmetic sum of confidence on 10 institutions [Banks, NGOs, Military, Police, Judiciary, Local government, Executive Government, Political parties, Rural power elites, Educational institutions]: great deal of confidence=2, only some confidence=1, and hardly any confidence at all=0.	14.6	2	20	508
Religious participation	The respondent prays at least once a day.	0.67	0	1	511
Recent misfortune	The respondent has been victim to any of the following incidents in the last year: robbery /theft, mugging, personal assault, home attack, land fraud, false criminal accusation, or political harassment.	0.218	0	1	511
Membership of voluntary association	Has a membership in a voluntary group and/or association.	0.291	0	1	509

In Table 4 we present the estimates from OLS regressions to explain factors that may influence the fractions sent by the senders, as well as the fractions returned by the receivers.[13]Since we used several enumerators to conduct the fieldwork, we test whether there is an enumerator effect or not. We cannot reject the hypothesis of no enumerator effect in the trust game regressions in Table 4 (p-value 0.53 and 0.21, respectively, for Model 1 and Model 2, based on joint F-tests. However, for the remaining regressions on stated trust, we can reject the hypothesis of no enumerator effect at a 5% significance level, and therefore we included dummy variables for the enumerators to control for the enumerator effect.

As is clear from Table 4, we found again no significant influence of religious allegiance either on the proportion sent or on the proportion returned, and none of the dummy variables for religious status were significantly different from zero. We also conducted *F*-tests, in order to test whether these three dummy variables were jointly significantly different from zero or not. Again, we cannot reject the null hypothesis of no differences between the sub-groups for either the fraction sent (p-value = 0.46) or the fraction returned (p-value = 0.56).

[13] We also estimated a separate model for proportion returned, where each sub-group's reaction to the proportion sent is analysed by interacting proportion sent with dummy variables for different sender versions. We do not find any significant effect, however.

Table 4. Regression analysis: Proportions sent and returned in the trust game, and stated trust.

Dependent variable	Proportion sent	Proportion returned	Stated trust in general	Stated trust in people of own religion	Stated trust in people of other religion
Regression Type	OLS	OLS	Ordered probit [a]	Ordered probit [a]	Ordered probit [a]
Fraction sent		-0.037			
		(0.078)			
Muslim sender- Hindu receiver	-0.015	0.056			
	(0.056)	(0.067)			
Hindu sender- Muslim receiver	-0.030	0.026			
	(0.056)	(0.065)			
Hindu sender- Hindu receiver	0.054	-0.028			
	(0.056)	(0.064)			
Hindu religion			-0.361***	-0.254**	0.513***
			(0.100)	(0.101)	(0.101)
Age	0.016*	-0.007	-0.031	-0.006	0.003
	(0.010)	(0.011)	(0.023)	(0.024)	(0.023)
Age squared	-0.0001	0.0001	0.0003	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)
Illiterate	-0.115	-0.104	0.268	-0.150	0.060
	(0.073)	(0.080)	(0.173)	(0.175)	(0.174)
Low education	-0.037	-0.057*	-0.104	-0.070	0.175
	(0.064)	(0.069)	(0.149)	(0.151)	(0.150)
Income per-capita	0.212***	0.029	0.051	-0.010	0.100
	(0.073)	(0.061)	(0.153)	(0.150)	(0.154)
Stated trust	0.021	0.060***			
	(0.015)	(0.018)			
Trusting behavior	0.031	-0.008	0.058	-0.048	0.137***
	(0.019)	(0.022)	(0.046)	(0.047)	(0.047)
Confidence index	-0.003	0.002	0.045***	0.077***	0.032**
	(0.006)	(0.007)	(0.015)	(0.015)	(0.015)
Religious participation	-0.027	-0.085	-0.124	-0.174	-0.210*
	(0.045)	(0.052)	(0.110)	(0.111)	(0.110)
Member of voluntary association	-0.070	0.013	0.002	-0.100	0.167
	(0.044)	(0.050)	(0.107)	(0.110)	(0.110)
Recent misfortune	-0.046	0.087	-0.307***	0.054	0.011
	(0.045)	(0.058)	(0.115)	(0.116)	(0.114)
Constant	0.230	0.500			
	(0.272)	(0.291)			
R^2 [Pseudo R^2]	0.146	0.165	0.100	0.100	0.100
No. of observations	251	233	502	501	501

Standard errors are in parenthesis. Superscripts *, **, *** denote statistical significance at 10%, 5%, and 1% level, respectively.

[a]We control for enumerator effects, but the coefficients are omitted from the presentation.

The amount sent increases with age and this effect is significant at a 10% level. Thus, the age profile of the amount sent is similar to the findings by Kocher and Sutter (2003), who report an age pattern where the amount sent increases from childhood to early adulthood, but stays almost constant thereafter. However, our findings are in contrast to Bellemare and Kröger (2003) and Gleaser et al. (2002), as well as to the age pattern of stated trust reported in Putnam (2000), and to Fehr et al. (2003) who found that older people send significantly less. We also find a strong positive effect of income. As with Glaeser et al. (2000), we found that stated trust predicts trustworthiness, measured as the fraction returned, much better (1% significance level) than it predicts trust, as measured by the fraction sent.

These results can be compared to those obtained from the ordered probit regressions of the *stated* trust 6-point scale survey questions, presented in the last three columns in Table 4. The results show that Hindus in general trust less, consistent with the findings of Alesina and Ferrara (2002) that minorities trust less. Hindus also trust other Hindus significantly less than Muslims trust other Muslims, contrary to the finding of the non-parametric tests reported in Section 3. It is perhaps more surprising that Hindus trust Muslims more than Muslims trust Hindus. Given that trust increases with interaction, the pattern may be explained in part by the fact that, in general, Hindus are more or less forced to interact with Muslims more than Muslims have to interact with Hindus. [14]

We find that an index of confidence in institutions positively and significantly influences stated trust, which is consistent with the top down perspective of Rothstein

[14] However, we cannot rule out that this result may in part be driven by an enumerator effect. Since the enumerators were all (except one) Muslims, possible attempts to please the enumerators would then clearly bias the result; cf. Bertrand and Mullainathan (2001).

and Stolle (2001), while trusting behavior only increases stated trust significantly in the case of trust in people from other religions.

6. Discussion and conclusion

We find no significant evidence that religious allegiance affects the level of trust or trustworthiness in a trust experiment conducted in rural Bangladesh, as measured by the proportions sent and returned, respectively. This may simply reflect that social distance with regard to religious belief does not matter for trust and trustworthiness in rural Bangladesh, or that it matters only to a small degree. However, the survey data provides a very different picture where Hindus trust less in general, Hindus trust other Hindus less than Muslims trust other Muslims, and Hindus trust Muslims more than Muslims trust Hindus. There are also differences between the methods regarding how one would interpret the average degree of trust. An obvious question, then, is which result should we believe in, or trust?

Some analysts, such as Glaeser et al. (2000), seem to take it as self-evident that trust experiments are superior to using survey questions, since the latter are not consequential in terms of real money. We are less certain, however. One advantage to using survey questions is that they measure concerns about trust more directly, because this is what they explicitly ask for, however noisy and biased the measurement may then be. Behavior in a trust experiment, on the other hand, can have several different driving forces. For example, Cox (2004) presented evidence that behavior in trust games partly measures other-regarding preferences, whereas Karlan (2005) found that it largely measures risk preferences, rather than trust. On the other hand, Eckel and Wilson (2004)

found no significant relationship between the decision to send money in a simplified trust game and two behavioral risk measures used.

On the basis of the results here, we cannot a priori argue that one way to measure trust is better than the other. It is also possible that they measure different aspects of trust. What can be said so far is that for a random sample in rural Bangladesh, the two most frequently used methods to measure trust give very different results, in part confirming the discussed findings by Buchan and Croson (2004) and Buchan et al. (2004). Consequently, it is an important task for future research to provide better insights into how to interpret the results of the different measures, and possibly also into how to develop better trust measures and methods.

References

Alesina, A., and E. La Ferrara, 2002, Who trusts others? Journal of Public Economics 85, 207-234.

Barr, A., 2004, Kinship, Familiarity, and Trust: an Experimental Investigation. In Henrich, J. R. Boyd, S. Bowles, C. Camerer, E. Fehr and H. Gintis (eds.): Foundations of Human Sociality, Oxford: Oxford University Press.

Bellemare, C. and S. Kroger, 2003, On Representative Trust, Center Discussion Paper Series No. 2003-47

Berg, J., Dickhaut, J., McCabe, K., 1995, Trust, Reciprocity, and Social History, Games and Economic Behavior, 10, 122-42.

Bertrand, M. and Mullainathan, S., 2001, Do People Mean What They Say? Implications for Subjective Survey Data, American Economic Review, Papers and Proceedings, 91, 67-72.

Beugelsdijk, S., Groot, H.L.F. de, Schaik, A.B.T.M. van, 2004, Trust and Economic Growth: A Robustness Analysis, Oxford Economic Papers 56, 118-134.

Bohnet, I., Harmgart, H., Huck, S. and Tyran, J.-R., 2005, Learning Trust, Journal of the European Economic Association, forthcoming.

Bouckaert, J. and G. Dhaene, 2004, Inter-Ethnic Trust and Reciprocity: Results of an Experiment with Small Businessmen, European Journal of Political Economy 20, 869-86.

Buchan, N. and R. Croson, 2004, The Boundaries of Trust: Own and Others' Actions in the US and China, Journal of Economic Behavior and Organization 55, 485-504.

Buchan, N., E. Johnson and R. Croson, 2004, Getting to Know You: An International Experiment of the Influence of Culture, Communication, and Social Distance on Trust and Trustworthiness, Journal of Economic Behavior and Organization, forthcoming.

Burns, J., 2004, Rebuilding Trust in Post-Apartheid South Africa, Mimeo, University of Cape Town.

Camerer, C., 2003, Behavioral Game Theory: Experiments on Strategic Interaction, Princeton University Press.

Cardenas, J.C. and Carpenter, J., 2004, Experimental Development Economics: A Review of the Literature and Ideas for Future Research, mimeo.

Cox, J., 2004, How to Identify Trust and Reciprocity, Games and Economic Behavior 46, 260-281

Dufwenberg, M. and Muren, A, 2005, Generosity, Anonymity, Gender, Journal of Economic Behavior and Organization, forthcoming.

Easterly, W. and R. Levine, 1997 Africa's Growth Tragedy: Policies and Ethnic Divisions. Quarterly Journal of Economics, 112, 1203-1250.

Eckel, C. C. and R.K. Wilson, 2004, Is Trust a Risky Decision? Journal of Economic Behavior and Organization 55, 447-65.

Encyclopaedia of Psychology, 2000, (Eds) Alan E Kazdin, Oxford University Press.

Fukuyama, F., 1995, Trust, New York: Free Press.

Fehr, E., U. Fischbacher, B. von Rosenbladt, J. Schupp and G. G. Wagner, 2002, A Nation-Wide Laboratory: Examining Trust and Trustworthiness by Integrating Behavioral Experiments into Representative Surveys, Schmollers Jahrbuch 122, 519-542.

Fershtman, C. and U. Gneezy, 2001, Discrimination in a Segmented Society" Quarterly Journal of Economics 116, 351-77.

Fershtman, C., U. Gneezy and F. Verboven, 2002, Discrimination and Nepotism: The Efficiency of the Anonymity Rule, mimeo.

Glaeser, E. L., D.I. Laibson, J.A. Scheinkman, C.L. Soutter, 2000, Measuring Trust, Quarterly Journal of Economics 115, 811-846

Guhathakurta, M., 2002, Assault on Minorities in Bangladesh: An Analysis, Meghbarta 3(2), http://www.meghbarta.org

Guhathakurta, M., 2004, Minority Rights, Identity Politics and Gender in Bangladesh: Current Problems and Issues, Meghbarta 5(2), http://www.meghbarta.org

Hoffman, E., K. McCabe, and Smith, V., 1996, Social distance and Other-regarding Behavior in Bargaining Games, American Economic Review 86, 653-660.

Holm, H. J., 2001, What's in a Name? An Ethnical Discrimination Experiment, Mimeo, Department of Economics, Lund University, Sweden.

Holm, H. J. and Danielsson, A., 2005, Tropic Trust versus Nordic Trust: Experimental evidence from Tanzania and Sweden, Economic Journal, forthcoming.

Karlan, D. S., 2005, Using Experimental Economics to Measure Social Capital and Predict Financial Decisions, American Economic Review, forthcoming.

Knack, S., Keefer, P., 1997, Does Social Capital Have an Economic Payoff? A Cross-Country Investigation. Quarterly Journal of Economics 112, 1251-1288.

Siegel, S. and Castellan, N. J., 2000, Nonparametric Statistics for the Behavioral Sciences, Second edition, McGraw-Hill.

Transparency International, 2003, Transparency International Corruption Perception Index 2003 Press Release, http://www.transparency.org

Transparency International, 2004, Transparency International Corruption Perception Index 2004, Press Release, http://www.transparency.org

Willinger, M., Keser C., Lohmann C. and Usunier J.-C., 2003, A Comparison of Trust and Reciprocity between France and Germany: Experimental Investigation Based on the Investment Game, Journal of Economic Psychology 24, 447-466.

Zak, P.J. and Knack, S., 2001, Trust and Growth, Economic Journal 111, 295-321.

Appendix 1

Histograms on the proportions of money sent and returned by different sub-samples

Figure 1. Proportion of money sent by Hindu senders to Muslim receivers

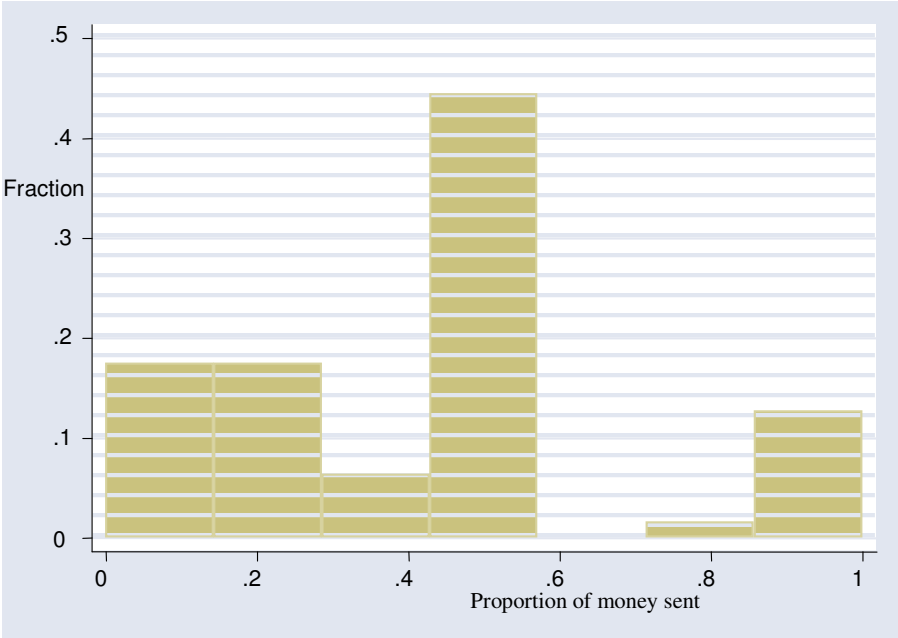


Figure 2. Proportion of money sent by Hindu senders to Hindu receivers

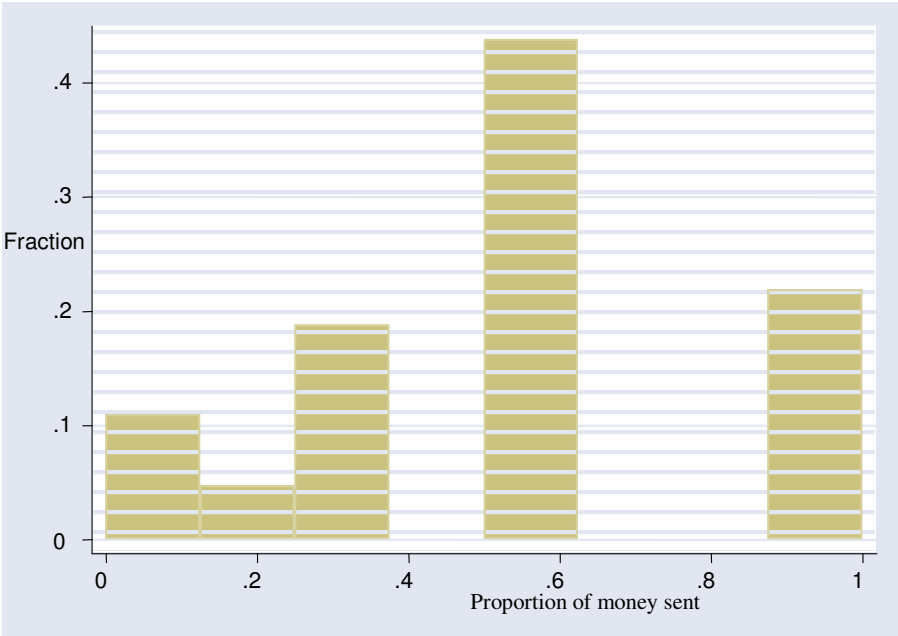


Figure 3. Proportion of money sent by Muslim senders to Muslim receivers

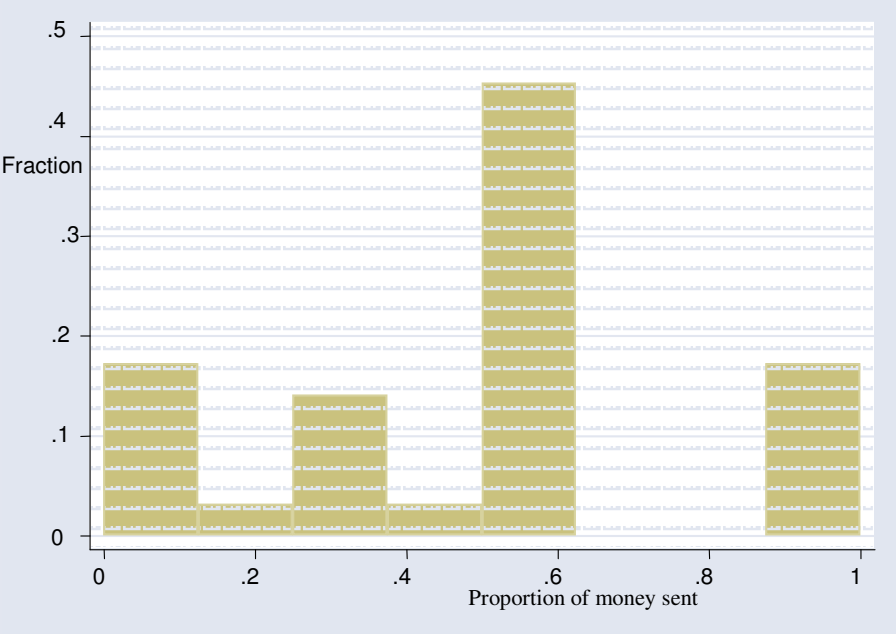


Figure 4. Proportion of money sent by Muslim senders to Hindu receivers

Figure 5. Proportion of money returned by Muslim receivers to Hindu senders

Figure 6. Proportion of money returned by Muslim receivers to Muslim senders

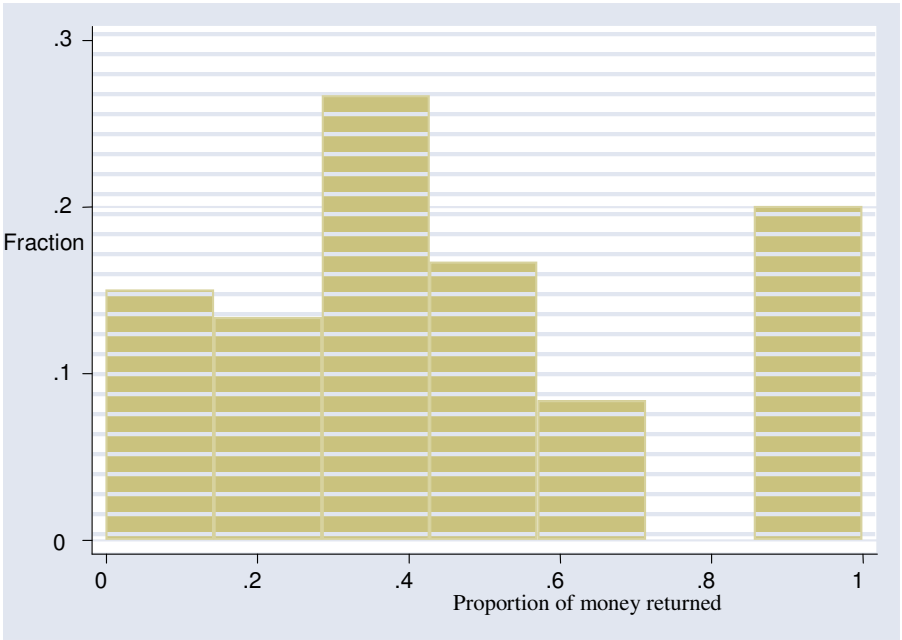


Figure 7. Proportion of money returned by Hindu receivers to Muslim senders

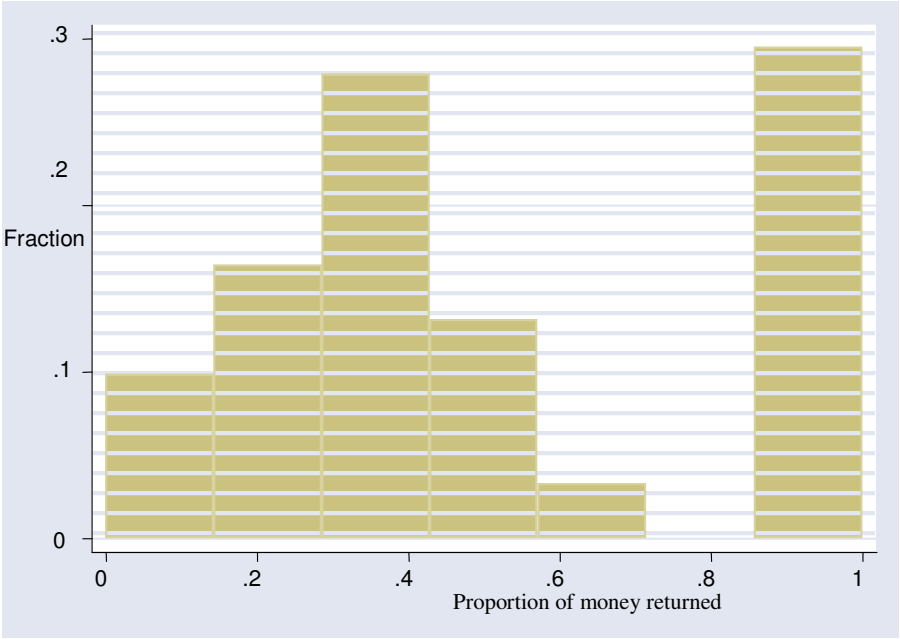
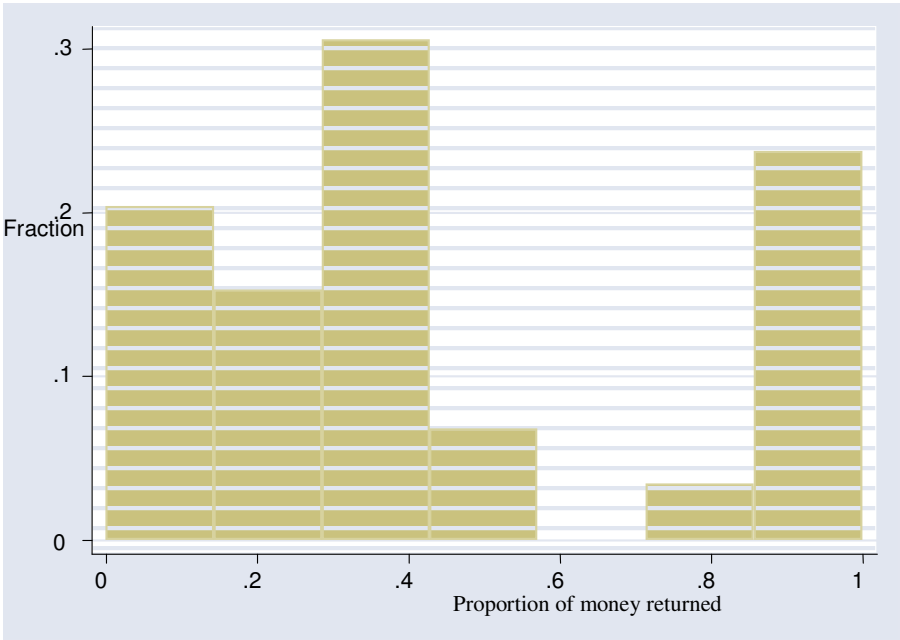


Figure 8. Proportion of money returned by Hindu receivers to Hindu senders



Appendix 2

2. A. Wilcoxon and Mann-Whitney test and Kruskal Wallis test of difference between the proportions sent by different groups

Hypothesis	p-value
Proportion sent by Muslim to Muslim = Proportion sent by Muslim to Hindu	0.814
Proportion sent by Muslim to Muslim = Proportion sent by Hindu to Muslim	0.602
Proportion sent by Muslim to Muslim = Proportion sent by Hindu to Hindu	0.590
Proportion sent by Hindu to Hindu = Proportion sent by Hindu to Muslim	0.310
Proportion sent by Hindu to Hindu = Proportion sent by Muslim to Hindu	0.410
Proportion sent by Muslim to Hindu = Proportion sent by Hindu to Muslim	0.871
Proportion sent by Muslim to Muslim = Proportion sent by Muslim to Hindu = Proportion sent by Hindu to Muslim = Proportion sent by Hindu to Hindu	0.760

2. B. Wilcoxon and Mann-Whitney test and Kruskal Wallis test of difference between the proportions returned by different groups

Hypothesis	p-value
Proportion returned by Muslim to Muslim = Proportion returned by Hindu to Muslim	0.427
Proportion returned by Muslim to Muslim = Proportion returned by Muslim to Hindu	0.557
Proportion returned by Muslim to Muslim = Proportion returned by Hindu to Hindu	0.400
Proportion returned by Hindu to Hindu = Proportion returned by Hindu to Muslim	0.804
Proportion returned by Hindu to Hindu = Proportion returned by Muslim to Hindu	0.153
Proportion returned by Muslim to Hindu = Proportion returned by Hindu to Muslim	0.183
Proportion returned Muslim to Muslim = Proportion returned by Muslim to Hindu = Proportion returned by Hindu to Muslim = Proportion returned by Hindu to Hindu	0.426

Trust, Trust Games and Stated Trust: Evidence from Rural Bangladesh

Olof Johansson-Stenman, Minhaj Mahmud and Peter Martinsson
Department of Economics, Göteborg University
Vasagatan 1, Box 640, SE- 405 30 Göteborg, Sweden[*]

Abstract

Levels of trust are measured by asking standard survey questions on trust and by observing the behavior in a trust game using a random sample in rural Bangladesh. Follow-up questions and correlations between the sent amount in the trust game and stated expectations reveal that the amount sent in the trust game is a weak measure of trust. The fear of future punishment, either within or after this life, for not being sufficiently generous to others, was the most frequently stated motive behind the respondents' behavior, highlighting the potential importance of motives that cannot be inferred directly from people's behavior.

Key words: Trust; trust game; social capital; field experiment; Bangladesh.

JEL classification: C93, Z13

[*] The authors wish to thank Fredrik Carlsson, Håkan Holm and seminar participants at Göteborg University, Umeå University and the 2nd Budapest Workshop in Behavioral Economics for useful comments. Financial assistance from Adlerbertska Jubilieumsfonden, VINNOVA and the Swedish International Development Agency (SIDA) are gratefully acknowledged.

1. Introduction

There is much recent theoretical and empirical evidence that trust between people fosters co-operation and economic activity, hence it is also important for economic and social development; see e.g. Fukuyama (1995), Knack and Keefer (1997) and Bohnet et al. (2005). Not surprisingly, the interest in measuring trust, as well as in explaining how trust itself develops, has therefore also increased. Empirical studies that analyse the concept of trust have tried both to measure and to explain what determines trust per se (e.g. Alesina and Ferrara, 2002; Glaeser et al., 2000) as well as to use measured trust as a variable to explain economic outcomes (e.g. Beugelsdijk et al., 2004, and Knack and Keefer, 1997). The most frequently used question for attitudinal trust is the General Social Survey (GSS) question framed as "Generally speaking, would you say that most people can be trusted, or that you cannot be too careful in dealing with people?" The results of such studies can be, and frequently has been, criticised since the choices made are non-consequential for the respondents.

On the other hand, the decisions made in a trust game, as introduced by Berg et al. (1995), are consequential since the decisions will have a monetary effect on the participants. Briefly, the trust game involves two stages. The participants in the game act either as a 'sender' or as a 'receiver'. The sender is given a certain amount of money and has to decide how much of this to send to the anonymous receiver, and how much to keep.[1] Any (positive) amount sent by the sender is normally tripled before it is given to the receiver. The receiver must then decide how much of the total amount of money received, (i.e. how much of the tripled amount of money) to return to the sender, after

[1] In the original trust game by Berg et al. (1995), the receiver was also given the same initial amount of money. The procedure adopted in this paper follows e.g. Glaeser et al. (2000), where the receiver was not given any initial money in the trust game.

which the game is over. The theoretical sub-game perfect prediction of this game is that the sender should send nothing to the receiver, since the sender would realize, by using backward induction, that the receiver has no incentive to send anything back. However, a Pareto improvement is possible by sending some or all of the money, if the receiver returns at least one third of the tripled amount received. Thus, the structure of the game allows the sender to use trust in order to achieve an improvement over the sub-game perfect outcome. The amount sent by the sender is typically regarded as an indication of trust, and the amount returned by the receiver as an indication of trustworthiness.

In an innovative paper, Glaeser et al. (2000) combined a trust game with an attitudinal trust survey among undergraduate students at Harvard. They found poor correlation between stated trust and the amount sent in the trust game, while the amount returned was significantly explained by stated trust. One of their conclusions was "that most work using these survey questions needs to be somewhat reinterpreted" (p. 814). This conclusion should be read in the light of their interpretation that the trust game measures trust and trustworthiness.

On the other hand, Cox (2004) discussed and tested whether there are other motives behind the amount sent and returned in a trust game, such as other-regarding preferences. Using three separate sub-samples, he conducted a different experiment in each of them. First he conducted an ordinary trust game, in which the amount sent was tripled and both the sender and the receiver were endowed with the same amount. The same endowment was then used with another sample in a dictator game, which is a game where the receiver cannot send back any of the tripled amounts received from the

sender.[2] The difference between these two games was considered to reflect pure trust. Although the amount sent was lower in the dictator game compared to the trust game (implying a non-negligible amount of pure trust), the amount sent in the dictator game was found to be non-negligible, implying evidence of unconditional other-regarding preferences. The third game mimics the second stage of a trust game by endowing the sender and the receiver with exactly the same as has been sent by a sender in the first stage of the trust game. However, the individuals in this game were not informed that their endowment was based on the first stage decisions made by another pair, but only that they would be part of a dictator game, where the person who is endowed with the receiver's amount could send any amount to the other individual as in a traditional trust game. The difference between the amount returned in the trust game and the amount sent in the last described game is considered to measure reciprocity. However, less was sent in this dictator game compared to the amount returned in the second stage of the trust game, which implies a non-negligible degree of reciprocity. Moreover, a non-negligible amount was on average sent in this dictator game too, implying further evidence of unconditional other-regarding preferences. Thus, Cox (2004) found that the average amounts sent and returned in the trust game are motivated by factors beyond pure trust and pure reciprocity, respectively. These results are also consistent with a considerable literature on social preferences, which has concluded that other motives beyond pure self interest, such as fairness considerations, are important in explaining observed behavior (e.g. Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; and Charness and Rabin, 2002).

[2] In a standard dictator game the proportion sent is not multiplied by 3. It was done in this case for the results to be comparable with trust game, in order to trace out different motives.

Holm and Danielsson (2005) conducted two experiments using university students in Tanzania and Sweden. In the Tanzanian study both a trust game and a dictator game were conducted, while in the Swedish study the trust game and the dictator game were accomplished with what was called a Trictator game. In the Trictator game, the receivers in the trust game were the senders in the dictator game, while the senders in the trust game were asked to guess the amount sent in the dictator game using an incentive compatible mechanism following the Becker-DeGroot-Marschak approach. The amount sent in the trust game is not explained by the amount sent in the dictator game in the Tanzanian sample, while the amount reported in the Trictator game significantly explained the amount sent. The latter indicates that senders who believe that the receivers have a stronger degree of inequality aversion send more in the trust game. Thus, their results indicate that there are several motives behind the behavior in a trust game. They also found that survey-trust explains the amount sent better when controlling for donation motives in the trust game.

Bellemare and Kröger (2003) and Fehr et al.(2002), using a representative sample in the Netherlands and in Germany, respectively, found that first mover's expectation about the second mover's transfer is positively correlated with the amount sent by the first mover in the trust game. Ashraf et al. (2002), using South African, Russian and American students, conclude that both senders and receivers are motivated by factors other than trust and reciprocity, while Carter and Castillo (2002), using a non-student sample in South Africa, find that the amount sent in the trust game is related to both trust and altruism. Chaudhuri and Gangadharan (2003) asked the senders whether they expected to get any money back, and how much they thought this might be. They also asked the senders to state their explicit motivations behind their transfers. They

coded the stated motivations into four different groups, and found a strong positive relationship between the amount sent and the amount expected back. They also concluded that the senders sent more money when a stronger degree of trust was expressed as motivation. Taken together, the above evidence is a bit mixed, but most results indicate that motives other than pure trust appear to be important for the observed behavior in trust games.

The main objective of this paper is to contribute to the discussion about what trust games and trust surveys actually measure. In addition to analysing the outcome of the trust game *per se* by using both descriptive and regression approaches, we asked what the senders' expectations were about how much they would eventually get back from the receivers. Moreover, we asked the participants to state what their motives were for sending and returning the amounts they sent and returned using follow-up questions after the trust game. It is unconventional within economics to ask about expectations, and it is even more unconventional and controversial to ask about motives. Of course, we do not doubt that there are non-negligible potential problems with interpreting stated motives, of which some will be discussed below. Still, as argued by Babcock and Loewenstein (1997), focusing solely on revealed behavior in markets and experiments often implies rather coarse tests of hypotheses since there are often many theories consistent with observed behavior. Moreover, we agree with Sen (1973, p.258) that "we have been too prone, on the one hand, to overstate the difficulties of introspection and communication and, on the other, to underestimate the problems of studying preferences revealed by observed behavior."

The analysis in this paper is based on a combined attitudinal trust survey and trust game conducted in the field among household heads in rural Bangladesh. There are

at least three advantages from using this sample, compared to the more frequently used choice of a student sample in a western country:[3] (i) We can afford to use high financial stakes compared to their normal wages, implying that the participants have strong incentives to treat the game seriously, and to think carefully about how to act, (ii) We obtain more variation in the socio-economic background variables and (iii) Bangladesh is a particularly interesting country in itself for the study of trust because of its top rank in the Transparency International's corruption perception index[4] implying the highest measure of corruption of public officials four years in a row (2001-2004). Given that the individuals associated with these institutions are perceived to be corrupt, this may affect lower levels in the society, and as argued e.g. by Alesina and Ferrara (2002), trust in existing institutions may affect trust in other people. In communities, where laws are well established and enforced, people may be relatively trusting because they feel well protected from non-co-operative behavior. Rothstein and Stolle (2001) hypothesised that the development of institutional characteristics, such as corruption, is the most important factor for the spread of distrust and general suspicion in a society. This top-down perspective is different from the bottom-up perspective put forward by Putnam (1993), who argued that trust develops largely through people's interactions in local voluntary organisations. We found in our study that stated trust in terms of the most

[3] Needless to say, there are disadvantages too, including lower education levels among the respondents, which may induce more cognitive errors, and is the expense and logistical difficulty of setting up a large-scale field experiment.

[4] Based on several polls and surveys, the Transparency International's CPI ranks countries in terms of the degree to which corruption is perceived to exist among public officials and politicians, which also relates to the abuse of public office for private gain, e.g. bribe-taking by public officials in public procurement. It is a composite index, which reflects the views of business people and analysts from around the world, including experts who are resident in the respective country. However, its sources do not distinguish between administrative and political corruption or between petty and grand corruption (Transparency International 2003, 2004).

frequently used General Social Survey (GSS) question framed as "Generally speaking, would you say that most people can be trusted, or that you cannot be too careful in dealing with people?" results in a strikingly low fraction, only around 3%, answered that most people can be trusted. This is consistent with the hypothesis that corruption creates distrust between people at grass-root level. On the other hand, the fraction sent in the trust game is of a similar order of magnitude as in most previous studies and thus does not provide any support for this hypothesis.

Overall, our results support and extend the recent conclusions by e.g. Cox (2004) and Holm and Danielsson (2005), that the motives behind the observed behavior in the trust games are much more complicated and mixed than often believed. For example, in our case we found that what might seem to be altruism may sometimes reflect long-term self-interest, since an important stated motive for both sending and returning money was that people believed that they would be punished, either in this life or in the after life, if they acted too selfishly in the trust game. This type of information, we believe, would have been impossible to obtain based on revealed behavior in various experiments.

The rest of this paper is organised as follows. Section 2 presents our survey and experimental design. Section 3 presents the main descriptive results from both the attitudinal trust survey and the trust game, followed by a discussion around the issue of what trust games actually measure. Section 4 provides econometric analysis while Section 5 summarises and concludes the paper.

2. The survey and the trust game

Our survey and experiment were conducted in five districts of the Dhaka division[5]; Netrokona, Mymensingh, Manikganj, Gazipur and Narayanganj. The trust game was conducted at the end of a rather extensive household survey. The participants were paid Taka 100 to complete the whole survey and the trust game. Although one can never rule out that this survey might have affected subsequent behavior in the trust game, we have no such indications. Since the amount of money was substantial, people clearly concentrated very hard to make good decisions. To avoid the risk of self-selection, and a possible over-representation of relatively trusting and trustworthy participants in the experiment (Holm and Danielsson, 2005), we chose a random sample strategy to match the senders and receives in the trust game. Using ex-ante matching of first and second mover, we ran the experiment among household heads in rural Bangladesh, where senders and receivers come from different but nearby villages and this information is common knowledge to all participants. The enumerators were allocated to different parts of the village to start with and were then asked to perform the interviews for the household survey and to run the experiment in every fourth household.[6] If the household

[5] Bangladesh consists of six divisions, and each division is made up of several districts. In total, Bangladesh has 64 districts, 16 of which are located in the Dhaka division. The sample is of course not representative of the total Bangladesh population.

[6] We actually used four different samples based on religion with intention to test for effects of social distance. By using a split sample technique, we match Muslim senders with either a Muslim or a Hindu receiver and in the same way with the Hindu senders. In each case, senders and receivers were informed about the religious identity of their matched partner. However, as explored in Johansson-Stenman et al. (2005), we found no significant differences between the amount sent and sent back in the different sub-samples. Therefore, we do not analyse these effects in this paper. Still, this implies that Hindus are over-represented in our sample.

head was not around, the enumerators were instructed to go back later and then, if still unsuccessful, to use a replacement household.[7]

At the end of the household survey, the senders and receivers were informed of the trust game. We start by describing the procedure in the sender's household. The respondents in these households were explicitly requested for a private environment, free from interruption by others, during the experiment. The enumerator then read the instructions to the sender, who was assured perfect confidentiality of the responses. The enumerator then presented the outcomes of different decisions to the sender, both related to the amount sent by the sender, and the amount returned to him by the receiver. The senders were then given two envelopes. One of the envelopes contained the original endowment and the other one was empty. The senders used the empty envelope to put the amount of money that they wanted to send to the receiver, and they were assured that the enumerator would not know what their decision was. In the experiment, we used thick envelopes in order to prevent enumerators guessing by eye how much had been sent for the receiver, and thus implicitly the amount kept. The senders were informed that they would be paid within three days.[8] The enumerator ensured that the decision was private by turning his back to the sender while the money was being put into the envelope. The enumerator waited until the sender was ready, after which the sender was asked to close the envelope that was to be sent to the receiver, and seal it

[7] We have approximately 23 percent replacement households in our final sample, which was almost solely because of lack of availability. Only 2% of those being available did not want to participate in the survey. In the villages, people from the same family chain normally live in a cluster of 4-5 households. Thus a replacement from the next household, or the next to next household should not bias our results for this reason.

[8] There is always a potential risk of participants distrusting the people running the experiment. In order to minimise this, university students were used as enumerators. Furthermore, it was specifically mentioned that this was a joint research project being run by a local and a Swedish university.

with a stamp provided before returning it to the enumerator. The sender was instructed to do so even if he/she had decided to send nothing.

After returning the envelopes, the senders were asked about their expectations regarding how much they believed they would eventually get back. The senders were given a piece of paper with three boxes numbered one to three, and an empty envelope into which they were to put the paper back after privately choosing the appropriate box indicating their expectation regarding the back transfer. The alternatives read to them, which were repeated if needed, were: (i) tick the first box if you expect back less than what you sent, (ii) tick the second box if you expect about the same back as the amount you sent and (iii) tick the third box if you expect more back than the amount you sent. In the event that they did not send any money, they were asked to tick the second box. Then the sender was asked to close the envelope, and seal it with a stamp provided, before returning it to the enumerator.

The enumerator then moved on to a discussion about the motives behind the amount chosen to send to the receiver. The senders were given a piece of paper with four boxes numbered from one to four, and an empty envelope into which they were to put their answers to the motive questions. It could of course be argued that more groups should have been used when eliciting the motives. However, the number of alternatives had to be balanced against our desire to maintain complete anonymity among the senders and enumerators, while working with a population where a large fraction is either illiterate or has a very low level of formal education. Therefore, we decided to use four broad groups only. The following alternatives were read to them, and repeated if needed, (i) it would be unfair not to send anything (fairness motive), (ii) the receiver probably needs the money more than you do (need motive), (iii) you believe that you

would get punished either during your lifetime or in an afterlife if you are not generous to others (punishment motive) and (iv) you believe that you would gain from sending the money (trust motive), where the fourth alternative is consistent with the conventional trust interpretation of the game. They were told that they could fill in more than one box if they wanted; the exact wordings are presented in Appendix 1. The punishment motive relates to effects beyond the trust game per se. In both Muslim and Hindu religions, people believe that they will be judged after death. In Muslim religion, on the day of judgement, every human being will be assessed and a decision will be made about whether he or she will go to heaven or hell. In Hindu religion, beliefs in rebirth and reincarnation of souls are essential, and the decision about what will happen in the next life largely depends on the actions in this life.

Finally, they were explicitly asked not to discuss the experiment with anyone else, because people who did not have this opportunity might envy them, which might cause future problems both for them as well as for the organisers. At the end of the day, the enumerator handed the envelopes to the principal researcher, who opened them and put the tripled amount into new envelopes with a household code.

The following day, the enumerators were given these new envelopes ready to be delivered to the assigned receiver. At the end of the household survey, the receivers were informed of the trust game. After ensuring a private environment, they were given the instructions for the experiment, including the same examples that were given to the sender. Then they received an envelope containing the tripled amount from the sender as well as an empty envelope. The enumerator then turned his back to the receiver who had been instructed to use the empty envelope to put the amount of money that he/she wanted to return to the sender. After putting the chosen amount in the envelope, he/she

was to close it and seal it with a stamp provided by the enumerator before returning the envelope.

Then the receiver was asked about his/her motives to send back money. The senders were given a piece of paper with three boxes numbered one to three, and an empty envelope into which they were to put the paper back after privately choosing the appropriate box(es) regarding their motives. He/she was then given an empty envelope into which he/she was to put the privately made answers to the motive questions and seal it. The alternatives presented to them were the same as for the senders, except for the alternative stating that they would gain financially from sending back any money (which is impossible). Thus, they were presented with the following alternatives: (i) it would be unfair not to send anything (fairness motive), (ii) the sender probably needs the money more than you do (need motive), (iii) you believe that you would get punished either during your lifetime or in an afterlife if you are not generous to others (punishment motive). At the end of the day the enumerators returned the envelopes to the principal researcher who checked and wrote down the amount to be transferred back. During the following day, i.e. on the third day, the envelopes were delivered back to the senders with the amount that the receiver had decided to return.

3. Results

This section contains our results from the part of the survey on the attitudinal trust as well as on the other social capital variables, in addition to the trust game. The average amount sent is 92.5 Bangladesh Taka (TK)[9], corresponding to 46 percent of their initial endowment of 200 TK. As compared to the comprehensive overview of trust games

[9] 57.88 TK=1 USD, at the time of the trust game (October, 2003).

conducted in developing countries and countries undergoing transition in Carpenter and Cardenas (2004), the fraction sent is a little bit lower than the average, but on the other hand many of these other studies were conducted with more homogenous respondents, typically students. Out of 256 senders who participated in the experiment 18 senders (7%) sent nothing while 46 senders (18%) sent everything. Out of the 237 receivers[10], 11 receivers (5%) sent back nothing to the senders while 9 receivers (4%) sent back everything to the senders. The average amount returned was 134.6 TK. Thus, it was on average profitable for the senders to send money in this trust game. The results in the literature are otherwise mixed on this point; Cardenas and Carpenter (2004) found that senders have benefited from trusting in 17 out of the 25 trust games reported.

We asked the senders about how much they expected to get back from the receivers. Since this is a rather difficult question to answer concisely, and since most respondents probably had some kind of subjective probability distribution regarding the receivers' behavior rather than a point estimate, we only asked them to choose between three alternatives. We asked them whether they expected less, the same or more than the amount that they had sent.[11] As can be seen from Table 1, slightly less than one third of the senders believed they would gain from sending money, and slightly less than one third believed they would loose, but still sent their money.

Hence, already these results provide some indication that the first stage of a trust game may not measure pure trust alone. Furthermore, as can be seen from the table, the pattern regarding the proportion of endowment sent for the three different expectations is reversed from what one would expect if pure trust was the only influence, since the

[10] One of the pre-assigned receivers refused to take part.

[11] One respondent chose all three responses; therefore we drop this observation when analysing senders' expectations.

senders who on average expect less sent more and vice versa. This is in sharp contrast to Bellemare and Kröger (2003), Chaudhuri and Gangadharan (2003) and Fehr et al. (2002), who all found a positive correlation between the amount sent and the expectations. One possible reason for this discrepancy is of course that, for whatever reason, the result here may be flawed. However, another possible reason is due to the fact that we did not ask the respondents to state a number when expressing their expectations and thus the amount they just sent will not obviously work as an "anchor." The other studies reported, on the other hand, asked the respondents to state the amount of money they expected the receiver would send back, after they had decided how much to send. It is well known from psychological research that respondents can be extremely sensitive to provided anchors, even in situations where it is obvious that the anchors carry no relevant information at all, such as the random number coming up from a spinning wheel (Tversky and Kahneman, 1974). In the case of a trust game, the most straightforward anchor is the amount they have just sent. Thus, the positive correlation found in several studies may, at least partly, be due to anchoring.[12]

[12] See Selten and Ockenfels (1998) for other reasons why the causality may go from the sent amount to expectations, instead of the other way around.

Table1. Average proportion sent in the trust game for different sub-samples of senders

	Share of respondents	Mean proportion sent
Whole sample (n=256)	100%	46%
Senders' expectations on sending a non-zero amount to the receiver (n=237) [a]		
Expects to get back less than the amount sent	29%	55%
Expects to get back about the same as amount sent	42%	51%
Expects to get back more than the amount sent	29%	44%
Motivations for sending non-zero amounts to the receiver (n=238)		
It would be unfair not to send anything (fairness motive)	21%	49%
The receiver probably needs this money more than you do (need motive)	26%	56%
You will get punished, either during your lifetime or afterwards, if you are not generous to others (punishment motive)	50%	51%
You believe you will gain from sending the money (trust motive)	23%	41%
Stated trust as the level of agreement with the statement "most people can be trusted" (n=256)		
Strongly disagree	15%	43%
Disagree	22%	40%
Partly disagree	31%	46%
Partly agree	14%	53%
Agree	7%	57%
Strongly agree	11%	49%

[a] Eighteen senders did not send anything. The expectation of one sender, with positive transfer, is not analysed for choosing all three expectations.

Table 2. Average proportion returned in the trust game for different sub-samples of receivers.

	Share of respondents	Mean proportion returned
Whole sample (n=237)	100%	46%
Motivations for returning non-zero amounts to the sender (n=221) [a]		
It would be unfair not to send anything back (fairness motive)	30%	44%
The sender probably needs this money more than you do (need motive)	32%	43%
You will get punished, either during your lifetime or afterwards, if you are not generous to others (punishment motive)	54%	50%
Stated trust as the level of agreement with the statement "most people can be trusted" (n=237)		
Strongly disagree	11%	31%
Disagree	30%	43%
Partly disagree	30%	49%
Partly agree	14%	47%
Agree	8%	52%
Strongly agree	7%	60%

[a] Eleven receivers sent back nothing. Among those who sent back positive amounts, five receivers did not express any motive behind the transfer.

To test more formally whether the proportions of money sent by the senders come from populations with similar distributions when separated according to their expectations, we conduct a series of pair-wise group comparisons using the Wilcoxon-Mann-Whitney test.[13] At 5% significance level, we can reject the null hypotheses that the proportions sent in the pair-wise comparisons come from the same distributions, except for the comparison of the proportions sent when the senders expect less and about the same, respectively. Moreover, we also reject the null hypothesis that all three samples come from populations with the same distribution at 5% significance level using the Kruskal-Wallis test (see Appendix 2 for detailed results of the tests).

The senders were then free to select one or several of the four available alternatives explaining their motives behind their behavior, while the receivers had three alternatives (since the alternative that they would gain financially is of course not applicable). 75 senders and 62 receivers chose more than one motive to describe their behavior in the trust game.[14] Perhaps surprisingly, as already indicated in Table 1, only 23% of those sending money stated that they thought that they would actually gain from sending the money, i.e. the motive that is typically seen to reflect trust. A strong driving force for both the senders and the receivers seems to be the fear of punishment, either in this life or after this life, and this alternative is chosen by about half of the senders as well as receivers. It would not have been possible to identify this type of motivation if

[13] For a more detailed description of the tests used, see e.g. Siegel and Castellan (2000).

[14] Three senders chose all four alternatives while the other combinations chosen were; fairness, need and punishment (5), fairness, punishment and trust (4), need, punishment and trust (3), trust and need (8) trust and fairness (4), fairness and punishment (17) and fairness and need (11), need and punishment (12), punishment and trust (8). Twelve receivers chose all three motives, the other contributions picked are: fairness and need (13), fairness and punishment (19) and need and punishment (18). It should be noted that five receivers did not indicate any motive i.e. they returned a blank answer sheet.

we had restricted our interpretation to revealed behavior. Hrung (2004) provides evidence, based on the time pattern of religious and non-religious charitable giving, that the behavior of Americans may also be influenced by expected after-life consequences; thus, our results may not be unique to Bangladesh society or even to the Muslim and Hindu world, even if it appears likely that the strength of this motive varies culturally.

Table 1 and Table 2 also present the proportion sent and returned, respectively, for the different sub-samples classified by their motives. Remarkably, we find that the proportion sent is actually lowest for those who reported that they sent the money because they believed that they would gain from sending money, i.e. the pure trust motive. At 5 % significance level, we can reject the null hypothesis that the proportion sent in cases where need or trust are the motivations arise from similar distributions, and similarly for the null hypothesis that the proportions sent are equal for punishment and trust motivations (see Appendix 2 for detailed results and histograms). From Table 2, it seems that what might be denoted as punishment considerations influence the amount returned by the receivers, but in pair-wise tests we cannot reject the null hypothesis that the proportion returned arises from the same distribution for these two sub-groups (see Appendix 2 for detailed results and histograms). The results we report here support that suggestion that there are other motives behind actions in a trust game, not just pure trust and reciprocity/trustworthiness, thus supporting e.g. Cox (2004) and Holm and Danielsson (2005). These findings are also extended by showing the importance of motives that are not possible to deduct from observed behavior, such as a fear of subsequent punishment.

Based on the most frequently used standard GSS question: “Generally speaking, would you say that most people can be trusted or that you cannot be too careful in

dealing with people?", only 3 percent answered that most people can be trusted, which seems to indicate a very low level of trust. Since this measure, for obvious reasons, is very crude, we also asked a similar question by allowing answers along a six-point scale ranging from "strongly disagree that most people can be trusted" to "strongly agree that most people can be trusted" as presented in Table 1 and Table 2.[15] Although less extreme responses, this question also reflects rather low levels of general trust.[16] We can also see that there is no strong relation between stated trust and the amount sent, even though a visual inspection indicates a weak positive relation suggesting that behavior observed in a trust game may be influenced by motives other than trust. There is a clearer pattern between stated trust and the amount sent back, which is similar to the reported results in Glaeser et al. (2000).

4. Econometric analysis

Table 3 defines the explanatory variables used in the econometric analysis together with their mean values. In Table 4, we present the estimates from the OLS regressions to explain factors that may influence the fraction sent by the senders, as well as the fraction returned by the receiver.[17]

[15] However, meanings of intermediate points in the scale were not mentioned to the respondents.

[16] We also asked the same question on trust (not reported) in a situation with either low or high stakes. Perhaps not surprisingly, people trust others more at low stake events.

[17] In the trust game, there are four different types of matching between Hindu and Muslim participants. Based on a joint test, we cannot reject the null hypothesis that the same amounts are sent in the four groups (p-value=0.46), nor can we reject the hypothesis that the same fractions are returned in the four groups (p-value=0.56). Therefore we do not include the corresponding dummy variables in the proceeding analysis.

Table 3. Sample statistics

Variable	Definition	Mean	Minimum	Maximum	N
Age	Age of the respondent	44.7	19	87	512
Illiterate	Cannot read and write	0.28	0	1	512
Low education	Not illiterate and /or education up to high school level	0.57	0	1	512
High education	Education above high school level	0.15	0	1	512
Household equivalent income	Annual household income (in 100000 Taka) adjusted with equivalence and economies of scale. Total yearly household income was divided by [(number of adults + 0.5× number of children)$^{0.75}$]	0.23932	0.010	3.64	511
Stated trust	Level of agreement with the statement that most people can be trusted (1= strongly disagree, 2= disagree, 3=partly disagree, 4=partly agree, 5=agree, and 6 = strongly agree).	3.05	1	6	512
Trusting behavior	Frequency of lending money to friends and neighbours: 1=once a year or less, 2= about once every other month, 3= about once a month, 4= about once a week, 5= more than once a week.	1.81	1	5	512
Confidence index	Arithmetic sum of confidence on 10 institutions [Banks, NGOs, Military, Police, Judiciary, Local government, Executive Government, Political parties, Rural power elites, Educational institutions]: great deal of confidence=2, only some confidence=1, and hardly any confidence at all=0.	14.56	2	20	508
Recent misfortune	The respondent has been a victim of any of the following incidents in the last one year: robbery /theft, mugging, personal assault, home attack, land fraud, false criminal accusation, and political harassment.	0.232	0	1	511
Member of voluntary association	Has membership in voluntary groups and/ or association	0.291	0	1	509

Table 4. Regression analysis of proportion sent, proportion returned and stated trust

Dependent variable	Proportion sent Model 1	Proportion sent Model 2	Proportion returned Model 3	Stated trust [a]
Regression type	Least square	Least square	Least square	
Non-zero amount sent by the sender	0.535***	0.531***		
	(0.073)	(0.087)		
Proportion sent by sender			-0.034	
			(0.073)	
Non-zero amount returned by the receiver			0.534***	
			(0.110)	
Sender expects about the same back as sent	-0.043	-0.039		
	(0.042)	(0.043)		
Sender expects a higher amount back than sent	-0.100**	-0.081		
	(0.048)	(0.051)		
Fairness motive		-0.022	-0.038	
		(0.047)	(0.054)	
Need motive		0.056	-0.060	
		(0.048)	(0.053)	
Punishment motive		0.003	-0.019	
		(0.047)	(0.053)	
Trust motive		-0.048		
		(0.056)		
Age	0.019**	0.016*	-0.005	-0.027
	(0.008)	(0.010)	(0.010)	(0.231)
Age squared	-0.0002**	-0.0002*	0.0001	0.0003
	(0.0001)	(0.0001)	(0.0001)	(0.0002)
Illiterate	-0.037	-0.052	-0.154**	0.380**
	(0.063)	(0.063)	(0.068)	(0.170)
Low education	-0.048	-0.064	-0.010*	-0.047
	(0.056)	(0.057)	(0.061)	(0.147)
Household equivalent income	0.158**	0.143**	0.010	0.055
	(0.063)	(0.064)	(0.058)	(0.152)
Stated trust	0.021*	0.020	0.057***	
	(0.012)	(0.012)	(0.016)	
Trusting behavior	0.022	0.017	-0.006	0.079*
	(0.016)	(0.017)	(0.020)	(0.045)
Confidence index	-0.001	-0.001	0.004	0.039***
	(0.005)	(0.010)	(0.006)	(0.015)
Member of voluntary association	0.038	0.036	-0.025	0.001
	(0.038)	(0.038)	(0.046)	(0.107)
Recent misfortune	-0.047	-0.044	0.061	-0.310***
	(0.039)	(0.040)	(0.0510)	(0.114)
Constant	-0.505**	-0.429*	-0.101	
	(0.242)	(0.251)	(0.273)	
R^2 (Pseudo-R^2 in Model 4)	0.270	0.271	0.204	0.051
No. of observation	250	249	227	503

Standard errors are in parentheses. Superscripts *, **, *** denote statistical significance at 10%, 5%, and 1% level.

[a] We control for enumerator effects, but the parameter estimates for enumerators are omitted from the presentation.

Since we used several enumerators to run the trust experiment, we tested if there was an enumerator effect. However, we cannot reject the hypothesis of no enumerator effect in the trust game regressions in Table 4 (p-value 0.82, 0.86, 0.41, respectively, for Model 1, 2 and 3), based on joint F-tests. However, for Model 4, the stated trust regressions, we can reject the hypothesis of no enumerator effect at 1% significance level, and hence we include dummy variables for the enumerators to control for enumerator effect.

In Model 1 in Table 4, we use dummy variables for senders' expectations[18] finding a significant *negative* relation, implying that a significantly lower amount (or fraction) is sent when the sender expects more in return, which is in line with our descriptive results reported in Section 3. The age profile of the amount sent is consistent with Bellemare and Kröger (2003) and Gleaser et al. (2002), as well as the age pattern of stated trust reported in Putnam (2000). Kocher and Sutter (2003), on the other hand, report an age pattern where the amount sent increases from childhood to early adulthood, but stays almost constant afterwards, whereas Fehr et al. (2002) found that older people send significantly less.

We also find that the amount sent increases with income, and weakly with stated trust. Glaeser et al. (2000) found no significant effect of stated trust on the amount sent, whereas Holm and Nystedt (2005) found that the correlations between stated trust and the amount sent in the trust game increased significantly when monetary incentives were taken away from the trust game, i.e. the correlation is stronger with a hypothetical trust game.

[18] We also include a dummy variable for sending a positive amount, since we want to measure the influence of motives conditional on having sent a positive amount.

In Model 2, we also include motivational dummy variables, but none of the associated parameters are significant.[19] Stated trust significantly explains the fraction returned (Model 3 in Table 4), which is similar to Glaeser et al. (2000) who also found that stated trust better explains the fraction returned than the fraction sent. Finally, we attempt to explain stated trust measured on a 6-point scale using the same control variables. Contrary to the findings in the trust game, we find that stated trust is significantly and positively affected by confidence in institutions and significantly and negatively affected by the occurrence of a recent misfortune. Also, past trusting behavior (frequency of lending money) weakly predicts stated trust in the survey. The positive relationship of stated trust and trusting behavior is consistent with the findings in Bellemare and Kröger (2003) and Fehr et al. (2002). We also find that illiterate people have a higher level of stated trust for which we do not have any adequate explanation.

5. Discussion and conclusion

Although an extremely high level of corruption has been observed in Bangladesh, our results from the trust game, having fractions being sent and returned that are comparable to most other studies, do not support the hypothesis that corruption has been transferred to the individual level and developed into a general low level of trust in others. However, our survey responses do indeed reflect very low levels of stated trust. The obvious question, then, is: Which measure should we trust? The answer is far from

[19] One would expect that the sender's expectation and motive might be correlated. We estimated a separate model excluding expectation dummies. The results are roughly the same. Hence, we present the results controlling for both expectations and motives in Model 2. Moreover, we estimated a model by interacting respondents' religion with the stated motives, to see whether certain motives are linked to religion or not. We do not find any significant effect (p-value 0.56), based on F-test, and hence we do not include these interaction terms in the model presented.

straightforward. As economists, we might have a bias towards relying on observed behavior with monetary incentives. For example, the fact that Glaeser et al. (2000) discovered a poor correlation between stated trust and the amount sent in the trust game was seen as an indication that stated trust does not measure real trust, not the other way around. As somewhat provocatively expressed by McCloskey (1985, 181): "Economists are so impressed by the confusions that might possibly arise from questionnaires that they have turned away from them entirely, and prefer the confusions resulting from external observation." In our view, pros and cons of each method should be discussed without prejudices.

Evidence put forward in favour of stated-trust questions is the strong measured correlation between the fraction agreeing that most people can be trusted and the number of wallets that were returned in a lost-wallet experiment in these cities, reported in Knack and Keefer (1997). They also found a rather strong correlation between measured stated trust and economic growth. Hence, one may argue that whatever stated trust measures, it appears to be something important, and something that is overall good for society. On the other hand, the fractions of wallets coming back is not really a measure of trust, but rather a measure of a particular social norm, which seems, if anything, to be more closely related with trustworthiness than with trust.[20] This is consistent with the finding in Glaeser (2000), and this paper too, that stated trust is a better predictor of the amount sent back by the receivers, than of the amount sent by the senders in the trust game.

Trust, as measured by the fraction sent in a trust game, has the clear advantage of relying on real monetary incentives, implying that it is costly for the participants to

[20] This may possibly also be a reflection of the trust in the police authorities. If people believe that a handed in wallet will just benefit the local policemen then they may find it meaningless to hand it in.

deviate from their true preferences and perceptions. This implies, for example that trust as measured by the amount sent in a trust game is less likely to be vulnerable to self-serving bias, in terms of self-signalling (cf. e.g. Benabou and Tirole 2003, 2004) and possible self-presentation effects, compared to survey questions. As discussed, there is also empirical evidence, e.g. when comparing the behavior in trust games with the behavior in dictator games, that the sent amount, to some extent, does measure trust. What is less clear is the extent to which different motives affect the behavior. Some seem to conclude, implicitly or explicitly, that even though other motives may also matter, they are probably relatively small, implying that they might be ignored. On the other hand, the results here, as well as the recent evidence by e.g. Cox (2004), indicate that this is too strong a conclusion. In this study, we show, for example, that the fear of subsequent punishment if behaving too selfishly seems to be a very important motive for both senders and receivers. Even though survey questions have their inherent problems, we cannot see how we could have obtained this kind of result if we were restricted to observing people's actual behavior in the game. At the end of the day, our conclusion is somewhat negative, suggesting that neither stated trust nor the fraction sent in trust games may be particularly good measures of trust, as we normally think about the word. This is, of course, not a reason to stop analysing and trying to measure trust, just as the obvious fact that, just because welfare is difficult to measure, we should not be prevented from doing welfare analysis. We believe, however, that our findings provide further arguments for trying to find other, better, measures of the important phenomenon of trust.

References

Alesina, A., E. La Ferrara, 2002, Who Trusts Others? Journal of Public Economics 85, 207-234.

Ashraf, N., I. Bohnet, N. Piankov, 2002, Decomposing Trust, Harvard University, mimeo. October.

Babcock, L. and Loewenstein, G., 1997, Explaining Bargaining Impasse: The Role of Self Serving Biases, Journal of Economic Perspectives 11, 109-126

Bellemare, C. and Kröger, S., 2003, On representative trust, University of Tilburg, Working Paper.

Benabou, R. and Tirole, J., 2002, Self-confidence and personal motivation, Quarterly Journal of Economics 117, 871-915.

Benabou, R. and Tirole, J., 2004, Willpower and personal rules, Journal of Political Economy 112, 848-886.

Berg, J., Dickhaut, J., McCabe, K., 1995, Trust, Reciprocity, and Social History, Games and Economic Behavior 10, 122-42.

Beugelsdijk, S., Groot, H.L.F. de, Schaik, A.B.T.M. van, 2004 Trust and economic growth: A robustness analysis, Oxford Economic Papers 56, 118-134.

Bohnet, I., Harmgart, H., Huck, S. and Tyran, J.-R., 2005 Learning trust, Journal of the European Economic Association, forthcoming.

Bolton, G. E. and Ockenfels, A., 2000, ERC: A theory of equity and reciprocity, and competition, American Economic Review 90, 166-193.

Bouckaert, J. and Dhaene, G., 2004, Inter-Ethnic Trust and Reciprocity: Results of an Experiment with Small Businessmen, European Journal of Political Economy 20, 869-886.

Cameron, L. A., 1999, Raising the stakes in the ultimatum game: Experimental evidence from Indonesia, Economic Inquiry 37, 47-59.

Cardenas, J.C. and Carpenter, J., 2004, Experimental development economics: A review of the literature and ideas for future research, mimeo. Middlebury College, Middlebury, Vermont.

Carter, M., and Castillo, M., 2002, The economic impacts of altruism, trust and reciprocity: An experimental approach to social capital, Working paper, Wisconsin-Madison Agricultural and Applied Economics Department, USA.

Cox, J. ,2004, How to identify trust and reciprocity, Games and Economic Behavior 46, 260-281.

Chaudhuri, A. and Gangadharan, L., 2003, Sending Money in the Trust Game: Trust or Other Regarding Preferences? University of Auckland Working Paper, December.

Dufwenberg, M. and Gneezy, U., 2000, Measuring Belief in an Experimental Lost Wallet Game, Games and Economic Behavior 30, 163-182.

Fehr, E., and Schmidt, K.,1999, A Theory of Fairness, Competition, and Cooperation, Quarterly Journal of Economics 114, 817-868.

Fehr, E., Fischbacher, U., Rosenbladt, B., Schupp, J., and Wagner, G.G., 2002, A Nation-Wide Laboratory: Examining Trust and Trustworthiness by Integrating Behavioral Experiments into Representative Surveys, Schmollers Jahrbuch 122, 519-542.

Fukuyama, F., 1995, Trust, New York: Free Press.

Fershtman, C. and Gneezy, U., 2001, Discrimination in a segmented society, Quarterly Journal of Economics 116, 351-77.

Glaeser, E. L., Laibson, D. I., Scheinkman, J.A., Soutter, C.L., 2000, Measuring Trust, Quarterly Journal of Economics 115, 811-846.

Glaeser, E. L., D.I. Laibson, Sacerdote, B., 2002, An Economic Approach to Social Capital, Economic Journal 112, 437-458.

Holm, H. J. and Danielsson, A., 2005, Tropic trust versus Nordic trust: Experimental evidence from Tanzania and Sweden, Economic Journal, forthcoming.

Holm, H.J. and Nystedt, P., 2005, Intra-generational trust—a semi-experimental study of trust among different generations, Lund University, mimeo.

Hrung, W. B., 2004, After-Life Consumption and Charitable Giving, American Journal of Economics and Sociology 63, 731-45.

Johansson-Stenman, O., M. Mahmud, P. Martinsson, 2005, Trust and Religion: Experimental Evidence from Bangladesh, Department of Economics, Göteborg University, Mimeo.

Knack, S., Keefer, P., 1997, Does Social Capital Have an Economic Payoff? A Cross-Country Investigation, Quarterly Journal of Economics 112, 1251-1288.

McCloskey, D., 1985, The Rhetoric of Economics. Madison: The University of Wisconsin Press.

Putnam, R. D., 1993, Making Democracy Work- Civic Tradition in Italy, Princeton University Press.

Putnam, R. D., 2000, Bowling Alone: The Collapse and Revival of American Community, New York.

Rothstein, B. and Stolle, D., 2001, Social Capital and Street –Level Bureaucracy: An Institutional Theory of Generalized Trust, Prepared for the ESF conference Social Capital: Interdisciplinary Perspectives, in Exeter, United Kingdom, September 15-20.

Selten, R. and. Ockenfels, A., 1998, An Experimental Solidarity game, Journal of Economic Behavior and Organization 34, 517-539

Sen, A. K., 1973, Behavior and the Concept of Preference, Economica 40, 241-59.

Sutter, M. and Kocher, M.G., 2003, Age and the development of trust and reciprocity, University of Innsbruck, mimeo, December.

Siegel, S. and Castellan, N.J., 2000, Nonparametric Statistics for the Behavioral Sciences, Second edition, McGraw-Hill.

Tversky, A. and D. Kahneman, 1974, Judgment under uncertainty, Science 185, 1122-1131.

Appendix 1. Eliciting subjects' motives in the trust game

Senders' motive

There are four boxes on this piece of paper. By putting a tick mark in any of the boxes you can indicate why you sent money. You may put tick marks in more than one box.

- Tick the first box if you think it would be unfair not to send anything.
- Tick the second box if you think the receiver probably needs this money more than you do.
- Tick the third box if you believe that you will get punished, either during your lifetime or afterwards, if you are not generous to others.
- Tick the fourth box if you believe that you will gain from sending the money

After you tick a box, please put the paper in the envelope, seal it with the stamp and return it to me.

Receivers' motive

There are three boxes on this piece of paper. By putting a tick mark in any of the boxes you can indicate why you sent back money. You may put tick marks in more than one box. You can put tick marks in any of the boxes even if you did not send back any money.

- Tick the first box if you think it would be unfair not to send anything back.
- Tick the second box if you think the sender probably needs this money more than you do.
- Tick the third box if you believe that you will get punished, either during your lifetime or afterwards, if you are not generous to others.

After you tick a box, please put the paper in the envelope, seal it with the stamp and return it to me.

Appendix 2. Detailed test statistics of proportions sent and returned

A.2.1 Wilcoxon and Mann-Whitney rank test and Kruskal- Wallis test of difference of proportion sent when motivations are different

Hypothesis	P-value
Proportion sent when sender expects less than amount sent = Proportion sent when sender expects about the same amount sent	0.89
Proportion sent when sender expects less than the amount sent = Proportion sent when sender expects more than the amount sent	0.025
Proportion sent when sender expects about the same amount sent= Proportion sent when sender expects more than the amount sent	0.002
Proportion sent when sender expects less than the amount sent= proportion sent when sender expects about the same amount sent= proportion sent when sender expects more than the amount sent	0.009

A.2.2 Wilcoxon and Mann-Whitney rank test and Kruskal- Wallis test of difference of proportion sent when motivations are different

Hypothesis	P- value
Proportion sent when fairness motive is involved = Proportion sent when need motive is involved	0.649
Proportion sent when fairness motive is involved = Proportion sent when punishment motive is involved	0.752
Proportion sent when fairness motive is involved = Proportion sent when trust motive is involved	0.133
Proportion sent when need motive is involved = Proportion sent when punishment motive is involved	0.736
Proportion sent when need motive is involved = Proportion sent when trust motive is involved	0.043
Proportion sent when punishment motive is involved = Proportion sent when trust motive is involved	0.019
Proportion sent when fairness motive is involved = Proportion sent when need motive is involved = Proportion sent when punishment motive is involved = Proportion sent trust motive is involved	0.085

A.2.3. Wilcoxon and Mann-Whitney rank test and Kruskal- Wallis test of difference of proportion sent back when motivations are different

Hypothesis	p- value
Proportion returned when fairness motive is involved = Proportion returned when need motive is involved	0.995
Proportion returned when fairness motive is involved = Proportion returned when punishment motive is involved	0.373
Proportion returned when need motive is involved = Proportion returned when punishment motive is involved	0.402
Proportion returned when fairness motive is involved = Proportion returned when need motive is involved = Proportion returned when punishment motive is involved	0.564

Appendix 3. Histograms of proportions sent and returned

Figure 1. Proportion of money sent by the senders in the trust game

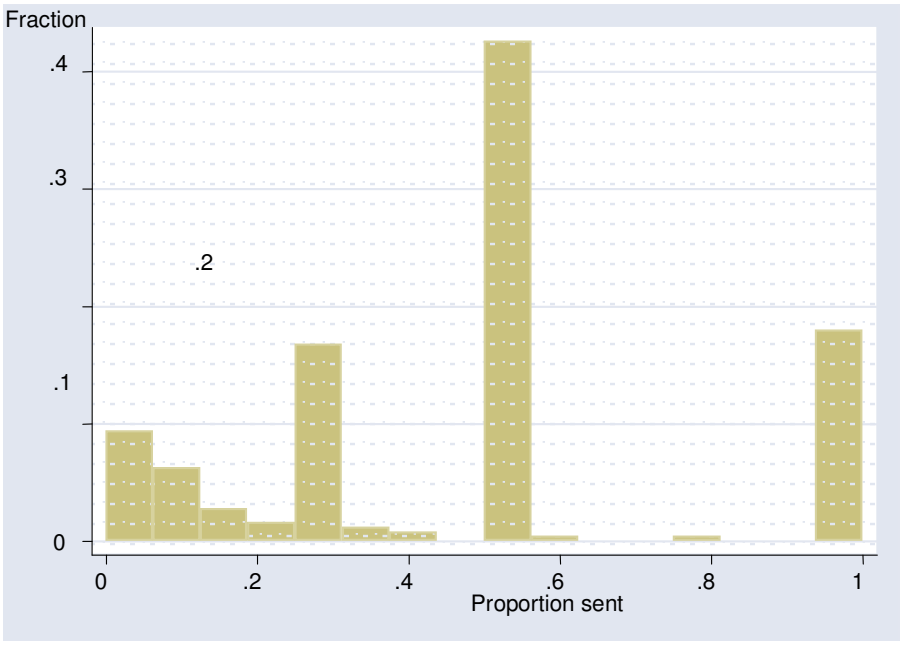


Figure 2. Proportion of money returned by the receivers in the trust game

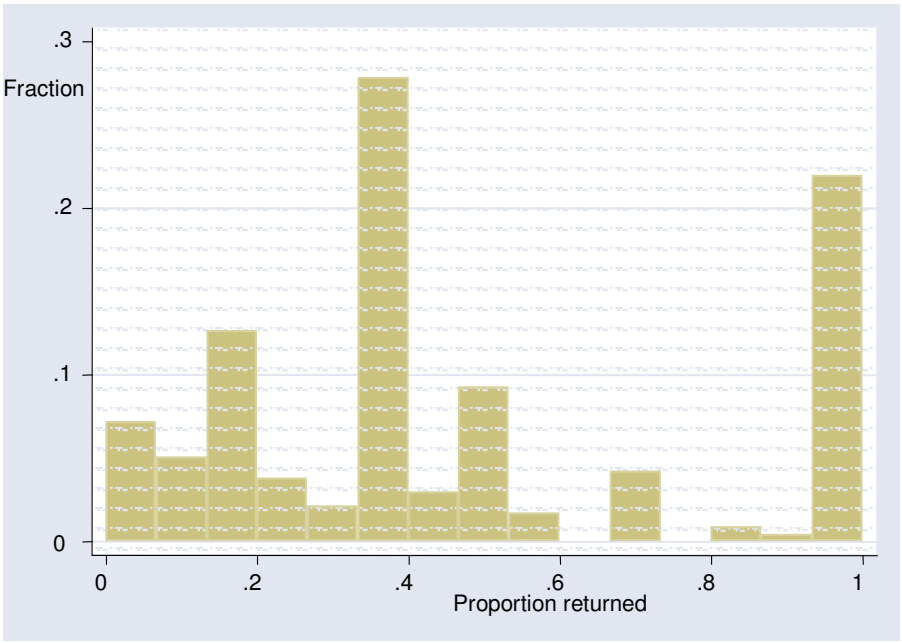


Figure 3. Proportion of money sent by the senders with different stated expectations

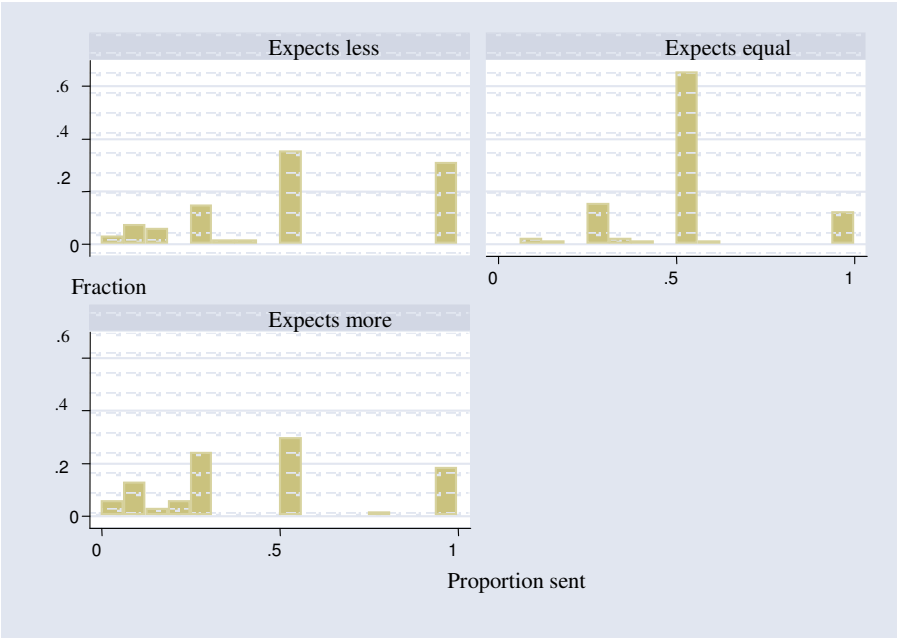


Figure 4. Proportion sent by senders with different stated motives

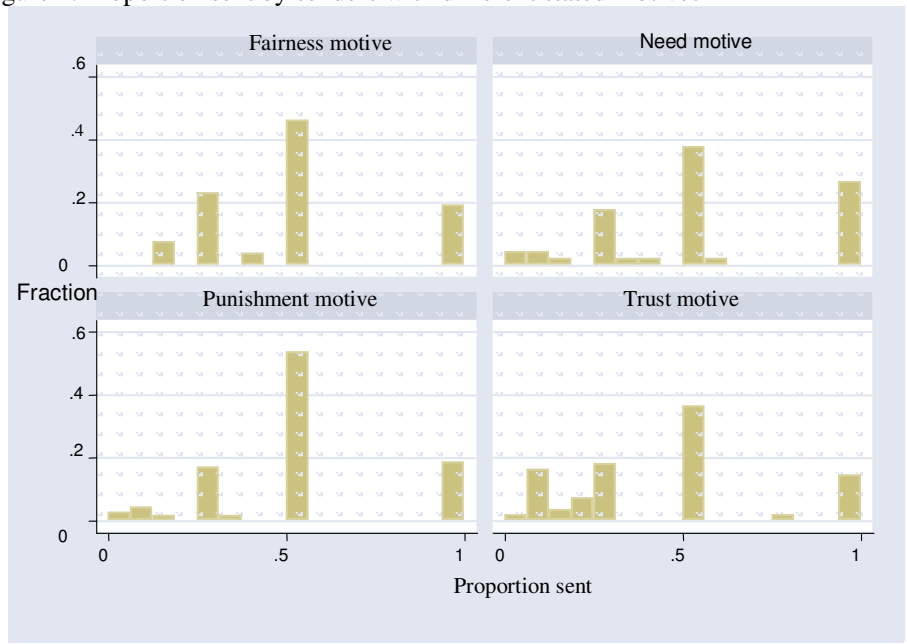
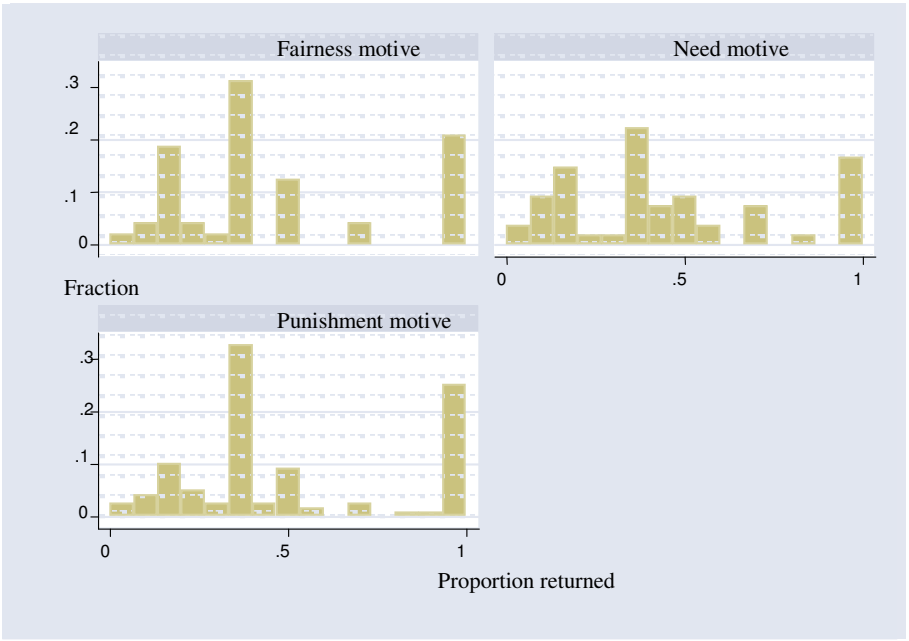


Figure 5. Proportion returned by the receivers with different stated motives



Contingent Valuation of Mortality Risk Reduction in Developing Countries: A Mission Impossible?

Minhaj Mahmud*
Department of Economics, Göteborg University
Vasagatan 1, Box 640, SE- 405 30 Göteborg, Sweden

Abstract

Using the contingent valuation method in developing countries to value mortality risk reduction is particularly challenging because of the low level of education of the respondents. In this paper, we examine the effect of training the respondents regarding probabilities and risk reductions, in addition to using visual aids to communicate risk and risk reductions, in a contingent valuation survey. Our results indicate a significantly higher WTP for the trained sub-sample, and WTP is sensitive to the magnitude of risk reduction both with and without the training.

Keywords: contingent valuation; risk reduction; WTP; effect of training; sensitivity to scope; Bangladesh.

JEL Classification: I1, D6, D8, H4

* I am most grateful for very useful suggestions and comments from Olof Johansson-Stenman, Peter Martinsson, Fredrik Carlsson, Håkan Holm, and seminar participants at Göteborg University. Financial assistance from the Swedish International Development Agency (SIDA) is also gratefully acknowledged.

1. Introduction

The contingent valuation method (CV) has been widely used to value mortality risk reduction, but mostly in developed countries (e.g. Corso et al., 2001; Persson et al., 2001; Krupnick et al., 2002). The CV method involves eliciting people's willingness-to-pay (WTP) for a hypothetical reduction in the risk of dying during a given time period (see Hammitt and Graham, 1999). The individual's rate of trade-off between own money and a small risk change[1] i.e. the marginal rate of substitution, is defined as the value of a statistical life (VSL) (Weinstein et al., 1980; Viscusi, 1993; Johansson, 2002). The individuals' WTPs for a reduction in the mortality risk are converted to VSL by dividing the WTPs by the risk change in question. However, most previous CV studies have found unreasonably low sensitivity of WTP to the size of the risk reduction. One likely reason for the lack of sensitivity is a poor understanding of probabilities and a lack of intuition regarding small risk changes (see Hammitt and Graham, 1999). However, recent evidence by Corso et al. (2001) suggests that there are ways to increase the sensitivity by using visual aids in the presentation of risks in the CV survey.

It is particularly challenging to use the CV method to value mortality risk reductions in developing countries.[2] The main reason is the difficulty to communicate probabilities and risk reduction to the respondents, since many either have very low levels of education or are illiterate. Moreover, most people are unfamiliar with the concept of trading income for the risk reduction and therefore might face greater uncertainty in placing a value on the risk reduction. A brief training of the respondents in the survey regarding probability, risk, and risk changes may enable the respondents to

[1] For example, a change in risk of dying from 5 in 10,000 to 4 in 10,000.

[2] However, by a careful survey design, it is generally possible to conduct high quality CV surveys in developing countries (see Whittington, 1998; 2002).

process risk information better and thus the respondents will become more elaborate about their preferences for risk reduction. Therefore, one might expect that by reducing any uncertainty regarding the object of valuation, the training would yield lower variation in the responses, as well as an increase in the sensitivity to scope. This paper reports on a CV study of mortality risk reduction conducted among a random sample of rural households in Bangladesh. The objective of the study was to examine the effect of training regarding probability, risk, and the implication of risk reductions on the WTP responses and not to obtain an absolute magnitude of the VSL, as well as to investigate whether training affects the sensitivity to scope.

The validity and reliability of the CV method is intensely debated (see e.g. Kahneman and Knetsch, 1992; Diamond and Hausman, 1994; Hanemann, 1994). However, some of the criticisms attached to the CV method, such as "warm glow," or the "purchase of moral satisfaction" phenomenon for contributing to overall social causes, is not applicable when valuing individuals' own risk reductions through vaccinations.[3] On the other hand, it can be highly cognitively demanding for the respondents to compare expected welfare effects from the risk reductions to the effects of monetary changes (Beattie et al., 1998; Hammit and Graham, 1999; Carlsson et al., 2004). The sensitivity of the estimated WTP to the magnitude of the good in question is regarded as a test of the validity of CV estimates (Mitchell and Carson, 1989; NOAA, 1994; Diamond and Hausman, 1994). Assuming that risk reduction is a desired good, the theoretical expectation is that WTP for mortality risk reduction should be positively associated with the magnitude of the risk reduction. Furthermore, for sufficiently small risk changes, WTP should be proportional to the magnitude of risk reduction (Weinstein

[3] Although vaccinations can be seen as a good with a positive externality, we find no indication that people consider this while deciding for their own vaccination.

et al., 1980; Hammitt and Graham, 1999; Hammitt, 2000). This means e.g. that WTP should be twice as high for a two-fold reduction in risk. However, a problem of stating risk in a survey is when the respondents treat the given probabilities as not applicable to them and hence form posterior risk estimates based on their prior beliefs and on information contained in the scenario. In this case, the stated WTP will not be proportional to the magnitude of the stated risk reduction presented; rather, it would rather be proportional to changes in the perceived risk (Viscusi, 1985; 1989).

The use of visual aids has proven to be useful (e.g. Corso et al., 2001; Krupnick et al., 2002) in obtaining responses consistent with the theoretical expectations (sensitive to scope). Examples of visual aids include verbal probability analogies, risk ladders, and graph paper with squares and an array of dots. Corso et al. (2001) found that graph paper and a risk ladder with logarithmic representation of the risk performed the best.

Accordingly, we used graph paper to communicate risk and risk reduction in our survey. We stated either a 25% or a 50% reduction in the risk that corresponds to the respondents' subjective risk of dying during the next five years; subjective risk is based on age-related statistical risks of dying for the same period presented in the CV scenario. Therefore, the insensitivity of scope problem can be expected to be smaller in view of the fact that we include a relatively large risk change together with training on risk reduction, particularly if the insensitivity is related to the poor understanding or lack of intuition about small risk changes on the part of the respondents. Another reason why the theoretical proportionality prediction would not hold in our case is that we deal with substantial risk changes; WTP would increase but less than proportionally to the risk changes and hence the resulting VSL is expected to be smaller compared to the case

when small risk changes are valued. The results indicate a significant difference in the distribution of the WTP between the sub-sample receiving training and the sub-sample receiving no training. However, we find no significant difference in the variance, but rather that WTP increases with training. We also find that estimated WTP is sensitive to the scope of the risk reduction, both with and without training.

The remainder of the paper is organized as follows: Section 2 presents the design of the CV survey including sample characteristics, respondents' risk perceptions etc., Section 3 discusses WTP results, Section 4 presents the econometric analysis, and the paper is concluded in Section 5.

2. Design of the CV survey

Two versions of the CV survey were constructed: one version including a brief training vis-à-vis probability and risk, and the other without such training.[4] The enumerators, used to conduct the survey, were trained beforehand using the guidelines of Whittington (2002). In particular, the enumerators were trained regarding the risk presentations, and the CV methodology in brief, i.e. the purpose of the CV survey, the notion of maximum WTP, etc. The same enumerators were used in the pilot and in the final survey, and they were closely supervised during the fieldwork. We furthermore test for possible enumerator effects when analyzing WTP responses. The CV questionnaire was tested using focus groups and two pilot studies, which, together with the feedback from the enumerators, enabled us to simplify the risk presentations, the CV scenario and the CV question. The survey-questionnaire and the CV scenario were translated back to English

[4] Ideally, it would be better to train people for a longer period; however, we intend to see whether a brief training as part of the questionnaire makes any significant difference in the responses since such education is the most realistic kind that can be pursued.

from Bengali to ensure the exact meaning of the original English version. The final survey was administered, using a random sample technique, among rural households in the following five districts of Bangladesh: Netrokona, Mymensingh, Manikganj, Gazipur, and Narayanganj. The sample is therefore not representative of the Bangladesh population, which consists of 64 districts. Moreover, the villages were chosen so that the respondents of the Hindu religion are over-represented (compared to the national average of 11 %), in order to facilitate religious comparisons. The enumerators were allocated to different parts of the selected villages and were then asked to perform household surveys and the CV experiments. The interviews were conducted with the household head as the decisions made within households are normally made, or at least approved of, by the household head. If a household head was not around, the enumerators were instructed to return later. If the respondent was not home at the second visit, the enumerator moved to next household.[5] The participation rate of household heads approached for interviews was 99 %. The respondents were paid (100 Taka) as an appreciation of their time and cooperation. Table 1 presents the sample statistics for the full sample. The household survey also included detailed questions on respondent health and risk perceptions, in addition to the socioeconomic questions. The CV survey took on average 15 to 30 minutes to complete. 774 individuals were interviewed. Table 1 presents the sample statistics.[6]

[5] However, in the villages people from the same family-chain normally live in a cluster of 4-5 households. Thus, a replacement from the next household or next to the next household (in some cases) should not bias the results. There were 22 % replacement households in our sample.

[6] We have excluded observations related to very old individuals (older than 75), as the risk reduction presented in the CV survey is for a five -year period.

Table 1. Sample Statistics (N=767)

Variable	Mean	Standard deviation	Minimum	Maximum
Male	0.91	0.28	0	1
Muslim religion	0.66	0.47	0	1
Hindu religion	0.34	0.47	0	1
Age	43.6	12.4	19	75
Income per capita [a]	22,594	29,117	807	3,63,650
Illiterate (cannot read and write)	0.31	0.46	0	1
Low education (not illiterate and/or education up to high school level)	0.55	0.50	0	1
High education (education above high school level)	0.14	0.35	0	1
Having chronic illness [b]	0.39	0.49	0	1
Currently smoking	0.56	0.50	0	1
Self reported happiness [c]	5	2.2	0	10

[a] Total yearly household income was divided by [Number of adults + 0.5*number of children] $^{0.75}$, to adjust for household size. N=772. Median 15508 Taka. 57.8 Taka=1 USD at the time of the survey (October 2003).

[b] If the respondent has been suffering from any of the chronic diseases: heart disease, high blood pressure, asthma, bronchitis, cancer, or diabetes.

[c] Responses, on an 11-point scale, to the question: "As a whole, how happy would you say you are? The scale is described as follows: 0 means "extremely unhappy," 10 means "extremely happy'," and 5 indicates average happiness such that half the population in Bangladesh is above 5 and half is below five.

The respondents were asked to state their maximum WTPs for obtaining a stated risk reduction that corresponded to their stated subjective risk of dying during the next five years, which was elicited after the average age-related objective risk had been presented to them. We choose the open-ended format as it provides more information than the closed-ended (dichotomous choice question) format, although many researchers would favor the latter (see e.g. Bateman et al., 2002; Hanley et al., 2003). Moreover, it has been shown in experiments, that dichotomous choice overestimates values more

than the open-ended questions in the case of auction values as well as private goods (see Balistreri et al., 2001).

Based on a *t*- test, we do not find any statistically significant differences (p-value>0.05) in terms of socio-economic characteristics between the populations of the two sub-samples, i.e. training and no training. Therefore, differences in the WTP responses (relating to a specific risk reduction), between these two sub-samples could be attributed to receiving training in the survey.

The responses can be divided into the following categories: (1) training and a 50% risk reduction, (2) training and a 25% risk reduction, (3) no training and a 50% risk reduction, and (4) no training and a 25% risk reduction.

2.1 Training and risk understanding

In the CV questionnaire, the training involved concepts of probability of different events occurring, risks, and implications of risk changes (presented in Figure 1). In particular, we used coin flipping, dice throwing and a lottery example to introduce the concept of probabilities to the respondents. Mortality risk was discussed using the example of risk of dying from traffic accidents. The chance of winning in a lottery and the mortality risk example were explained with the use of graph paper containing 100 and 1,000 squares, respectively. The respondents were asked test questions after each example. If the respondent had a correct answer, the enumerator continued to the next example. To facilitate understanding the respondents received more explanation following a wrong answer, before being asked the same question again. If a respondent still did not have a correct answer after the third, then the enumerator continued to the next example after explaining the correct answer.

Figure 1. Training - Probability and risk examples [a]

Now I will discuss the chances and risks of events occurring using some examples.

Example 1: Sometimes we toss a coin to decide which of two things to choose. When we toss a coin [Enumerator: show tossing a coin], we get either a head or a tail. We cannot be sure of the result of the toss. As there are two things that can happen from a coin toss, the chance of getting a head is 1 in 2. The same is true for getting a tail.

Similarly, when we roll a dice (chakka) [Enumerator: show throwing a chakka] we may either see on the top 1, 2, 3, 4, 5, or 6, but we don't know which one beforehand. Since there are six different numbers from 1 to 6, we may see any of them on the top. The chance of seeing a 5 on the top is 1/6.

Is this example clear to you?

[Enumerator: If no, explain again and make sure that the respondent understands. Write down how many times you had to explain. If the respondent has not understood after three times, write" 4" and continue.]

PT1. Now, if I throw this "chakka" (dice), what is the chance that 2 will be shown on top?

Answer:

[Enumerator: If the answer is wrong, explain with example until the correct answer is given. Write down how many times you had to explain. If the respondent did not have it right after a third explanation, explain the answer and write 4.]

Example 2: Consider buying a lottery ticket. Many people buy lottery tickets and most people do not win. Suppose that there is only 1 prize in a lottery and 100 people buy one lottery ticket each. [Enumerator: Show grid table 1]. In this case we say that the chance of winning the prize will be 1 in 100.

Is this example clear to you?

[Enumerator: If no, explain again and make sure that the respondent understands. Write down how many times you had to explain. If the respondent has not understood after three times, write" 4" and continue.]

PT2. Now, suppose there are two lotteries. The chance of winning in one lottery is 5 in 1000 and the chance of winning in the other lottery is 10 in 1000. [Enumerator: Show the grid table- 2 and grid table 3, when explaining]. Which lottery has the larger chance of winning?

5 in 1000 1

10 in 1000 2

Answer: ……..

[Enumerator: If the answer is wrong, explain with example until the correct answer is given. Write down how many times you had to explain. If the respondent did not have it right after a third explanation, explain the answer and write 4.]

Example 3: Question

PT3. Now, suppose there are two roads that are both very prone to accidents. The risk of dying on road A is 1 in 1000 and the risk of dying on road B is 3 in 1000. [Enumerator: Show the grid table-4 and grid table 5, when explaining]. Which road is more risky to take?

Road A 1

Road B 2

Answer: ……

[Enumerator: If the answer is wrong, explain with example until the correct answer is given. Write down how many times you had to explain. If the respondent did not have it right after a third explanation, explain the answer and write 4.]

[a] The training is read by the enumerators.

The results of the probability test questions are summarized in Table 2. A respondent is considered to have passed the entire test if he/she provided the correct answer to each of the three probability test questions on the first attempt. Only 24% of the respondents passed the entire test.

Table 2. Understanding of probability and risk for the sub- sample with training

Probability/ risk questions [a]	% of respondents answered the test questions correctly	% of respondents answered correctly after 2nd explanation	% of respondents answered correctly after 3rd explanation	% of respondents never answered correctly
Dice throwing (PT1)	31	22	18	29
Lottery winning (PT2)	74	20	5	1
Mortality risk (PT3)	83	15	2	0.25

[a] See Figure 1 for exact wording of the test questions.

In the second stage of the training, the meaning of the risk reductions was explained to the respondents (presented in Figure 2, read by the enumerators). To begin with, the respondents were informed about the average risk of dying for an adult in Bangladesh in the next five years (40 in 1,000).[7] This risk was explained using a graph paper containing 1,000 squares of which 40 were colored black (see Appendix). Then the respondents were told that with appropriate public policy this mortality risk could be reduced to for example 35 in 1,000. The two risk levels were shown simultaneously using graph papers to explain the differences with five of the black squares becoming white in the second graph.

[7] This is based on Bangladesh life table estimates for the year 2000 provided by the World Health Organization (WHO, 2004).

Figure 2. Training - Explaining risk reduction

Example 4: Suppose the average risk of dying for an adult person during the next 5 years is 40 in 1000.

[Enumerator: show grid table 6 when explaining].

Suppose a reduction in mortality risk, through some kind of public measure, could reduce the mortality risk from 40 in 1000 to 35 in 1000

[Enumerator: show grid table 6 and grid table 7 together to explain the difference].

This means that, on average, 5 out of 40 would be saved by the measure.

PT4. Do you understand this risk reduction?

Yes 1

No 2

Answer:

[Enumerator: If no, explain again and make sure that the respondent understands and write down how many times you had to explain. If the respondent has not understood after three times, continue and write 4]

Example 5: Similarly, if the risk was reduced from 40 in 1000 to 20 in 1000[Show grid table 6 and grid table 8 together to explain the difference], then 20 out of 40 would be saved on average.

PT5. Do you understand this risk reduction?

Yes 1

No 2

[Enumerator: If no, explain again and make sure that the respondent understands and write down how many times you had to explain. If the respondent has not understood after three times, continue and write 4]

Answer:

Example 6: If the risk was reduced from 40 in 1000 to 10 in 1000[show grid table 6 and grid table 9 together to explain the difference], then 30 out of 40 would be saved on average.

PT6. Do you understand this risk reduction?

Yes 1

No 2

[Enumerator: If no, explain again and make sure that the respondent understands and write down how many times you had to explain. If the respondent has not understood after three times, continue and write 4]

Answer:

Example7- Question

PT7. Which of the above risk reductions would you prefer?

[Enumerator: Show the cards and let the respondent point]

a) 40 in 1000 to 35 in 1000

b) 40 in 1000 to 20 in 1000

c) 40 in 1000 to 10 in 1000

The implication of the risk changes from 40 in 1,000 to 35 in 1,000 was explained to the respondents by saying that 5 out of 40 lives could be saved through a policy measure. In a similar fashion, the meaning of further risk reduction was explained to the respondents, i.e. reducing the risk from 40 in 1,000 to 20 in 1,000 and

reducing the risk from 40 in 1,000 to 10 in 1,000. Each risk reduction example was explained up to three times to facilitate the respondents' understanding. Almost 95% of the respondents revealed that they had understood all risk reduction examples after the first explanation. This might reflect a "yea saying" bias as the question (Do you understand this risk reduction?) is of a yes/no type. However, when asked in the end of the training to indicate which of the above three risk reduction examples they would prefer, almost 98 % of the respondents preferred the largest risk reduction example (reducing the risk from 40 in 1,000 to 10 in 1,000), which suggests that they had understood the risk reduction examples.

2.2 Objective risk and risk perception

All respondents in the survey, before being presented with the CV scenario, were first informed about the average mortality risk of persons aged 30-34 and persons aged 55-59 as 15 in 1,000 and 90 in 1,000, respectively, in the next five year period (see Figure 3). Then the respondents were asked to mention their perceptions of their own risks of dying during the same period taking into consideration their ages, health and lifestyles in particular.

Figure 3. CV questionnaire: Risk perception

It has been estimated that in Bangladesh, an average of 15 out of 1000 people in the 30-34 age group will die over the next five years from various causes, and 90 out of 1000 people in the 55-59 age group will die over the next five years from various causes Enumerator: [show grid table 10 and 11].

R1. *Thinking about your own life and the way you are living it, what do you think the risk of you dying in the next five years is? [Enumerator: Let the respondent also see the tables 10 and 11 again, at the same time].*

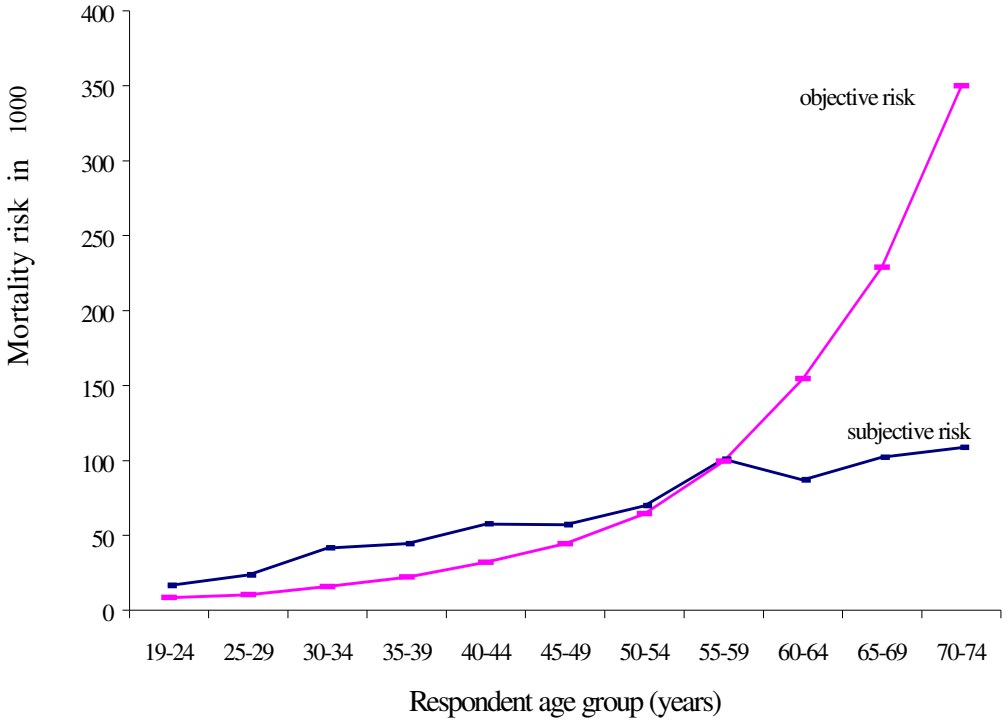
Answer: in 1000

[Enumerator: Use the grid table 12, which is an empty grid table to represent the respondent's subjective risk of dying in the next five years. Let the respondent look at it.]

Thus, we customize the mortality risk for each individual according to his or her own perception. Figure 4 shows the mean mortality risk, both objective and subjective, for various age groups of respondents during the next five years. As observed, people on average overestimate mortality risk at younger ages and underestimate it at older ages. This supports earlier findings (Viscusi, 1992) that people tend to overestimate small risks and underestimate large risks. Based on a non-parametric (Wilcoxon matched-pairs signed-rank) test, we can conclude that respondents' subjective and objective (age-related) risks are significantly different (p- value <0.001).[8] That people have a biased risk perception is also consistent with much research in psychology (Tversky and Kahneman, 1992; Kahneman and Tversky, 2000; Gilovich et al., 2002).

[8] The difference between objective and subjective risk is positive, negative and zero in 446, 290 and 35 cases, respectively.

Figure 4. Objective and subjective mortality risk during the next five years as a function of age



We estimate ordinary regression to see what characteristics explain the risk perception of individuals. For obvious reasons we do not focus on gender differences in risk perception since we only have 9% female respondents.[9] As age and age-related objective risk are highly correlated, we estimate two separate models. The first model includes age-related objective mortality risk and the second model includes respondents' ages, in addition to other explanatory variables. Based on a PE test (Greene, 2000), we can reject the null hypothesis of a linear specification in favor of a log-linear specification (PE coefficient 70.61, p-value 0.000), in the case of the first model. For the second model, however, we cannot reject neither the null hypothesis of a

[9] We do not focus on gender difference in any of the subsequent analyses of this paper.

linear specification (PE coefficient 18.97, p-value 0.285), nor the null hypothesis of a log-linear specification (PE coefficient 0.010, p-value 0.335). In Table 3, we present the results of log-linear specifications for both models.

Table 3. Ordinary least square estimates of subjective risk

Dependent variable	Log(subjective risk)	Log(subjective risk)
Received training in the survey	-0.104	-0.106
	(0.072)	(0.072)
Log (age-related objective risk)	0.501***	
	(0.038)	
Log (age)		-0.548
		(2.69)
Log(age)-squared		0.302
		(0.362)
Having chronic illness	0.149**	0.133*
	(0.074)	(0.075)
Low education	-0.011	-0.018
	(0.083)	(0.083)
High education	-0.043	-0.050
	(0.123)	(0.123)
Muslim religion	0.144*	0.147*
	(0.076)	(0.076)
Log (income per capita)	-0.070	-0.076
	(0.045)	(0.046)
Currently smoking	0.185**	0.153**
	(0.076)	(0.077)
Constant	2.08***	1.82**
	(0.462)	(5..01)
Adjusted R-square	0.206	0.220
Number of observations	765	765

Standard errors are in parentheses. Superscripts ***, **, * denote statistical significance at the 1 %, 5%, and 10% level, respectively.

We observe that the respondents' risk perception increases by only 0.5 % for a 1% increase in the average age-related objective risk. We find that respondent's health status significantly affects the perception of the risk of dying; people with chronic illness show a 16% higher risk perception (first model) compared to people not having any chronic illness. Although weakly significant, the Muslim respondents seem to have a 15% higher risk perception compared to the Hindu respondents.

We further observe that the smokers' perception of the risk of dying is 19% higher compared to non- smokers. Thus, it seems like smokers are quite aware of the health risk of smoking even in the rural areas of a developing country such as Bangladesh. On average, smokers' life expectancy is shorter than for non-smokers. For example, Shaw et al. (2000) estimate that average loss of life due to smoking is 6.5 years or 11 minutes per cigarette. Studies have shown that the risk of developing lung cancer is 22% higher for smokers and that the mortality risk from cardiovascular disease (heart disease) is almost double for smokers (including ex-smokers) compared to non-smokers (Newcomb and Carbone, 1992; ILO, 2002).[10] Given these estimates, it is hard to say whether smokers overestimate or underestimate the risk of dying in our case. The empirical evidence regarding smokers' risk perception is otherwise mixed.[11]

2.3 The CV scenario

Finally, the respondents were presented with the CV scenario (see Figure 5) and WTP questions to elicit their preferences for a risk reduction. The specific risk reduction (either a 25% or a 50% reduction) corresponds to the respondent's stated perceived risk of dying during the next five years and was communicated by separating the black squares representing perceived risk into 25-75 or 50-50 splits. The risk reductions, as told to the respondents, were to be achieved through participation in a program

[10] The risk for contracting other types of cancer is also relatively higher for smokers (Newcomb and Carbone, 1992).

[11] For example, studies in developed countries find that smokers overestimate the risk of getting lung cancer from smoking and that their assessed loss in life expectancy due to smoking is quite high (Viscusi, 1990; Viscusi 1992). For a sample of smokers in Sweden, Hammar and Johansson-Stenman (2004), however, did not find support for the conclusion that smokers overestimate the health risk from smoking. Slovic (2000) discusses the fact that particularly young smokers considerably underestimate the health risk due to smoking.

involving various vaccinations. Each respondent was asked about his/her maximum WTP for the stated risk reduction. If the respondent stated zero WTP for the risk reduction, he or she was asked several follow-up questions in order to ascertain possible scenario rejections.

Figure 5. CV scenario

Preventative vaccines could reduce the risk of dying from many infectious diseases.

Suppose that you could participate in a program involving various kinds of vaccinations against infectious diseases. The vaccines, if received, would reduce your risk of dying during the next five years.

Assume that the vaccines would be completely safe and would have no side effects. However, the effects of the vaccines would not last beyond the five-year period.

If received, such vaccines would reduce the risk of you dying over the next five years by one quarter/ one half.

[Enumerator: Show grid table 12 in which the stated risk from question C10 has been included by filling in the number of squares representing this subjective risk. Split the filled in area into 25-75%/50-50%. Then while mentioning the risk reduction, point at the 25% / 50% part of the split box and while mentioning the remaining risk on the other part of the split box.]

CV1. What is the maximum, as a one-time fee, you would be willing to pay to obtain such vaccines for yourself? You should also remember that if you were to pay for the vaccines, you would have less money left for other purposes.

Maximum ……Taka

3. WTP results

Approximately 10 % of the respondents stated zero WTP. Those who stated zero WTP were asked if they instead would want to receive the vaccination free of cost; 77 % of respondents with zero stated WTP mentioned that they would rather want the risk reduction free of cost. To ascertain scenario rejection, all these respondents (including the ones who would not want free vaccination) were asked why they would not be willing to pay for vaccinations; possible answers (reasons) were read to the respondents

and the respondents could choose more than one answer (see Table 4). Among those 79 respondents, 84 % had chosen more than one answer; we believe these responses indicate scenario rejection.[12] We do not include these responses in our further statistical analysis of this paper.

Table 4. Follow-up questions asked to respondents who stated zero WTP (N = 79)

		Sub-sample of respondents who would want free vaccination (77%)	Sub-sample of respondents who would not want free vaccination (23%)
	Reasons for not being willing to pay for vaccination [a]	% of respondents agree	
i)	I cannot afford vaccinations, even though I believe it is good to have them.	79%	-
ii)	I think the government should pay for the vaccinations.	77%	11%
iii)	I do not think the vaccine would really be safe.	-	17%
iv)	I do not think it is possible to reduce the mortality risk by vaccines.	7%	33%
v)	I do not believe in reducing mortality risk by any means.	7%	44%
vi)	Other reasons stated by the respondents: Reluctant to answer, not interested, dislike vaccination, not sure if (s) he would be willing to pay for vaccination.	3%	39%

[a] All the respondents who stated zero WTP were asked, "Why would you not be willing to pay for vaccinations?". Then a list of possible reasons were read to them and the respondents were allowed to choose more than one reason. They were also allowed to express other reasons.

However, we analyze the probability of scenario rejection using a standard probit model. In Table 5, we see that the only significant variable that explains scenario rejection is the respondent's religious belief; Muslim respondents are more likely to provide a protest zero. We also see that respondents receiving training in the survey are less likely to reject the scenario at the 10% level.

[12] Respondents who had chosen any response than (i) or had chosen more than one responses are believed to have provided protest zeros when answering the WTP question.

Table 5. Probit regression of scenario rejection (N=79)

Variable	Marginal effects	Standard error
Received training in the survey	-0.120	0.085
50% risk reduction	-0.067	0.081
Age in years	0.005	0.002
Low education [a]	0.144	0.093
Muslim religion	0.416***	0.145
Log (income per capita)	0.477	0.482
Having chronic illness	-0.065	0.084
Currently smoking	-0.138	0.109

Standard errors are in parentheses. Superscripts ***, **, * denote statistical significance at the 1 %, 5%, and 10% level, respectively.

[a] We cannot estimate the marginal effect of high education as all eight observations from this group are dropped because, for this group, all the zero WTPs imply a scenario rejection.

Alberini (2004) discusses about the robustness of CV estimates and different types of outliers. We look for WTP outliers in relation to income. There is no *a priori* reason to assume that WTP for reducing the risk of dying should be a small part of income, since the payment for the risk reduction was to be made once for a five-year period. We choose to keep WTP responses equal to or less than 50% of respondent annual income per capita. The coefficient of income remains roughly the same with this exclusion.[13] We finally have 692 observations for further statistical analysis. Table 6 reports the descriptive statistics of WTP for different sub-samples; in the Appendix, we present histograms of the distribution of WTP, where a visual inspection of WTP data reveals the existence of some focal points.

[13] However, the mean WTP (and not median WTP) for the sub-samples reduces between 1- 27 % with the exclusion of these responses.

Table 6. WTP results for different sub samples [a]

	No training		Training	
	25% risk reduction	50% risk reduction	25% risk reduction	50% risk reduction
Sample size	162	168	175	189
Mean WTP	487	672	671	970
Standard deviation	1531	1934	1377	1324
Median WTP	100	200	500	500
Mean WTP ratio [b]	1.38		1.45	
Null Hypothesis: Mean WTP ratio=1 [c]	p- value <0.001		p- value <0.001	
VSL based on changes in subjective risk				
Mean VSL	1,03,074	1,06,585	1,68,905	1,07,697
Median VSL	20,000	13,333	33,333	30,000
95% confidence interval for mean VSL	43,742 – 1,62,407	32,164 – 1,81,005	1,03,714 – 1,34,097	81,167 - 1,34,228
VSL based on changes in average age related objective risk				
Mean VSL	81,861	56,539	1,31,353	75,617
Median VSL	18,118	9,615	36,363	18,461
95% confidence interval for mean VSL	32,499 – 1,31,225	30,142 – 82,936	82,932 – 1,79,774	54,268 – 96,967

[a] WTP and VSL are expressed in Bangladesh Taka. 57.8 Taka =1 US $, at the time of survey (October 2003).

[b] Ratio of mean WTP for a 50% risk reduction to mean WTP for a 25% risk reduction.

[c] Using both the non-parametric Wilcoxon –Man-Whitney test and the *t* test.

We find that the mean WTP is TK. 487 (TK. 672) for a 25% (50%) risk reduction, for the no-training sub-sample. For the trained sub-sample, the mean WTP is TK. 671 (TK. 970) for a 25 % (50%) risk reduction. Thus, we find that the training in the CV survey increases the mean and reduces the variances of WTP. We can reject the hypothesis that for the specific risk reduction (either 25% or 50%), WTP for the two sub-samples (training and no training) comes from the same underlying distribution (in both cases, p-value<0.001). As reported in Table 6, the results also indicate a smaller variation in the CV responses for the training sub-sample, which in turn implies that a

brief training facilitates respondent ability to better process the risk information, and hence yields lower variation in the CV responses. We conduct more tests regarding the differences in WTP in the econometric analysis in Section 4.

3.1 Implied value of statistical life

We calculate the individual VSL from the data by dividing individual WTP by the changes in the individually perceived risk; the mean VSL is in the US$ 1,783 to US$ 2,922 range. As depicted in Table 6, we also calculate the individual VSL by dividing the changes in the individual's average age-related mortality risk. We observe a different pattern for the mean VSL (but not for the median VSL) based on the subjective risk changes, between the trained and the not trained sub-samples. In the no-training sub-sample, the mean VSL is higher for the higher (50%) risk reduction level, and in the trained sub-sample, the mean VSL is higher for the lower (25%) risk reduction levels. If the observed WTP were less than proportional to the magnitude of the risk change, we would obtain a lower VSL for the higher risk reduction level. The result we mentioned above is due to some very high WTPs related to very low perceived risks, in the case of the no-training sub-sample. This pattern is not observed for VSL based on objective risk changes; the mean VSL is lower for the higher risk reduction level.

The magnitude of the VSL, however, is very low compared to the available estimates for developing countries.[14] For example, using results from several VSL studies, Miller (2000) predicts a VSL for Bangladesh in the range of US$ 30,000 to US$ 1,000,000. The lower absolute values of VSL in our case may be attributed to the fact

[14] Using data from the Indian labor market, Shanmugam (2000) provides VSL in the range of US $ 0.76 million -$1.026 million and Simon et al. (1999) provide VSL for India from an independent wage-risk study in the range of US$ 0.15- US$ 0.35 million.

that unlike many other studies we had relatively large risk reductions, if we assume that there is inadequate sensitivity to scope. For example, Carlsson et al. (2004) suggests that VSL tends to decrease rapidly when the size of the risk reduction increases.

Moreover, stating WTP for a risk reduction is somewhat difficult for people unfamiliar with the concept of trading income for risk reduction. Therefore, it is likely that people would suffer from initial anchoring when constructing an answer as to how much they would be willing to pay (Tversky and Kahneman, 1974; Kahneman et al., 1982; Green and Tunstall, 2001). For example, the respondents might anchor on the price of vaccination or on their other expenditures. People in rural areas, in a developing country like Bangladesh, have a general perception that vaccinations are to be provided free by the government, as is the case with the children's vaccination programs. For example, Cropper et al. (2004) estimated demand for a hypothetical malaria vaccine in rural Ethiopia and their results suggest that at very low prices few vaccines are purchased; "at an annualized price of US$ 3, half of the households in Tigary would purchase no vaccines." Moreover, in the context of a developing country, household consumption expenditures are usually low on average, particularly in the rural areas. Therefore, when placing a value on a desired and substantial risk reduction, the respondents might anchor initially to such low expenditures, and adjust thereafter. In addition, the incidence of financial limitations coupled with poorly functioning credit markets also results in lower WTPs, particularly if the respondents are asked for one-time payments rather than continuing monthly or yearly payments.[15] Training seems to

[15] As observed by Carson (2000), "A one-time payment generally produces more conservative estimates since it does not offer the opportunity to spread payments over time," compared to a continuing payment (p. 1416).

reduce this potentially downward bias in WTP as we observed a significantly higher WTP for the training sub-sample.

4. Econometric analysis

An appropriate econometric model for analyzing the WTP data that also includes zeros would be Tobit with selection, as it allows for modeling zero and positive WTP separately (for a discussion, see Carlsson and Johansson-Stenman, 2000). However, with only 13 (non-protest) observations with a zero WTP, it is doubtful whether a sample selection model can be justified in our case. Moreover, it is unclear whether there is any true negative WTP. Therefore, we estimate a truncated regression model where WTP is truncated at zero. Based on a PE test (p-value = 0.017) we can reject a linear specification in favor of a log linear one. We use log (WTP + 1) as the dependent variable, which is truncated at zero.

In the previous section, we observed that the distributions of WTP significantly differ between the sub-samples concerning training. Given this result, we first estimated a model allowing for heteroscedasticty concerning training, where dummy variables are included identifying training in pooling the data. However, the heteroscedastic term is not significantly different from zero (p value=0.96). Therefore, we cannot reject the null hypothesis of homoscedasticity; the variances of WTP between the two sub-samples do not differ significantly. Hence, we estimate the model assuming homoscedasticity and dummy variables are included identifying training and risk reduction levels in pooling the data. It is expected that people who passed the entire training would show higher sensitivity to scope compared to other respondents; hence, we include separate interaction variables.

We first estimate two models; one assuming that there are no enumerator effects on WTP responses, and the other taking the enumerator effects into consideration by including dummy variables for the enumerators (see e.g. Köhlin, 2001). Out of 13 enumerators, we found only one enumerator to be significant and negative. Based on a likelihood ratio test, we can reject the null hypothesis of "no enumerator effect" (p-value 0.013). Therefore, we present the results from the latter model in Table 7, where we control for the enumerator effects.

The coefficient "training" is highly significant and implies that WTP for a larger reduction in risk is 79% ($e^{0.58} - 1$) higher for the group receiving the training. The coefficient on a 50% reduction is highly significant and indicates that WTP for the larger risk reduction is 45% higher than the WTP for the smaller reduction. The WTP difference concerning risk reduction is even higher for the training sub-sample (16% higher) in general and for the group who passed the tests (6% higher) in particular, although these differences are not statistically significant. We find that people with higher levels of prior education have on average 7% higher WTPs compared to illiterate people; however, this difference is not statistically significant.

Table 7. Estimated WTP by sub-samples: truncated regression model [a]

Dependent variable Log (WTP+1)		
Variable	Coefficient	Standard error
Constant	15.12**	6.89
Received training	0.583***	0.146
Passed probability test	0.170	0.230
50% risk reduction	0.371***	0.137
Received Training × 50% risk reduction	0.146	0.203
Passed probability test × 50% risk reduction	0.060	0.310
Muslim religion	-0.005	0.092
Log(age)	-7.02*	3.73
Log(age)-squared	0.912*	0.502
Log (subjective risk)	0.021	0.052
High education	0.070	0.090
Low education	-0.064	0.074
Log(income per capita)	0.425***	0.063
Having chronic illness	-0.010	0.102
Currently smoking	0.016	0.102
Self reported happiness	0.070***	0.023
Disturbance standard deviation	1.22	0.033
Log-Likelihood	-1100.12	
Number of observations	692	

Superscripts *** and ** denote statistical significance at the 1 % and 5% levels, respectively.

[a] We control for the enumerator effects but do not present the parameter estimates of the enumerator dummies.

We also estimate a separate model controlling for the WTP difference concerning the training as well as the risk reduction levels by interacting these variables with respondents' educational background variables. However, the interaction terms are not significant, based on a likelihood ratio test (p-value 0.36); hence, we can accept the results of the model presented in Table 7. Although it can be expected that people with higher levels of education might have higher values for risk reduction, other studies, e.g.

Krupnick et al. (2002) and Alberini et al. (2004), have found that more highly educated people report lower WTPs.

The estimated marginal effects for income[16] are positive and significant, with an income elasticity of 0.43. The result that the income elasticity of the VSL is well below unity is also found in many other CV studies (e.g. Carlsson et al., 2004; Persson et al., 2001). In cross-country comparisons of VSL studies, Miller (2000) as well as Viscusi and Aldy (2003) found the income elasticity of VSL in the range of 0.85 to 1 and 0.5 to 0.6, respectively.

We find that the effect of age on WTP is negative at younger ages, until a minimum is reached at age 46, and then increases. Alberini et al. (2004) found, using a sample over 40-year olds that WTP does not decline until age 70. Although subjective risk positively affects respondents' WTPs for risk reduction, this is not significant.[17] The result that subjective risk does not have any significant effect on WTP for risk reduction is somewhat surprising, but consistent with other studies in developed countries (e.g. Corso et al., 2001). We also observe that having a chronic illness has no significant effect on WTP, which is consistent with the finding of Alberini et al. (2004) that having a chronic condition does not reduce the WTP for mortality risk. Although smokers' risk perceptions were higher compared to non-smokers, being a smoker does not significantly increase the WTP. We find that the level of overall individual

[16] It should be noted here that the distribution of income is highly skewed and hence we estimated separate models excluding relatively high income. However, as the coefficient of income is roughly the same, we decided to keep them in our final model presented here.

[17] As respondent age and subjective risk might be correlated (risk perception is based on age-related objective risk), we also estimate a separate model excluding the subjective risk. However, the results are roughly the same.

happiness significantly and positively affects WTP for risk reduction, all else remaining the same.

From the estimated results, we formally test for the sensitivity to scope. We calculate the mean WTP ratio, i.e. ratio of mean WTP for a 50% risk reduction to the mean WTP for a 25% risk reduction, using the regression coefficients. The WTP ratios for both the sub-samples are presented in Table 8. We obtain the standard errors for these ratios as well as for the differences between these two ratios using the Delta method (Greene, 2000) and hence, construct the corresponding confidence intervals.

Table 8. Sensitivity to scope

	No training	Training
Mean WTP ratio[a]	1.45	1.51
Standard error	0.198	0.160
95 % confidence interval	1.06 - 1.84	1.20 – 1.83
Difference of mean WTP ratio	0.065	
Standard error	0.073	
95% confidence interval	-0.08 – 0.21	

[a] Ratio of mean WTP for a 50% risk reduction to mean WTP for a 25% risk reduction. The ratio is calculated using the regression coefficients of the dummy variable for risk reduction and the mean values of other explanatory variables.

We can reject the hypothesis that WTP is insensitive to the magnitude of risk reduction for both the no-training and training sub-samples. We also find that sensitivity to scope is higher (higher WTP ratio) for the trained sub-sample; however, the difference between two WTP is not statistically significant.

5. Discussion and Conclusion

Past studies have discussed that it is problematic to consistently measure the value of statistical life using the CV method, largely due to the cognitive burden that the respondents face when comparing expected welfare effects of a small reduction in the risk to those of small monetary changes (Beattie et al., 1998; Hammit and Graham, 1999; Carlsson et al., 2004). However, the main objective of this study was not to obtain an absolute measure of VSL. Rather, we used substantial risk changes to be valued and measured the effect of training regarding probability and risks on the WTP responses and on the sensitivity to scope.

In the survey, we customize the mortality risk for respondents by asking them to report their perception of the risk of dying based on their health and information about age-related risks of dying provided in the survey. We find that people on average overestimate the risk at younger ages and underestimate the risk at older ages. This result is consistent with previous studies in the context of developed countries.

We find a significant difference in the WTP distributions between the sub-sample receiving training and the sub-sample receiving no training. However, we find no significant difference in the variance, but rather in the levels. We have found that estimated WTP is sensitive to the size of the risk reduction in both the sub-samples. Although sensitivity is higher for the trained sub-sample, the difference is not statistically significant, which implies that the training does not affect the sensitivity to scope. Although the implied VSL is higher for the trained sub-sample, it is still substantially lower compared to other studies, which may be attributed to the fact that compared to other studies we used relatively large risk reductions. Moreover, the respondents in the survey were asked for a one-time payment, rather than for continuing

monthly or yearly payments, and there is possibility of respondents anchoring on the often zero price that people pay for vaccination in reality.

Overall, it appears constructive to train the respondents regarding probabilities and risk reductions. Training reduces the extent of cognitive burden that the respondents face and thus increases the ability of the respondents to value the risk reduction in a situation where the respondents are not familiar with the notion of probabilities and/or risk/–money trade-offs. As discussed earlier, the respondents are likely to suffer from initial anchoring in stating WTP for the risk reduction, which may be related to their other expenditures that are usually low and which may result in a downward bias in WTP. Training seems to reduce this potentially downward bias, since WTPs are substantially higher in the trained sub-sample. However, there might also be some problems with providing training in the CV survey. The respondents may get tired if they find it boring and this may cause fatigue effects. Moreover, by talking a lot about uncertainties and probabilities, the respondent can get the impression that avoiding risks is very important. Hence, they will tend to state higher WTP in the training version. This is then not because they are better trained but because they think that it is expected of them. However, while some respondents may respond in this way, others are able to draw inferences about the risk reduction, and training facilitates a cognitive structure that is essential to draw such inference in such a situation.

Finally, using the CV method to elicit people's VSLs is not a "mission impossible." CV risk-reduction can be performed in a developing country with very low levels of education. A comprehensible training on probability and risk concepts, interspersing risk examples with questions to maintain respondent interest as well as to check understanding, should be given before presenting the CV scenario of risk

reduction to the respondents. There are remaining problems but most of these appear to be related to the CV methodology *per se*, rather than to CV studies being performed in developing countries.

References

Alberini, A. (2004), Robustness of VSL Values from Contingent Valuation Surveys, FEEM Working Paper No. 135.04, November.

Alberini, A., Cropper, M., Krupnick, A. and Simon, N. B. (2004), Does the value of a statistical life vary with age and health status? Evidence from the US and Canada, Journal of Environmental Economics and Management, 48(1): 769-792.

Balistreri, E., McClelland, G., Poe, G., and Schulze, W. (2001), Can Hypothetical Questions Reveal True Values? A Laboratory Comparison of Dichotomous Choice and Open-Ended Contingent Values with Auction Values, Environmental and Resource Economics, 18: 275-292.

Bateman, I.J., Carson, R.T., Day, B., Hanemann, W.M., Hanely, N., Hett, T., Jones-Lee, M., Loomes, G., Murato, S., Ozdemirogla, E., Pearce, D.W., Sugden, R. and Swanson, J. (2002), Guidelines for the Use of Expressed Preferences Methods for the Valuation of Preferences for Non-market Goods, Edward Elgar Publishing, Cheltenham.

Beattie, J., Covey, P. Dolan, L. Hopkins, M, Jones-Lee, G., Pidgeon, A. Robinson, and Spencer, A. (1998), On the Contingent Valuation of Safety and the Safety of Contingent Valuation: Part 1- Caveat Investigator, Journal of Risk and Uncertainty 17(3): 5-25

Carlsson, F. and Johansson-Stenman, O. (2000), Willingness to Pay for Improved Air Quality in Sweden, Applied Economics, 32: 661-670.

Carlsson, F., Johansson-Stenman, O. and Martinsson, P. (2004), Is Transport Safety More Valuable in the Air? , Journal of Risk and Uncertainty, 28:2; 147-163.

Carson, R.T. (2000), Contingent Valuation: A User's Guide, Environmental Science and Technology, 34(8): 1413-1418.

Corso, P.S., Hammitt, J.K. and Graham, J. D. (2001), Valuing Mortality –Risk Reduction: Using Visual Aids to Improve the Validity of Contingent Valuation, Journal of Risk and Uncertainty, 23(2): 165-184.

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

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

Gilovich, T., Griffin, D., and Kahneman, D (2002), Heuristics and Biases: The Psychology of Intuitive Judgment, Cambridge, Cambridge University Press.

Green, C., and Tunstall, S. (2001), A Psychological Perspective, In: Bateman, I.J., and Willis, K.G, (eds), Valuing Environmental Preferences, Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries, Oxford, Oxford University Press.

Greene, W. H. (2000), Econometric Analysis, Fourth Edition, Prentice Hall, USA.

Hammar, H. and Johansson-Stenman, O. (2004), The value of risk- free cigarettes- do smokers underestimate the risk? Health Economics, 13: 59-71.

Hammitt, J.K. (2000), Valuing Mortality Risk: Theory and Practice, Environmental Science and Technology, 34(8), 1396-1400.

Hammitt, J.K. and Graham, J. D. (1999) Willingness to pay for health protection: Inadequate sensitivity to probability, Journal of Risk and Uncertainty 18, 33-62

Hanemann, W. M. (1994), Valuing the Environment through Contingent Valuation, The Journal of Economic Perspectives, 8(4), 19-43.

Hanley, N., Ryan, M., and Wright, R. (2003), Estimating the monetary value of health care: lessons from environmental economics, Health Economics, 12, 3-16.

ILO (2002), Cardiovascular System: Relative mortality risk for smokers, http://www.ilo.org/public/protection/safework/tobacco/encyclo/car01fe.htm

Johansson, P.O. (2002), The value of a statistical life: theoretical and empirical evidence, Applied Health Economics and Health Policy, 1(1), 33-41.

Kahneman, D. and Knetsch, J.L. (1992), Valuing Public Goods: The Purchase of Moral Satisfaction, Journal of Environmental Economics and Management, 22(1), 57-70.

Kahneman, D., and Tversky, A. (2000), Choices, Values and Frames, Cambridge, Cambridge University Press.

Kahneman, D., Slovic, P. and Tversky, A. (1982), Judgment Under Uncertainty: Heuristics and Biases, Cambridge, Cambridge University Press.

Köhlin, G. (2001), Contingent Valuation in project planning and evaluation: the case of social forestry in Orissa, India, Environment and Development Economics, 6:237-258.

Krupnick, A., Alberini, A., Simon, N., O'Brien, B., Goeree, R. and Heintzelman, M. (2002), Age, Health and the Willingness to Pay for Mortality Risk Reductions: A Contingent Valuation Survey of Ontario Residents, Journal of Risk and Uncertainty, 24:2, 161-186.

Miller, T. R (2000), Variations between countries in values of statistical life, Journal of Transport Economics and Policy, 34, 169-188.

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

NOAA (1994), National Oceanic and Atmospheric Administration, Natural Resource Damage Assessments: Proposed Rules, Federal Register 59, 1062-1191

Newcomb, P.A. and Carbone, PP (1992), The health consequences of smoking: cancer, In: Fiore MC, Ed. Cigarette Smoking: A Clinical Guide to Assessment and Treatment, Philadelphia, PA: WB Saunders Co: 305-331, Medical Clinics of North America, http//www.cancer.org.

Persson, U., Norrinder, A., Hjalte, K. and Gralen, K. (2001), The Value of a Statistical Life in Transport: Findings from a New Contingent Valuation Study in Sweden, Journal of Risk and Uncertainty 23(2): 121-134.

Shanmugam, K.R. (2000), Valuations of Life and Injury Risks: Empirical Evidence from India, Environmental and Resource Economics 16:379-389.

Shaw, M., Mitchell, R. and Dorling, D. (2000), Time for a smoke? One cigarette reduces your life by 11 minutes, British Medical Journal, 320(1).

Siegel, S. and Castellan, N.J, 2000, Nonparametric statistics for the Behavioral Sciences, McGraw-Hill.

Simon, N.B., Cropper, M. L., Alberini, A. and Arora, S. (1999), Valuing Mortality Reductions in India: A Study of Compensating Wage Differentials, World Bank.

Slovic, P. (2000), The Perception of Risk, London, Earthscan.

Tversky, A. and Kahneman, D. (1974), Judgment Under Uncertainty: Heuristics and Biases, Science, 185: 1124-1131

Tversky, A. and Kahneman, D. (1992), Advances in Prospect Theory: Cumulative Representation of Uncertainty, Journal of Risk and Uncertainty, 5, 297-323

Viscusi, W. K., (1985), A Bayesian Perspective on Biases in Risk Perception, Economics Letters 17, 59-62.

Viscusi, W.K., Magat, W. and Huber, J. (1987), An Investigation of the Rationality of Consumer Valuations of Multiple Health Risk, RAND Journal of Economics, 18, 18, 465-479.

Viscusi, W. K. (1989), Prospective Reference Theory: Toward an Explanation of Paradoxes, Journal of Risk and Uncertainty, 2, 235-263.

Viscusi, W. K. (1990), Do smokers underestimate risks? Journal of Political Economy, 98(6), 1253-1269

Viscusi, W. K. (1992), Smoking: Making the Risky Decision, Oxford University Press, New York.

Viscusi, W.K. (1992), Fatal Tradeoffs: Public and Private Responsibilities for Risk, New York, Oxford University Press.

Viscusi, W. K. (1993), The Value of Risks to Life and Health, Journal of Economic Literature 31, 1912-1946.

Viscusi, W.K. and Aldy, J.E. (2003), The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World, Journal of Risk and Uncertainty, 27(1): 5-76.

Weinstein, M.C., Shepard, D.S., and Pliskin, J.S. (1980), The Economic Value of Changing Mortality Probabilities: A Decision Theoretic Approach, Quarterly Journal of Economics, 94(2), 373-396.

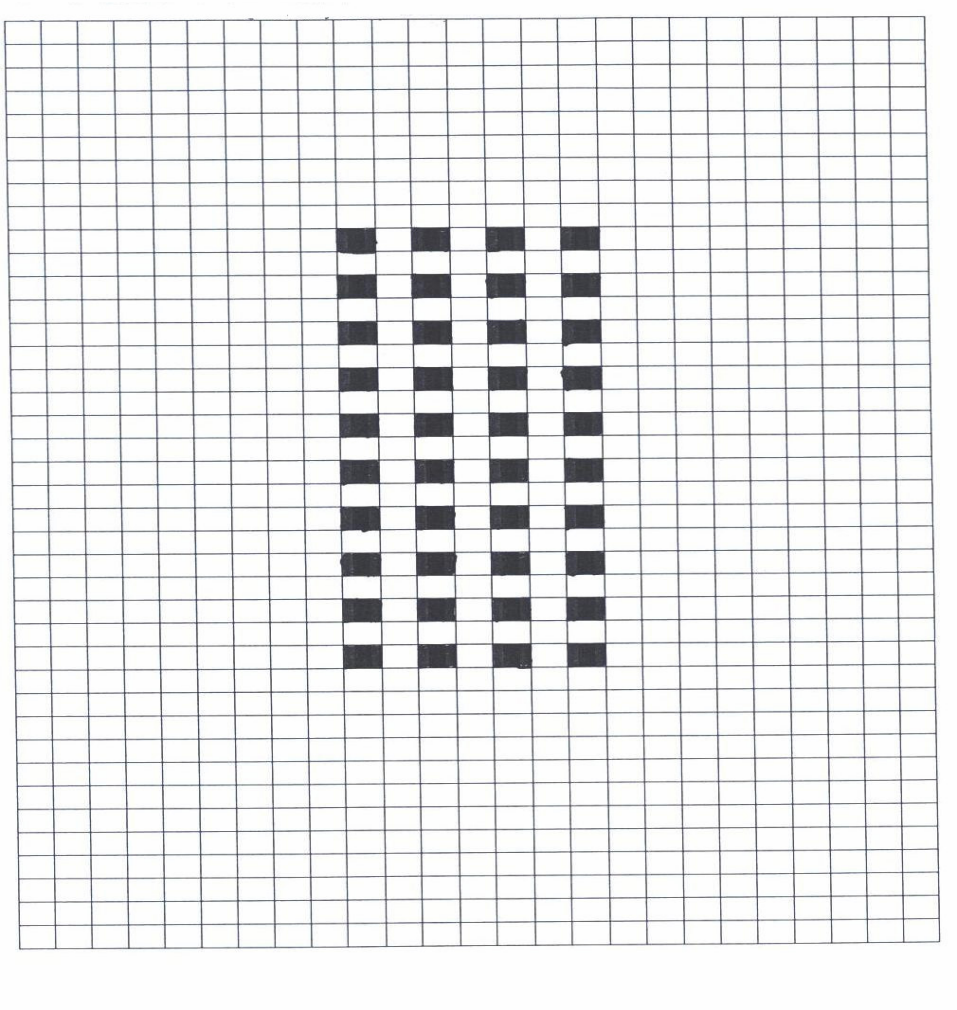
Whittington, D (1998), Administering Contingent Valuation Surveys in Developing Countries, World Development, 26(1), 21-30.

Whittington, D. (2002), Improving the Performance of Contingent Valuation Studies in Developing Countries, Environmental and Resource Economics, 22, 323-367.

WHO (2004), World Health Organization, Life Tables for 191 Countries, World Mortality in 2000, http://www3.who.int/whosis/life_tables

Appendix 1. Visual aid for risk understanding.

Figure 1. Grid table showing mortality risk of an adult in the next five years as 40 in 1000



Appendix 2. Distribution of WTP (Histogram)

Figure 1. Distribution of WTP for sub-sample: no training and 25% risk reduction

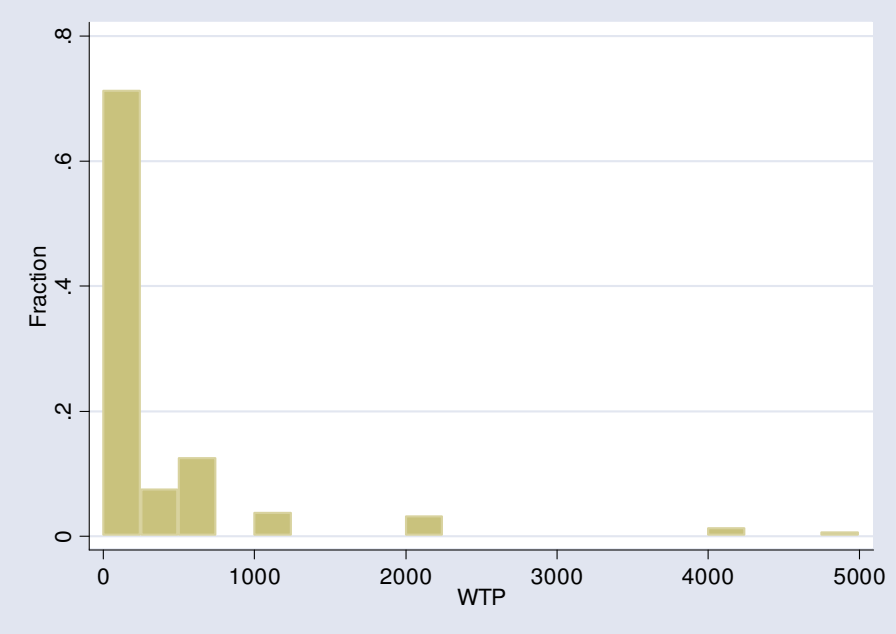


Figure 2. Distribution of WTP for sub-sample: training and 25% risk reduction

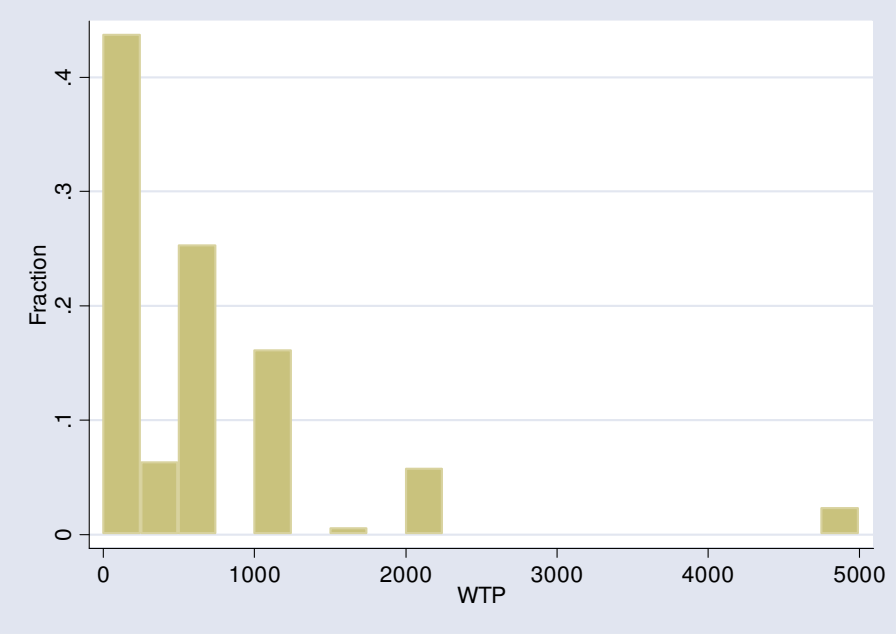


Figure 3. Distribution of WTP for sub-sample: no-training and 50% risk reduction

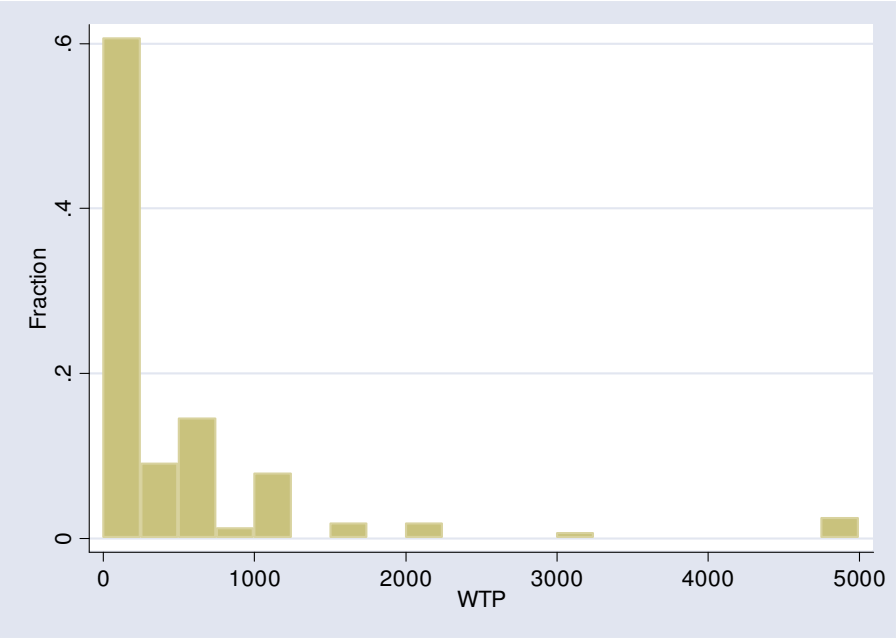
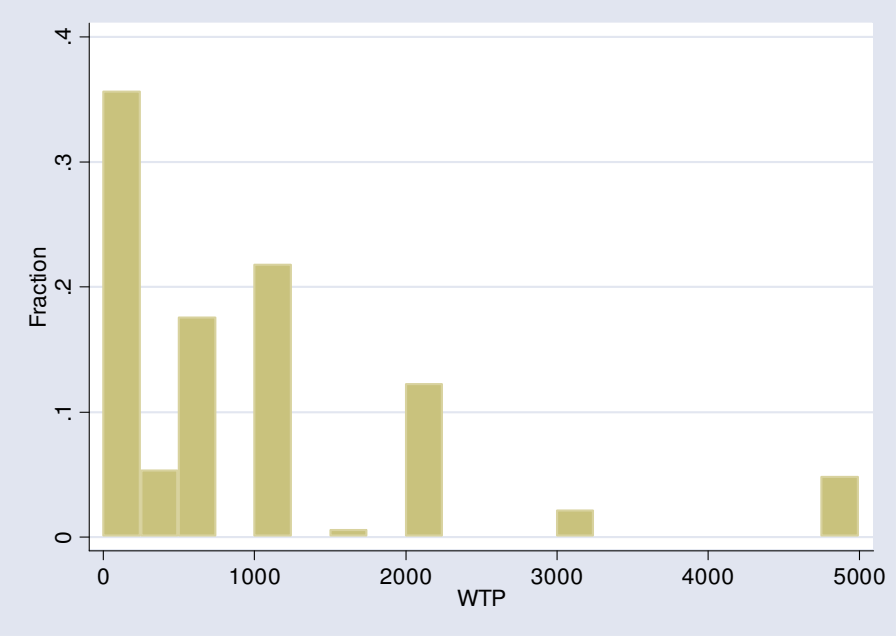


Figure 4. Distribution of WTP for sub-sample: training and 50% risk reduction



Using Choice Experiments to Measure Relative Values of Statistical Lives: Evidence from Bangladesh

Minhaj Mahmud*
Department of Economics, Göteborg University
Vasagatan 1, Box 640, SE- 405 30 Göteborg, Sweden

Abstract

By assuming that an individual has preferences concerning different states of the world and these preferences can be described by an individual social welfare function, we explore the relative value of statistical life using survey data from Bangladesh. We apply a pair-wise choice experiment on life-saving programs to elicit individuals' preferences regarding differences in the values of statistical lives related to age. We find that the relative value decreases strongly with age and that people have strong preferences for saving more life-years, rather than lives *per se*. Moreover, in specific follow-up questions, it is again elicited that a majority of the respondents believe that it is better, from a social point of view, to save younger individuals.

Key words: social preference; choice experiment; life-saving programs; relative value of life; Bangladesh.

JEL Classification: D63, I18, J17

* I am most grateful for very useful suggestions and comments from Olof Johansson-Stenman, Peter Martinsson, Håkan Holm, Fredrik Carlsson, and seminar participants at Göteborg University. Financial assistance from the Swedish International Development Agency is gratefully acknowledged.

1. Introduction

The value of a statistical life (VSL), in monetary terms, has been studied for a long time using both revealed and stated preference approaches (see e.g. Weinstein et al., 1980; Viscusi 1993; Hammitt and Graham 1999; Viscusi and Aldy, 2003). The idea of using varying values of statistical lives in decision making is more recent, and has been frequently discussed since the publication of the 1993 World Development Report (World Bank, 1993). The report states on page 213 that "Most societies attach more importance to a year of life lived by a young or middle-aged adult than to a year of life lived by a child or an elderly person." Murray and Lopez (1994) argued that "In all societies social role varies with age. The young, and often the elderly, depend on the rest of society for physical, emotional and financial support. Given different roles and changing levels of dependency with age, it may be appropriate to consider valuing time lived at a particular age unequally"(p. 8). The disability adjusted life year (DALY)[1] age weights, proposed by Murray (1994; 1996) and used by the World Bank (1993) in measuring the global burden of diseases, imply that the value for each year of life lost rises steeply from zero for newborns until a peak at the age of 25 and then declines gradually with increasing age, while remaining positive. The World Health Organisation regularly publishes burden of disease results in the World Health Reports, and hence endorses its application and development (see e.g. World Health Report, 2004). The DALY age weights also imply that the value of a statistical life, commencing from infanthood, strongly decreases with age (see Figure 1 and Appendix 1). Such age

[1] DALY measures health outcome in terms of losses. The sum of DALYs lost across all ages, conditions, and regions are referred to as the global burden of disease, which reflects both the number of "years of life lost due to premature death," as well as the number of "years of life lived with disabilities," due to each disease (see World Bank 1993; Murray, 1994, 1996; Murray and Acharya, 1997).

weights are defended by the argument that they seem to be appealing to most people and also that the discriminatory effect is ethically quite different from setting un-equal age-weights regarding sex, ethnicity, education, or income, since everybody is supposed to experience each age (Murray, 1996). Priority setting in health care on the basis of the burden of disease calculations (thus on the basis of the DALY age weights) is still a matter of debate (see e.g. Williams 1999, 2000; Murray and Lopez, 2000; Mooney and Wiseman, 2000). Some critics such as Williams (1997, 1999, 2000) suggest that such age weights may be relevant for developing countries.

There now exists a large theoretical literature in economics and in philosophy that discusses the issue of age related preferences (or ageism) in the context of health care and life-saving (for review see Hausman and McPherson, 1996; Broome, 1999; Tsuchiya, 1999, 2000; Cookson and Dolan, 2000; Williams and Cookson, 2000; Tsuchiya et al., 2003). For example, Tsuchiya (1999) discusses two different rationales for age weighting health benefits as "efficiency based age weighting" and "equity based age weighting," The three main forms of ageism discussed are: (i) "health maximization ageism"- giving priority to younger versus older since younger will experience an analogous health gain for a longer period; (ii) "productivity ageism"- giving priority to young adults as they are more productive in the family and in the society; and (iii) "fair innings ageism"- relates to quality adjusted life years[2] (QALYs) and gives priority to a younger person over an older because the former has a smaller number of expected lifetime QALYs than the older.[3] "Fair innings weights" reflect aversion to inequality in

[2] QALYs measure health outcomes in terms of gains in health, as opposed to DALYs (see Weinstein and Stason, 1997; Dolan, 2000; Hammit, 2002).

[3] This is based on Williams (1997) fair innings argument: "It reflects the feeling that everyone is entitled to some normal span of health (usually expressed in life years, e.g. 'three score years and ten'). The

lifetime QALYs and are supposed to decrease with age to reflect that the expected lifetime QALY increases with survival (Tsuchiya, 2000, p. 58). The first two forms of ageism are related to "efficiency based age weighting," while the last form of ageism is related to "equity based age weighting." The DALY age weighting has been referred to as" efficiency based age weighting," which relates to "productivity ageism" (Tsuchiya, 1999).

However, there are also some empirical studies that elicit peoples' preferences regarding the question of age related priority setting, mostly conducted in developed countries. For example, in the case of health care many people prefer that young should be given priority over the old (for reviews see Tsuchiya, 1999; Williams and Cookson, 2000). Studies quantifying trade-offs between saving lives at different ages also observe that people place more weight on saving younger people's lives than on saving the lives of older people (Cropper et al., 1994; Johannesson and Johansson, 1997). Both these seminal studies, conducted in the USA and Sweden respectively, estimated trade-offs between lives saved at different ages. Cropper et al. (1994) found that saving eleven 60 year olds was judged equivalent to saving one 30 year old. Johannesson and Johansson (1997) found that saving five 50 year olds or saving thirty-four 70 year olds was judged equivalent to saving one 30 year old. The result from these two studies also predicts that the value of saving lives decreases promptly with advancing age. More recently, Johansson-Stenman and Martinsson (2004), using an ethical preference approach and a choice experiment about safety enhancing road investment in Sweden, also found that the relative value of a saved life decreases by age but at a much slower rate; e.g. saving two 70 year old pedestrians was judged equivalent to saving one 30 year old pedestrian.

---

implication is that anyone failing to achieve this has in some sense been cheated, whilst anyone getting more than this is living on a borrowed time" (p. 119).

The objective of this paper is to, in a developing country context, investigate people's preferences regarding using relative values of statistical lives when it comes to different ages of individual being saved. This is elicited by using a pair-wise choice experiment between different life-saving programs. Moreover, our approach allows us to test if people have preferences for saving more life-years rather than only saving lives. The findings of this study are expected to provide new insight into the issue of age weighting in decision making regarding public health policies, particularly from a developing country perspective.

We elicit individuals' preferences for life-saving programs using a random sample of households in rural Bangladesh. As a developing country, Bangladesh has a much less developed social security system, reflecting almost non-existence of child-care as well as elderly-care systems. The social institutions that these vulnerable groups rely on are the family or the extended family. Bangladesh also has a long cultural and religious tradition of looking after elderly people, where families and communities are expected to take care of their own elderly members, and a cultural norm that also suggests that older people should be treated with respect (Banglapedia, 2003). Unlike developed countries, the income in a developing country like Bangladesh, peaks in the 20-40 age range since the productivity of workers, which often is of manual nature, peaks in the same age range. With the population ageing[4] these prime age adults simultaneously bear the responsibility of supporting their children as well as their elderly parents. Given such socio-economic and cultural setting, one might expect that people in a developing country like Bangladesh would have fairly different perceptions

[4] In a developing country like Bangladesh, chronological age has little importance in defining old age, but the age of 60 years seems to be a sensible statistical definition for ageing, particularly taking into account average retirement age, health conditions etc. (Banglapedia, 2003).

of valuing lives of different ages, i.e. perceive smaller differences in the relative values of life, compared to people in most developed countries such as the USA or Sweden. This is not what we find in our empirical analysis, however.

Following an approach similar to Johansson-Stenman and Martinsson's (2004), we apply a pair wise choice experiment (see e.g. Alpizar et al., 2003, and Louviere et al., 2000) to elicit individuals' preferences for relative values of statistical lives related to age. The respondents were presented with two life-saving programs at a time, each containing information on the number of individuals saved and the ages of the saved persons. To each respondent, several choice sets were presented from which he/she was asked to choose the preferred life-saving program. We find that relative value decreases strongly with age of the individuals saved and people have stronger preferences for saving more life-years, rather than lives.

The rest of the paper is organized as follows: Section 2 discusses the theoretical and empirical model, Section 3 presents the design of the choice experiment, Section 4 discusses the econometric results, and Section 5 discusses the responses from follow-up questions asked after the choice experiment. Finally, Section 6 concludes the paper.

2. The model

We assume that an individual has ethical social preferences concerning different states of the world, which can be described by an individual social welfare function (ISWF). We also assume that individuals act as social planners who maximize their own ISWF with regards to life-saving programs. We start with a general ISWF that includes the number of saved lives in different age groups, following Johansson-Stenman and Martinsson (2004). Henceforth, we call this the general model. Based on the ISWF and

the assumption of no discounting, we can calculate the individual social marginal rates of substitution (SMRS) between saved lives of people of different ages. In this paper, we also test a more restrictive model, denoted restricted model, where the ISWF depends on the total number of saved lives (irrespective of age) and the total amount of life-years saved. This model allows us to test and compare the hypotheses *only lives matter* and *only life-years matter*.

We begin with the general model. Let us consider small changes in the number of saved people of different ages, and a corresponding local linearization in these variables, implying that we can write

$$W_i = \hat{W}_i + \alpha_i^1 s^1 + ... + \alpha_i^n s^n \,, \qquad (1)$$

where $\hat{W}_i$ denotes *i*'s social welfare at status quo, i.e. without any saved lives from the programs, s^j is the number of saved people in the group *j* (e.g. corresponding to a certain age), and α_i is the vector of coefficients for individual *i*. The SMRS between group *j* and group *k* for individual *i* is expressed as follows:

$$SMRS_i^{jk} = \frac{\partial W_i}{\partial s^j} \Big/ \frac{\partial W_i}{\partial s^k} = \frac{\alpha_i{}^j}{\alpha_i{}^k} \qquad . \qquad (2)$$

The $SMRS_i^{jk}$ thus gives the ratio of saved lives of age group *j* to age group *k*, for which the individual remains indifferent.

In line with the random utility approach (McFadden 1974), we assume that the true ISWF is not directly observable and hence consists of both an observable and a non-observable (stochastic) part. By introducing a random error term, ε_i, to reflect unobservable characteristics, we can rewrite the above equation as

$$W_i = \hat{W}_i + \alpha_i^1 s^1 + ... + \alpha_i^n s^n + \varepsilon_i \,. \qquad (3)$$

An ISWF-maximizing respondent prefers a project *A* over a project *B* if $W_i(A) > W_i(B)$. Based on the observable information, we can then model the probability that *A* is chosen as follows:

$$\Pr(A \text{ is chosen}) = \Pr(W_i(A) > W_i(B)) = \Pr\left(\alpha_i^{\,1}\Delta s^1 + ... + \alpha_i^{\,n}\Delta s^n > \phi_i\right), \quad (4)$$

where $\Delta s^k = s^k(A) - s^k(B)$ and $\phi_i = \varepsilon_i(A) - \varepsilon_i(B)$. Given that ϕ_i is standard normal distributed, equation (4) and all individual-specific coefficients can then in principle be estimated by a standard probit regression for each individual separately.

A more restricted way of modelling observable heterogeneity is to use segmentation, which is used in equation (5) by interacting the individuals' observable personal characteristics with the saved number of people of a certain age such as

$$W_i = \hat{W} + \alpha^1 s^1 + ... + \alpha^n s^n + \sum_{j=1}^{n} \gamma^j x_i s^j + \varepsilon_i \ , \quad (5)$$

where γ is a vector of coefficients to be estimated and x_i denotes the observable characteristics of individual *i*, or people in subpopulation *i*, which may for example reflect age, level of education and whether the individual has any children or not. Thus we can then estimate the probability that an individual *i*, or people of the sub-population *i*, will choose *A* as follows:

$$\Pr(A \text{ is chosen}) = \Pr(W_i(A) > W_i(B)) = \Pr\left(\alpha^{\,1}\Delta s^1 + ... + \alpha^{\,n}\Delta s^n + \sum_{j=1}^{n}\gamma^j x_i \Delta s^j > \phi_i\right). \quad (6)$$

The relative value of saving a life belonging to group *j* compared to group *k* can then be calculated as:

$$SMRS_i^{\,jk} = \frac{\alpha^j + \gamma^j x_i}{\alpha^k + \gamma^k x_i} \ . \quad (7)$$

If we disregard the observable characteristics of the respondents, an SMRS equal to 1 would for example mean that only the number of lives matter. That is, the more lives saved the better, irrespective of the ages of the saved individuals.

An alternative and somewhat more restricted model is when the ISWF depends on saved lives (irrespective of ages) and the total amount of life-years saved. Here we assume that a saved person will attain the corresponding life expectancy of his/her age. Compared to the model in equation (1), in this case we impose a specific functional form for ISWF, where we assume that all individuals have the same preferences as measured in the deterministic part, and heterogeneity in preferences is captured by the error term:

$$W_i = \hat{W} + \beta l + \delta y + \varepsilon_i \ , \qquad (8)$$

where l is the total number of lives saved and y is the total number of life-years saved. This model facilitates a direct test between the "only-lives-matter-hypothesis," l-hypothesis for short, and the "only-life-years-matter-hypothesis," y-hypothesis for short. According to the l-hypothesis, we would expect that $\beta > 0,\ \delta = 0$, whereas the y-hypothesis implies that $\beta = 0, \delta > 0$. Some intermediate alternatives, of course, also exist, where $\beta > 0$ and $\delta > 0$. We can estimate the probability that individual i, or people of sub-population i, will choose program A as follows

$$\Pr(A \text{ is chosen}) = \Pr(W_i(A) > W_i(B)) = \Pr(\beta \Delta l + \delta \Delta y > \phi_i) \ , \qquad (9)$$

where $\Delta l = l(A) - l(B),\quad \Delta y = y(A) - y(B)$ and $\phi_i = \varepsilon_i(\mathrm{A}) - \varepsilon_i(B)$.

The estimated relative value of saving a life belonging to a certain age group, compared to another, is then given as follows:

$$SMRS_i{}^{jk} = \frac{\beta + \delta y^j}{\beta + \delta y^k}, \qquad (10)$$

where y^j and y^k denote the total remaining expected life-years of the saved individual belonging to age *j* and *k*, respectively.

3. The choice experiment

In a choice experiment framework, respondents make repeated choices between different alternative goods or projects described by their attributes (see Louviere et al., 2000; Alpizar et al., 2003). Some of the advantages of using a choice experiment rather than a single question experiment are that it is easier to estimate the marginal impact of different attributes on the decision and that more information is provided per respondent. However, there are negative aspects such as being cognitively more demanding for the respondents, and that the complexity of task can affect the decision of the respondents (Adamwicz et al., 1998). The complexity of the task is determined by the following factors: the number of choice sets, the number of alternatives in the choice sets, the number of attributes describing the alternatives, and the correlation between attributes (Swait and Adamwicz, 1996).

The design of a choice experiment necessarily involves defining attributes and levels of attributes, experimental design, questionnaire development, and designing a sample and sampling strategy. The attributes in our case of life-saving programs are the age of the life saved individuals and the number of lives saved. The experimental design involves creating the choice sets in an efficient way by combining attribute levels into alternatives in the choice sets. One important issue here is to minimize task complexity and obtain a manageable number of choice sets. We created the choice sets in SAS, only

considering main effects (i.e. considering the effects of each attribute on utility, by using a D-optimal design approach). D-optimal design considers the importance of the levels of the attributes in the choice sets and ensures that the alternatives give more information about the trade-off between the different attributes (see Carlsson and Martinsson, 2003). Forty-two choice sets were created, and blocked into 7 groups so that respondents answered different sub-sets of the main design. These blocks were randomly distributed among the respondents.

To facilitate the design (choice of attributes, attribute levels, and the choice scenario), we conducted focus groups and two pilot studies in order to test the choice experiment in the field. The attributes and the levels used in the final choice experiment are presented in Table 1.

Table 1. Attributes and levels used in the choice experiment.

Attribute	Levels
Number of people saved	200, 400, 700, 1000, 1300 and 1700
Age group of people saved	0 - 1, 1 - 10, 10 - 20, 20 - 30, 20 - 40, 40 - 60 and 60 - 80

The scenario used in the choice experiment (see Appendix 2) explains that the financial constraints often necessitate setting priorities in conducting life-saving programs, and people's preferences regarding such priorities are of essential importance for policy makers. Moreover, the respondents are told that it is possible to target people within certain age groups for these life-saving programs. To be more realistic, we presented saved lives groups spanning a range of ages rather than using specific ages. The age intervals used are: 0-1 year, 1-10 years, 10-20 years, 20-30 years, 20-40 years, 40-60 years and 60-80 years. Respondents were presented with six different pairs of life-saving programs differing with respect to the number of lives saved and the age

group of the saved persons; however, the programs were similar in other aspects, including the costs. Since we are genuinely interested in the attributes mentioned, it is important to mention that other aspects of the programs, including the costs, are the same. We specifically mentioned that the life-saving programs do not change the total amount of suffering of ill or injured people, so that the respondents would not assume different diseases/injuries (and their associated sufferings) for different age-groups. Finally the respondents were asked to choose their preferred alternative in each of the six choice sets assuming that each choice set was the same in all aspects, including cost, except for the attributes included.

The choice experiment was conducted among a random sample of rural households in the following five districts of Bangladesh: Netrokona, Mymensingh, Manikganj, Gazipur, and Narayanganj. The sample is therefore not representative of the Bangladesh population, which consists of 64 districts. Moreover, the villages were chosen so that the respondents of the Hindu religion are over-represented (34% compared to the national average of 11 %) in order to facilitate religious comparisons. The enumerators, used to conduct the survey, were allocated to different parts of the selected villages and were then asked to perform a survey and the choice experiment. The survey included questions on respondent health and risk perceptions, in addition to the socioeconomic questions, plus contingent valuation questions on risk reduction. The interviews were conducted with the household heads as the decisions made within households are normally made, or at least approved of, by the household heads. If the household head was not around, the enumerators were instructed to go back later. If the respondent was not at home at the second visit, the enumerator moved to the next

household.[5] The participation rate of household heads who were approached for interviews is 99 percent. The respondents were paid (100 Taka) as an appreciation of their time and cooperation. We present the sample characteristics in Table 2.

Table 2. Sample Characteristics

Variable/Characteristics	Mean	Minimum	Maximum
Male	0.91	0	1
Age	44.28	19	87
Illiterate (Cannot read and write)	0.32	0	1
Low education (Not illiterate and/or educated up to high school level)	0.54	0	1
High education (Educated above High school level)	0.14	0	1
Hindu religion [a]	0.33	0	1
Muslim religion	0.67	0	1
Per capita household income [a] (Taka)	22552	947	288848
Children in the household (up to age 16 years old)	0.87	0	1
Parent alive (either father or mother is alive)	0.09	0	1

[a] Yearly household income is adjusted with household equivalent and economies of scale, by dividing total yearly income by [Number of adults+0.5*number of children] $^{0.75}$. USD $ = 57.8 Taka, at the time of survey (October, 2003).

The enumerators were trained beforehand regarding the choice experiment, i.e. the purpose of the experiment and how to conduct it practically, the presentation of choice scenario, and the choice sets. The choice scenario was translated back to English from Bengali to ensure the exact meaning of the original English version. Each choice set was presented using a small card on which enumerators pointed each time they explained a new choice situation to the respondents. The respondents were allowed to

[5] However, in the villages, people from the same family-chain normally live in a cluster of 4-5 households. Thus, a replacement from the next household or next to the next household (in some cases) should not bias the results. There were 22 % replacement households in our sample.

mark their choices in the appropriate place on the questionnaire. After the choice experiment, we asked follow-up questions regarding priority setting on saving lives.

4. Econometric analysis and results

As our experiment involves only two choice alternatives we employ a binary probit to estimate the choice parameters. As noted earlier, each respondent made six choices, which therefore provides six observations for analysis. Thus the observations may be correlated at the individual level. To account for a possible overestimation of the statistical significance of the attributes, we use clustering at the individual level. That is, we allow observations to be independent across individuals (clusters), but not necessarily independent between responses for a specific individual (for clustering command see STATA Manual, 2003).

We first discuss the results from the general model. However, it should be noted, when discussing the results from the models, we do not present the results directly from the probit estimates; instead we present and discuss the SMRS (relative value) that we obtain using those coefficients. The probit results are available from the author on request. First, however, we present calculations of the ratio of life expectancy for the relevant ages in Table 3 (column 2). The ratio of life expectancy is calculated by dividing the average remaining life expectancy for a certain age by the average remaining life expectancy for a 1-year old. Using the estimated probit coefficients and equation 7 (disregarding the respondent characteristics), we calculate the relative value (*SMRS*) between age groups in the experiment, using saved 1-year olds as the base case. These relative values are reported in column 3 of Table 3. The results indicate that the relative value strongly decreases with age. For example, saving one 1-year old is judged

equivalent to saving 2.09 (1/0.478) 30-year olds. Using the Delta method (Greene, 2000), we obtain the standard errors and hence we can calculate the 95% confidence interval for the ratios. As shown in Table 3, except for the 50-year old saved, the *SMRSs* are significantly different from one at 1% level. The relative values of life obtained here are roughly similar to the ratios of life expectancy, up to a certain age. However, the results also indicate that relative value is negative and significant for the oldest group (70-years old) in the choice sets. While such a negative value is very surprising and theoretically un-intuitive, this might result from the fact that respondents have a very strong preference for saving more life- years by saving younger lives; therefore in a pair wise choice of life-saving programs they might be always choosing the alternative which saves relatively younger lives. Thus, this might result in negative coefficients even if their true preferences entail assigning very small weight on saving the old.

Given the results discussed above, we are particularly interested in the preference heterogeneity of saving the lives of older people. Therefore, we model observed preference heterogeneity (using equation 6) by interacting 70-year olds saved with the following characteristics of the respondents: age, education, religion, income, and whether a respondent's parent is alive or not.[6] From the estimated probit coefficients (not reported here), we calculate the relative value for all saved lives as well as an additional value for the 70-year olds saved in the choice experiment using equation (7), using saved 1-year old as the base case (fourth column of Table 3). We do not find any significant differences in values towards older saved lives with regards to respondent age and income. It may be mentioned that both Cropper et al. (1994) and Johannesson and Johansson (1997) found that respondent preferences are unaffected by

[6] As our sample consists of only 9 percent female respondents, we do not focus on gender difference.

own age, while Johansson-Stenman and Martinsson (2004) found that older respondents value older saved lives higher than younger respondents do. However, unlike Johansson-Stenman and Martinsson, we use age as a continuous variable. We find that a respondent values older saved lives relatively higher if his/her parent is alive. This seems quite reasonable given that people may care more for their own elderly parents than for the elderly in general.[7] On the other hand, the respondents with a higher level of education and the respondents of the Hindu religion on average assigned relatively lower values to 70- year olds. We do not have a good explanation for this result. We have also investigated observed preference heterogeneity of saving younger people, especially saving children, without finding any significant results to discuss here.

Then we estimate the parameters of the restricted model (equation 9). Contrary to the hypotheses mentioned earlier, we obtain a negative coefficient for the number of lives saved and a positive coefficient for the amount of life- year saved, and both are significant at the 0.1 % level. Hence, we obtain the relative value of life using equation (10) with 1-year olds saved as the base case. The results are reported in the fifth column of Table 3. The results are almost similar to the results reported in the third column of Table 3 in that we obtain a negative value for the oldest group in the choice experiment.

[7] There is a negative and significant (at the 1% level) correlation between the age of respondents and whether their parents are alive (r=-0.14).

Table 3. Relative value of (statistical) life [a] (SMRS)

Saved life	Life expectancy ratio [b]	SMRS [c]	SMRS [d]	SMRS [e]
1- year old saved	1	1	1	1
5- year olds saved	0.977	0.790*** (0.594 – 0.986)	0.819*** (0.634 – 1.1)	0.966*** (0.962 – 0.969)
15 -year olds saved	0.84	0.569*** (0.303 – 0.836)	0.602*** (0.349 – 0.856)	0.761*** (0.737 – 0.785)
25- year olds saved	0.698	0.490*** (0.194 – 0.786)	0.515*** (0.236 – 0.799)	0.548*** (0.502 – 0.593)
30- year olds saved	0.628	0.478*** (0.332 – 0.624)	0.494*** (0.355 – 0.634)	0.444*** (0.388 – 0.50)
50- year olds saved	0.363	0.003 (-0.154 – 0.160)	0.035 (-0.114 – 0.184)	0.047 (-0.048 – 0.143)
70- year olds saved	0.158	-0.397*** (-0.587 – -0.188)	-0.481** (-0.895 – -0.066)	-0.259*** (-0.385 – -0.132)
Observed heterogeneity				
70- year olds saved × respondent's age			0.004 (-0.003 – 0.012)	
70- year olds saved × respondent has high education			-0.374** (-0.717 – -0.074)	
70- year olds × Per-capita household income			0.0003 (-0.0009 – 0.002)	
70- year olds saved × respondent is of Hindu religion			-0.261** (-0.498 – -0.054)	
70- year olds saved × respondent's parent alive			0.229** (0.009 – 0.449)	

[a] 1- year old saved is the base case. 95% confidence intervals in the parentheses and the superscripts *** and ** denote statistical significance at 1% and 5% level, respectively.

[b] Life expectancy figures are adapted from the BBS (2001).

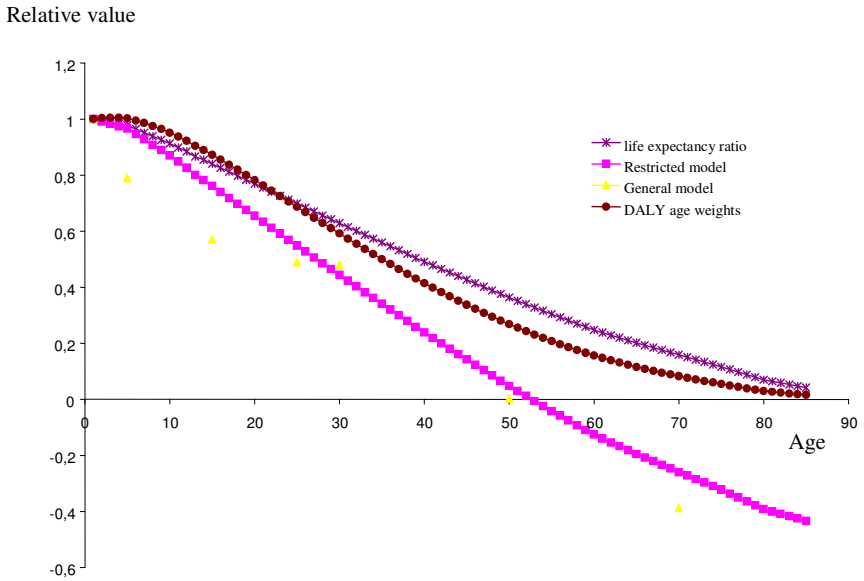
[c] Calculated using the estimated coefficients of the general model.

[d] Calculated using the estimated coefficients of the general model accounting for heterogeneity.

[e] Calculated using the estimated coefficients of the restricted model.

In Figure 1, we plot the relative value of a life obtained from the restricted model, from the general model, and from the DALY age weights, as a function of age, along with the ratio of life expectancy.

Figure 1. Relative value of life



As depicted in Figure 1, the relative value of life becomes negative around the age of 53 with both models used here. It is interesting to note that our results are quite closely related to the age pattern of the relative value of life obtained using the DALY age weighting function, up to a certain age. The fact that we obtained negative values for the oldest age group in both models may seem unreasonable. One can speculate about the reason. One possibility is that there are methodological reasons related to the choice experiment design, and that the results simply reflect that the respondents have very strong preferences for saving younger individuals, i.e. almost lexicographic

preferences. This may also reflect an expression of attitude, which is not related to trade-offs in the choice experiment or preferences of the individual (Kahneman et al., 1999). It may also be mentioned that we obtained a similar negative value using a slightly different design in two pilot studies that we conducted before the final choice experiment. On the other hand, it might be that respondents simply do not want to prolong an older life considering their apparent old age related sufferings out of economic hardship, socio-economic insecurity, senile diseases, and poor health care facilities. In the absence of modern social safety nets for the elderly and in view of changing socio-economic transformations and poverty, "the aged persons in an average Bangladeshi family are often treated as a burden" (Banglapedia, 2003), where almost 80% of the elderly live in rural areas. The result we obtain may be a manifestation of changing cultural norms, values and attitudes.

It appears unlikely, however, that we would have obtained a similar age pattern when focusing on health rather than on life-saving. In a recent study in Canada, it is shown that the intervention type seems to matter for age preference; responses for a life-saving scenario favored younger age groups while those for palliative care scenarios showed no age preference (Johri et al., forthcoming).

Moreover, we also estimated a further restricted model, where the value is a function of the number of saved life-years up to a certain age, and that saving life-years beyond this cut-off point yields no further value. In other words, saving lives above this age level yields zero value. Using a grid search procedure (see e.g. Greene, 2000), we estimated such a model (not reported here but available upon request), which resulted in an estimated cut-off point of about 66 years. In other words, beyond the age of 66 additional saved life-years have a social value of zero. Otherwise, the age-pattern was

quite similar to the previously estimated model without any cut-off point. On the whole, our results clearly show the importance of the number of life-years saved in the valuation of life, i.e. they support the *y*-hypothesis rather than the *l*-hypothesis, which is also consistent with the ratio of life expectancy and with the results implied by the DALY age weighting function used in World Bank (1993).

5. Follow-up questions

After the choice experiment, we asked follow-up questions regarding priority setting on saving lives. As noted earlier, World Bank (1993) observes that most societies in practice seem to attach higher values to a year of life of young and middle aged adults than to a year of life of a child or an elderly person, which also implies, without any discounting, that the relative value of life decreases with age. We therefore explicitly asked the respondents if they wanted to prioritize saving the younger people over the older. Then we asked about the respondents' preferences on prioritizing the younger, given that they agreed to prioritize the younger. It might be relevant to think that it is fair to save younger individuals so that they can live, all else going well, as many years as an old person has already lived. Moreover, as an older person has fewer expected life years left, by saving a considerably younger person, more life-years could be saved to achieve more societal welfare. Another important aspect is that young adults could contribute to the society both in terms of production and child rearing, and could also shoulder the responsibility of the older people. This aspect is related to productivity ageism. The exact wordings of follow-up questions and responses are presented in Table 4.

Table 4. Preferences for priority setting in saving lives

Question/Statement	% of respondents agree
1. Society should give a higher priority to saving younger people. [a]	98
2. Why do you think that society should give a higher priority to saving younger people?	
a) A younger individual has a longer time left to live. Hence the society saves more life-years by saving a younger individual compared to an older.	80
b) It is fairer that younger individuals are saved since they have not lived as many years as older individuals have.	66
c) It is better from a social point of view to save younger individuals since they will contribute more to the society in terms of production and of raising children.	98

[a] Given an agreement to this statement, question 2 is asked; the respondents could choose more than one response.

As elicited in the follow-up questions, almost all respondents agree with the statement that society should give higher priority to saving younger people. Out of these, about 80% support the view that society could save more life-years by saving a younger individual compared to an older, and 66% think that younger individuals should be given priority as they have not lived as many years as older individuals have (supporting "fair-innings ageism"). Almost all respondents support the view that society should give a higher priority to saving younger people primarily because they will contribute to the society in terms of production and raising children (supporting "productivity ageism").

Although we cannot rule out the possibility of "yea saying" in these responses, the findings are nevertheless consistent with our estimated choice experiment results. As discussed earlier, in a developing country context, prime age adults shoulder the responsibility of their children as well as elderly members in the family, where the absence of modern social safety nets makes the elderly clearly dependent on family or relatives. It may be mentioned that old age security provides potentially strong motives

for having more children (high birth rates) in poor countries as children are also useful as income earning assets (Dasgupta, 1993). A saved child has more years to live, all else going well, and therefore would contribute to the family as well as to society in terms of production and child rearing (Tsuchiya, 1999).

6. Conclusion

Our results strongly indicate that the respondents have preferences for life-saving projects that save younger people, and a model where people simply value the total number of saved life-years appears to explain our data quite well. Interestingly, our results are quite closely related to the age pattern of relative values, up to a certain age, implied by the disability adjusted life year (DALY) age weights used by the World Bank to reflect different social values of different ages of people. While the DALY age weights relate to "productivity ageism," the values we obtain, as a reflection of peoples' preferences, appear also to be related to other kinds of ageism. One may argue that decisions on priority setting should be informed by people's preferences on such priorities. Our results are informative for priority setting in public health projects in the context of developing countries. The rather surprising results that we obtain regarding the relative value of "old years" may partly be linked to measurement problems related to the choice experimental design and that respondents are expressing a view not related to the trade-offs in the choice experiment presented to them. It is essential for more research to be done on eliciting preference for relative value of life using the choice experiment method. In any case, we do certainly not recommend any negative weights to be used in practice. Moreover, the choice experiment approach on life-saving programs may seem rather demanding in the context of a developing country, where a

substantial proportion of the population has no formal education. Future research should also address such methodological issues when eliciting peoples' preference for using relative value of life in public decision making.

References

Adamwicz, V., Boxall, P., Williams, M, and Louviere, J. (1998), Stated preference approaches to measuring passive use values, American Journal of Agricultural Economics 80, 64-75

Alpizar, F., Carlsson, F. and Martinsson, P. (2003), Using Choice Experiments for Non-Market Valuation, Economic Issues, 8:83-110.

Banglapedia (2003), National Encyclopaedia of Bangladesh, Vol. 7. Asiatic Society of Bangladesh, Dhaka.

BBS (2001), Bangladesh Bureau of Statistics, Statistical Yearbook of Bangladesh, Statistics Division, Ministry of Planning, Government of Bangladesh, Dhaka.

Broome, J. (1999), Ethics out of Economics, Cambridge University Press.

Carlsson, F. and Martinsson, P. (2003), Design Techniques for Stated Preference Methods in Health Economics, Health Economics, 12:281-294.

Cropper, M. L., Aydede, S.K., Portney, P.R. (1994) Preferences for Life-saving Programs: How the Public Discounts Time and Age, Journal of Risk and Uncertainty 8, 243-265.

Greene, W.H. (2000), Econometric Analysis, Fourth Edition, Prentice Hall, New Jersey.

Cookson, R. and Dolan, P. (2000), Principles of justice in health care rationing, Journal of Medical Ethics, 26:323-329.

Dolan, P. (2000), The Measurement of Health Related Quality of Life, in: Culyer, A.J. and J.P. New house, (eds), Handbook of Health Economics, Elviser, Harvard University, Cambridge, MA, USA.

Dasgupta, P. (1993), An Inquiry into Well-Being and Destitution, Oxford University Press.

Hammitt, J. K. and Graham, J.D. (1999). Willingness to pay for health protection: Inadequate Sensitivity to Probability? Journal of Risk and Uncertainty 18, 33-62.

Hammitt, J.K. (2002), QALYs Versus WTP, Risk Analysis, 22(5):985-1001.

Hausman, D.M. and McPherson, M.S. (1993), Taking Ethics Seriously: economics and contemporary modern philosophy, Journal of Economics literature, 31(2):671-731.

Johannesson, M and Johansson, P.O. (1997) Is the Valuation of a QALY gained independent of age? Some empirical evidence, Journal of Health Economics 16, 589-599.

Johansson-Stenman, O. and Martinsson, P. (2004), Are Some Lives More Valuable? Working Paper, Department of Economics, Göteborg University.

Johri, M., Damschroder, L.J., Zikmund-Fisher, B.J., and Ubel, P. A. (forthcoming), The importance of age in allocating health care resources: does intervention-type matter? Health Economics.

Kahneman, D., Ritov, I., and Schkade, D. (1999), Economic Preferences or Attitude Expressions? An Analysis of Dollar Responses to Public Issues, Journal of Risk and Uncertainty, 19(1): 203-235.

Louviere, J.J., Hensher, D.A., Swait, J.D (2000), Stated Choice Methods, Analysis and Application, Cambridge University Press.

McFadden, D. (1974), Conditional Logit Analysis of Qualitative Choice Behavior, in P. Zarembka (ed.), Frontiers in Econometrics, New York: Academy Press.

Mooney, G. and Wiseman, V. (2000), Burden of Disease and Priority Setting, Health Economics, 9: 369-372.

Murray, C.J.L. (1994), Quantifying the burden of disease: the technical bases for disability adjusted life years, In: Murray, C.J.L., and Lopez, A.D. (eds.), Global Comparative Assessments in the Health Sector: Disease Burden, Expenditures and Intervention Packages, World Health Organization, Geneva.

Murray, C.J.L. (1996), Rethinking DALYs, In: Murray, C.J.L., and Lopez, A.D. (eds.), The Global Burden of Diseases, Harvard University Press.

Murray, C.J.L., and Lopez, A.D. (1994.), Global Comparative Assessments in the Health Sector: Disease Burden, Expenditures and Intervention Packages, World Health Organization, Geneva.

Murray, C.J.L and Acharya, A.K. (1997), Understanding DALYs, Journal of Health Economics, 16: 703-730.

Murray, C. J.L. and Lopez, A. D. (2000), Progress and Directions in Refining the Global Burden of Disease Approach: A Response to Williams, Health Economics, 9:69-82.

STATA Manual (2003), Cluster Analysis, Stata Manual, Stata Coprporation, Texas, USA.

Swait, J. and Adamowicz, (1996), The effect of choice environment and task demands on consumer behavior: discriminating between contribution and confusion, Working Paper, Department of Rural Economy, University of Alberta.

Tsuchiya, A. (1999), Age-related preferences and age weighting health benefits, Social Science and Medicine, 48:267-276.

Tsuchiya, A. (2000), QALYs and Ageism: Philosophical Theories and Age Weighting, Health Economics, 9:75-68.

Tsuchiya, A., Dolan, P., Shaw, R (2003), Measuring people's preferences regarding ageism in health: some methodological issues and some fresh evidence, Social Science and Medicine, 57: 687-696.

Viscusi, W. K. (1993), The Value of Risks to Life and Health, Journal of Economic Literature 31, 1912-1946.

Viscusi, W. K., and Aldy, J.E. (2003), The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World, Journal of Risk and Uncertainty, 27(1), 5-76.

Weinstein, M.C., Shepard, D.S., and Pliskin, J.S. (1980), The Economic Value of Changing Mortality Probabilities: A Decision Theoretic Approach, Quarterly Journal of Economics, 94(2), 373-396.

Weinstein, M.C. and Stason, W.B. (1997), Foundations of Cost-effectiveness Analysis for Health and Medical Practices, New England Journal of Medicine 296:716

Williams, A. (1997), Intergenerational Equity: An Exploration of the 'Fair Innings' Argument, Health Economics, 6: 117-132.

Williams, A. (1999), Calculating the Global Burden of Disease: Time for a Strategic Reappraisal? , Health Economics, 8: 1-8.

Williams, A. (2000), Comments on the response by Murray and Lopez, Health Economics, 9: 83-86.

Williams, A. and Cookson, R. (2000), Equity in Health, in: Culyer, A.J. and J.P. Newhouse, eds., Handbook of Health Economics, Elviser, Harvard University, Cambridge, MA, USA.

World Bank (1993), World Development Report 1993: Investing in Health, London, Oxford University Press.

World Health Report (2004), The World Health Report 2004 -changing history, The World Health Organisation, Geneva.

Appendix 1: Relative value of life using the DALY age weighting function

The relative value of a year of life at each age has been modelled using the following formula:

Age-weighting function = $Kxe^{-x\beta}$, (1)

where $K = \text{constant} = 1$

$\beta = \text{constant} = 0.04$

$x = \text{Age}$

This function rises quickly from zero at birth to a peak at age 25 and then declines asymptotically toward zero (World Development Report, 1993, p 213).

We calculate the value of life at each age, x, by integrating equation (1) as follows:

$$\int_{x}^{T(x)} Kxe^{-x\beta}dx, \quad (2)$$

$$= \left[-\frac{k}{\beta^2}\{x\beta + 1\}e^{-\beta x} \right]_{x}^{T(x)}$$, where $T(x)$ denotes life expectancy at age x.

The relative value of life at each age is obtained as a ratio by taking the value for 1- year old as the base case.

Appendix 2: Choice scenario

Governmental policy makers can prevent or postpone many deaths by increasing the financial resources for different kinds of life-saving programs. However, since the government's budget is limited, it has to choose which programs to prioritize. The purpose of this part of the survey is to gather information about people's preferences for such priorities.

Appropriate life-saving programs can prevent many causes of death. Many people die each year because of contaminated water, contaminated food, polluted air, smoking, or as a result of road accidents. More and better life-saving programs could reduce the number of deaths from each of these causes.

Suppose that there are two different life-saving programs and that it is possible to target different age groups of the population within each of these programs. Both programs save a different number of lives in different age groups. Both programs cost the same.

[Enumerator: Show Figure EXAMPLE and POINT to the attribute levels when they are mentioned]

As an example, assume that you were to choose between two available life-saving programs, A and B. The effects of the programs differ with respect to the number of lives saved and the age of those saved. The cost of both life-saving programs is the same. Program A saves 200 lives of people who are 20-40 years old, and program B saves 250 lives of people who are 40-60 years old. The programs do not change the total amount of suffering experienced by ill or injured people.

	Program A	Program B
Age-group of lives saved	20-40 years	40-60 years
Number of lives saved	200	250
Your choice		

QUESTION: If both programs cost the same, which life-saving program would you choose?

[Enumerator: Let the respondent put cross mark (X) in the blank box under the chosen program]

We will now present you with 6 different pairs of life-saving programs and each time we will ask you to choose the one you think would be the best for society. The effects of the programs differ with respect to the number of lives saved and their age-group, but they are similar in all other respects. The programs do not change the total amount of suffering experienced by ill or injured people.