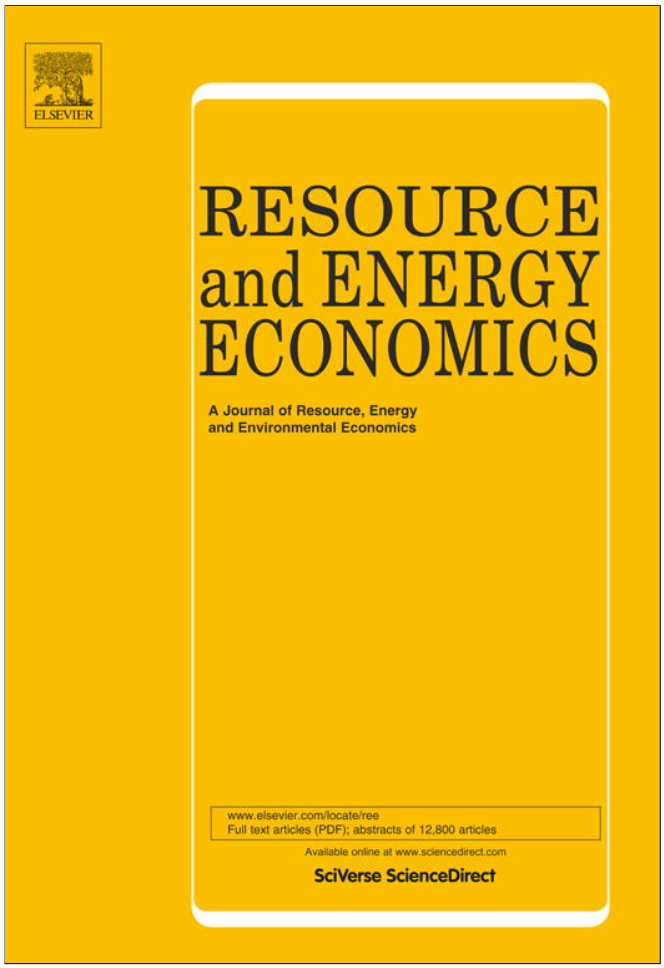

(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

Resource and Energy Economics 34 (2012) 493–513

Contents lists available at SciVerse ScienceDirect

Resource and Energy Economics

journal homepage: www.elsevier.com/locate/ree

The willingness to pay for mortality risk reductions in Mongolia☆

Sandra Hoffmann[a,1], Ping Qin[b,*], Alan Krupnick[a], Burmaajav Badrakh[c], Suvd Batbaatar[d], Enkhjargal Altangerel[e], Lodoysamba Sereeter[f]

[a] Resources for the Future, 1616 P St. NW, Washington, DC 20036, USA
[b] School of Economics, Renmin University of China, Beijing, China
[c] Mongolian Ministry of Health, Ulaanbaatar, Mongolia
[d] Public Health Institute, Ulaanbaatar, Mongolia
[e] Mongolian Public Health Association, Ulaanbaatar, Mongolia
[f] National University of Mongolia, Ulaanbaatar, Mongolia

ARTICLE INFO

Article history:
Received 15 February 2011
Received in revised form 19 December 2011
Accepted 30 April 2012
Available online xxx

JEL classification:
H41
I18
O13
Q51
Q53
Q56

Keywords:
Value of statistical life
Air pollution
Mortality risk
Environment and development
Public health

ABSTRACT

This paper reports results from a stated preference survey designed to estimate the willingness to pay for mortality risk reductions in Ulaanbaatar, Mongolia. The survey includes both contemporaneous and latent risk reductions of a magnitude typically achievable through clean air policy. The study is one of a series of national studies designed to provide comparable estimates around the world. One goal of this series is to build a more solid bridge for benefits transfer between developed and developing countries. The survey was conducted in winter 2010. Estimates of willingness to pay passed external and internal scope tests. Study results imply a value of statistical life of approximately $500,000 (based on a purchasing power parity exchange rate) for a contemporaneous 5-in-10,000 annual risk reduction.

☆ Views expressed in this work are the author's own and may not be attributed to the U.S. Department of Agriculture.

* Corresponding author.
E-mail addresses: shoffmann@ers.usda.gov (S. Hoffmann), pingqin2009@gmail.com (P. Qin), krupnick@rff.org (A. Krupnick), burmaajav@yahoo.com (B. Badrakh), suvd552001@yahoo.com (S. Batbaatar), ajargal8@gmail.com (E. Altangerel).
[1] Present address: Food Economics Division, U.S.D.A. Economic Research Service, 355 E St. SW, Washington, DC 20024, USA.

0928-7655/$ – see front matter
http://dx.doi.org/10.1016/j.reseneeco.2012.04.005

1. Introduction

Developing countries face the difficult problem of setting priorities among many pressing needs for both public and private capital. With heavy demands for improved infrastructure, education, health care, nutrition, and economic development, governments often view environmental quality as a luxury. Yet the citizens of developing countries may feel the need for improved environmental quality as acutely, if not more acutely, than citizens in developed countries. Information on these preferences should help governments do a better job of setting priorities to benefit their citizens, whether directly through cost-benefit analysis of proposed programs or projects, or indirectly, through building a sense among policymakers of how strongly the public feels about improving their health and reducing their chances of dying.

This paper estimates the willingness to pay (WTP) of residents of Ulaanbaatar, Mongolia for reductions in mortality risk of the size and type typically provided by air pollution policy. Recent air quality monitoring in Ulaanbaatar found annual mean particulate matter levels (PM10) that are the highest yet measured in the world. Mean annual PM10 levels are estimated to be 427.5 $\mu g/m^3$, over twenty times the World Health Organization (WHO) annual guideline of 20 $\mu g/m^3$ (World Bank, 2011). More polluted ger neighborhoods have annual average PM10 levels of 700 $\mu g/m^3$. Even the cleaner, central city neighborhoods are estimated to have mean annual PM10 levels of 300 $\mu g/m^3$. Winter mean daily PM10 levels in Ulaanbaatar were measured to be as high as 3800 $\mu g/m^3$ and frequently exceeded 500 $\mu g/m^3$ (World Bank, 2011). WHO guidelines recommend that PM10 levels not exceed a 24-h mean of 50 $\mu g/m^3$ (WHO, 2006). The WHO recognizes that national standards will need to vary with economic conditions in each country, but even Mongolian PM10 standards are set at an annual mean of 50 $\mu g/m^3$ and a daily mean of 100 $\mu g/m^3$ (WHO, 2006; World Bank, 2011).

The WTP estimates presented here are based on results from a stated preference survey of Ulaanbaatar residents, conducted in January and February 2010. In addition to providing Mongolia-specific value of statistical life (VSL) estimates, this survey is designed to help build a stronger empirical basis for transfer of environmental health valuation estimates among countries, particularly between developed countries, where many such studies have been conducted, and developing countries, where there have been far fewer studies. It does this by administering the same survey instrument used in studies in seven other countries: the United States and Canada (see Alberini et al., 2004; Krupnick et al., 2002), the United Kingdom, France, and Italy (see Alberini et al., 2006b), Japan (see Itaoka et al., 2005) and China (Krupnick et al., 2010). As in the China study, this computerized survey uses a payment card/screen rather than the dichotomous choice approach to eliciting respondents' WTP that was used in the other six studies. This approach was used to address problems with dichotomous choice elicitation found in recent stated preference studies conducted in China (Krupnick et al., 2010; Xu et al., 2003) and in India (Bhattacharya et al., 2007). The risk reductions offered respondents in the Mongolian survey are of the kind and size that are expected from air pollution policy (contemporaneous (e.g., respiratory) and latent (cancer)), ranging from a 5- to 10-in-10,000 change in annual risk reductions over a 10-year period.

This paper is organized as follows. Section 2 reviews contingent valuation studies of WTP for mortality risk reductions in eastern Asia, including one conducted in Inner Mongolia, and methodologically focused literature comparing dichotomous choice and payment card elicitation methods. Section 3 briefly describes the theoretical model underlying statistical analysis in this study. Section 4 describes the survey design, adaptation, and administration. Section 5 describes the characteristics of our sample, estimation methods and results, including estimates of WTP for current and future risk reductions. Section 6 is a discussion section, which recommends a VSL for use in policy analysis and compares our estimates with those from other countries using the same basic survey instrument. Finally, Section 7 offers conclusions.

2. Literature review

We are aware of no prior studies estimating willingness to pay for reduction in mortality risks in Mongolia. We found six such studies in China (Hammitt and Zhou, 2006; Krupnick et al., 2010; Li et al., 2002; Wang et al., 2001; Zhang, 2002; Zhang et al., 2004). Differences between cultural and

socioeconomic conditions in Mongolia and China argue for the usefulness of a specifically Mongolian VSL estimate rather than use of estimates based on those from China.

This study uses a payment card to elicit willingness to pay. Payment cards are one of the two standard approaches to eliciting willingness to pay using closed ended questions. The other is dichotomous choice. With payment cards, respondents are asked to choose the amount they are willing to pay for a benefit from values written on a card. In dichotomous choice questions respondents are asked whether they are willing to pay a specified amount and then may be asked a follow-up dichotomous choice question contingent on their answer to the first question. Champ and Bishop (2006) review the relative merits of each approach and conclude that, independent of the study context, neither method is clearly superior to the other. Krupnick et al. (2010) provides a more detailed discussion of the literature comparing payment card and dichotomous choice elicitation methods and their application in China.

Prior stated preference studies in China and India, including one in Inner Mongolia, found respondents tended to agree to pay the amount presented in dichotomous choice questions even when those values were raised to very high levels (Bhattacharya et al., 2007; Krupnick et al., 2010; Xu et al., 2003). This tendency toward "yea saying" has been seen in other settings and may be a common form of survey bias in which respondents seek to please the surveyors or act from other motives. These Asian studies found that payment cards provided a reasonably effective means of overcoming this problem. In a study in Chinese cities that used the instrument on which this Mongolian survey is based, Krupnick et al. (2010) found excessive "yea saying" and failure to pass the external scope test with a double-bounded dichotomous choice bid elicitation format even though that survey instrument had been successfully administered in six industrialized countries. This problem was not found when the same survey was administered substituting a "payment screen" elicitation format for the dichotomous choice format (Krupnick et al., 2010). Xu et al. (2003) encountered a similar problem with excessive "yea saying" when using a dichotomous choice question to elicit WTP for riverine ecosystem services in Inner Mongolia. They also found that use of a payment card resolved this problem. Similar results were found in India (Bhattacharya et al., 2007).

3. Conceptual model

The conceptual model underlying this study, as well as the other seven studies which use this survey instrument, is a life-cycle model of consumption with uncertain lifetime explained fully in Alberini et al. (2004). This work draws on a long line of similar models dating back to the 1960s (see Arthur, 1981; Cropper and Sussman, 1990; Shepard and Zeckhauser, 1982, 1984; Yaari, 1965). The life-cycle model assumes that an individual of age j maximizes expected lifetime utility by choosing a future consumption stream:

$$V_j = \sum_{t=j}^{T} q_{j,t}(1+\rho)^{j-t} U_t(C_t) \tag{1}$$

where V_j is the present value of expected utility lifetime consumption, $U_t(C_t)$ is utility of consumption at age t, $q_{j,t}$ is the probability that the individual who is now age j survives to age t, and ρ is the subjective rate of time preference. V_j is maximized subject to the constraint that the present value of expected consumption cannot exceed current wealth W_j plus the present value of expected lifetime earnings, y_t:

$$\mathcal{L}_j = \sum_{t=j}^{T} q_{j,t}(1+\rho)^{j-t} U_t(C_t) + \lambda_j \left[\sum_{t=j}^{T} q_{j,t}(1+r)^{j-t}(y_t - C_t) + W_j \right] \tag{2}$$

where λ_j is the marginal value of an increase in wealth at time j and r is the riskless rate of interest.

Policies, like air pollution control, which reduce the probability of dying, D_t, in any period, t, will thus affect V_j. Because reducing the probability of dying increases the likelihood of surviving in that and all subsequent periods, the life cycle model can be used to determine the amount of current wealth

that an individual would be willing to pay to reduce the probability of dying in any period or periods. The willingness to pay for a small change in probability of dying in any year k can be expressed as the marginal rate of substitution between wealth, W_j, and the probability of dying in year k times the change in D_k (Alberini et al., 2006a).[2]

$$WTP_{j,k} = -\frac{dV_j/dD_k}{dV_j/dW_j} dD_k \equiv VSL_{j,k} dD_k \tag{3}$$

Or, inversely, VSL is the WTP for a small change in mortality risk divided by the change in the risk. Alberini et al. (2004, 2006a) show that under standard assumptions regarding individuals' ability to borrow through their lifetime, (2) and (3) imply that:

$$WTP_{j,k} = \frac{dD_k}{1-D_k} \sum_{t=k+1}^{T} q_{j,t} \left[\lambda_j^{-1} (1+\rho)^{j-t} \frac{u_t(C_t)}{\partial u_j/\partial C_j} + (1+r)^{j-t} (y_t - C_t) \right] \tag{4}$$

Eq. (4) serves as the basis for empirical estimation in this study, as it did in other studies conducted using the survey instrument adapted for this study in Ulaanbaatar. In particular, Eq. (4) implies that WTP is increasing in wealth and income, in reductions in the risk of death, and aversion to mortality risks. Because the literature on gender and risk aversion generally finds women more averse to physical risks than men, Eq. (4) also implies that gender should help explain WTP to reduce mortality risks (Finucane et al., 2000; Flynn et al., 1994).

Alberini et al. (2004) show that (4) implies a theoretically ambiguous relationship between age and WTP to reduce mortality risks, and between health status and WTP to reduce mortality risks. Reviews of the empirical literature on the relationship between age and WTP to reduce mortality risks show a mixed picture. Aldy and Viscusi (2007) conclude that VSL estimates based on labor market data peak in mid-life and decline with age. Krupnick (2007) concludes that evidence from stated preference studies on this relationship is mixed and sensitive to model specification. An inverted U-shape relationship between age and WTP was found in some of the other studies conducted using the survey instrument on which this study was based (see, for example, Alberini et al. (2004) for Canada and Alberini et al. (2006a) for the U.S.), though not in Japan (Itaoka et al., 2005).

The empirical literature on the relationship between health status and WTP to reduce mortality risk is scant and the results mixed (NRC, 2008). However, using the same basic survey instrument we used for Mongolia, Alberini et al. (2004) reported that a family history of chronic heart or lung disease increased WTP by 26 percent in Canada, and 37 percent in the United States. Baseline risk presented in the survey is the mortality risk of someone of the respondent's age and gender living in the same place as the respondent. This suggests that theoretical findings regarding the ambiguous relationship between health status, age, and WTP may also hold for baseline risk. Past studies generally have not found baseline risk to have an independent effect on WTP.

4. Survey design and administration

4.1. The survey instrument

The survey questionnaire begins with demographic questions. Questions about the health history of the respondent and the respondent's parents are used to remind respondents that they face mortality risks. Responses to these questions are also used to test for the influence of health status on WTP to reduce mortality risks. The second section of the survey introduces the concept of chance and the probability of dying. Simple probability concepts are introduced, using virtually universal games of chance including coin tosses (Fig. 1) and die throws. Because mortality risks associated with air pollution and other environmental hazards are typically much smaller than those seen on a die toss, a representation using a much larger denominator is needed. The survey works towards this goal by first presenting a grid of 36 squares motivated by a discussion about a wheel used to select prizes

[2] D_k is considered as the altered probability of dying at age k. By definition, $q_{j,t} = (1-D_j)(1-D_{j+1})\cdots(1-D_{t-1})$.

Fig. 1. Introduction to probability concepts using coin tosses.

on a Mongolian TV game show (similar to a roulette wheel). It then presents the risk communication device used in the survey to elicit WTP for a mortality risk reduction—a 1000-square grid in which risks are represented using red squares (Fig. 2). To test their comprehension, respondents are asked to compare grids depicting mortality risks for two hypothetical people (person A and person B) and to determine which of the two has the higher risk of death (Fig. 2). They are also asked to select which of the two people they would rather be. The results of these questions are used to identify respondents' ability to understand the survey instrument and are used in cleaning the survey data.

The third section of the survey is designed to increase the saliency of mortality risk and mortality risk reduction efforts to respondents. It first presents respondents with age- and gender-specific leading causes of death and then introduces common risk-mitigating behaviors, illustrative risk reductions, and qualitative costs of reducing risks. These screens remind respondents that people regularly take actions that have real costs as a means of reducing their chance of dying. We present respondents with information on the effectiveness of common actions taken to reduce mortality risk and describe their relative costs qualitatively.

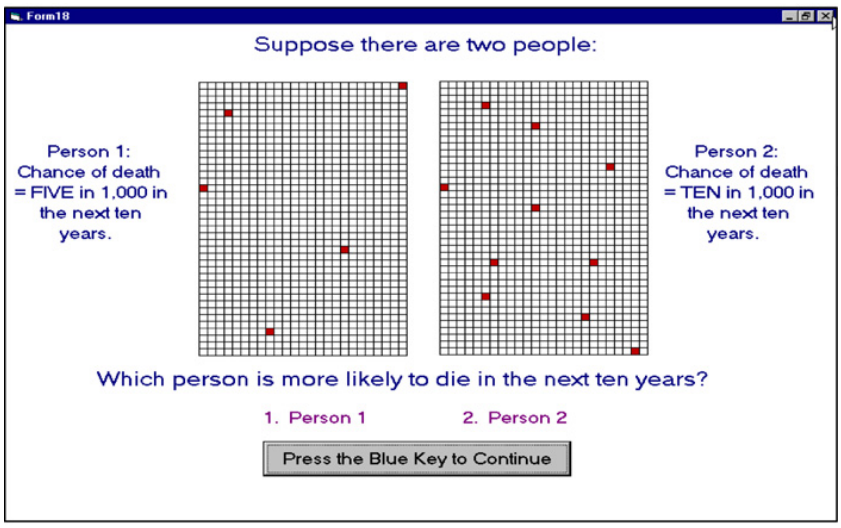

Fig. 2. The survey risk communication device: the 1000-square grid (American version used to develop Mongolian survey).

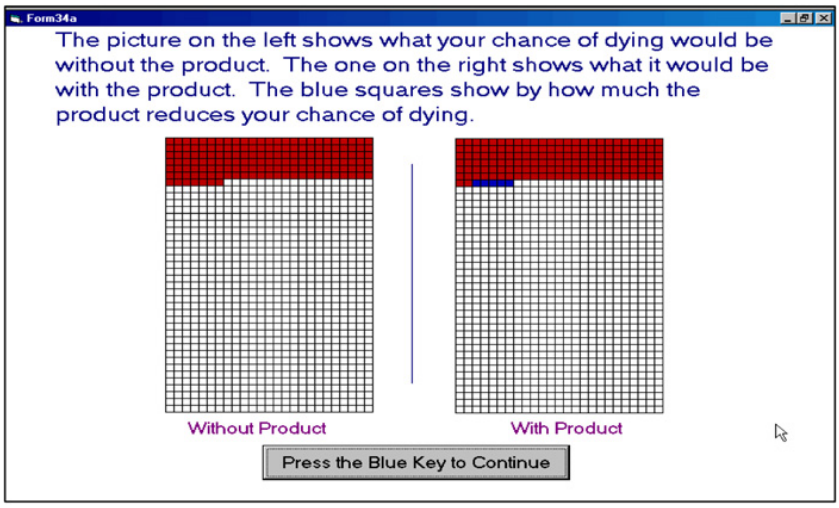

Fig. 3. Depiction of risk change (American version used to develop Mongolian survey).

The survey then presents respondents with information on baseline mortality risks for someone of their own age and gender living in Ulaanbaatar and asks them to accept this risk as their own for the purpose of the survey. Respondents' acceptance of the baseline risk is tested in debriefing questions.

Respondents are next told about a product that would reduce their risk of dying over a specified period. The description of the product is purposefully left as generic as possible so as not to have the characteristics of the product affect respondents' WTP (e.g., ease of use or side effects of a drug). We used a product as the means of reducing the risk rather than air pollution policy because the purpose of the survey is to provide a clean estimate of WTP to reduce mortality risk. Having respondents think that the risk reduction was due to an improvement in air quality would evoke a wide range of benefits in addition to mortality risk reduction. These might include the aesthetic benefits of clearer air, the altruistic utility respondents gain from protecting others' health or the benefits respondents themselves might experience from reductions in morbidity, rather than mortality, associated with reduced air pollution. It also might (as we have seen in focus groups) evoke zero bids if respondents think that they should not have to pay to clean up pollution not of their making. The 1000-square grid is used to show respondents their own baseline mortality risk and the reduction in their baseline risk that would result from use of this product (Fig. 3). Finally, respondents are asked about their WTP for this product using a payment card approach (see below).

At the time the survey was conducted it was unknown how large a risk reduction might be possible from air pollution policy in Ulaanbaatar. As a result, we considered use of mortality risk reductions of a magnitude typically seen from air pollution policies, i.e., in the range of 1 in 10,000 to 10 in 10,000 per year. As the survey was originally being developed in the United States and in other countries, it became clear that respondents had a difficult time understanding very small changes in risk such as these. As a result, a device was developed to convey this magnitude of risk reduction in a way comprehensible to respondents. Instead of a risk reduction of x in 10,000 over one year, respondents were offered a risk reduction of x in 10,000 annually for 10 years, so in total the risk reduction would be x in 1000 over 10 years. As discussed below, focus group results in Ulaanbaatar led us to use two levels of annual risk reduction, 5 in 10,000 and 10 in 10,000 per year or cumulatively a 5-in-1000 or 10-in-1000 risk reduction over 10 years.

The sample design has two waves or treatments. Respondents in the first wave receive a 5-in-10,000 annual risk reduction for 10 years as the initial WTP question and a 10-in-10,000 annual risk reduction for 10 years as a second WTP question. Respondents in the second wave receive a 10-in-10,000 annual risk reduction for 10 years followed by a question offering a 5-in-10,000 annual risk reduction for 10 years (Table 1). The initial WTP question offers a contemporaneous risk reduction to all respondents. The second WTP question offers a latent risk reduction to respondents age 40–65 and a contemporaneous risk reduction to respondents over age 65. This structure allows testing for ordering effects among those over 65.

Table 1
Study design (cumulative probabilities over a ten-year period).

Group of respondents	Initial current risk reduction valued	Second risk reduction valued	
		Future risk reduction Respondents age 40–65	Current risk reduction Respondents over age 65
Wave 1 (n = 309)	5 in 1000	10 in 1000	
Wave 2 (n = 320)	10 in 1000	5 in 1000	

Note: Respondents over 65 did not get a latency question.

The basic form of the contemporaneous risk question is how much the respondent would be willing to pay for a product that, when used and paid for annually over the next 10 years, would reduce their baseline risk by x in 10,000 each year during the same period. A timeline is used to illustrate the payment period and the period during which the benefit is experienced. The latent risk reduction questions ask respondents how much they would be willing to pay for a product that, when used and paid for annually over the next 10 years, would reduce their annual mortality risk by x in 10,000 for a ten-year period beginning at age 70. Fig. 4 illustrates elicitation of WTP for latent, 5-in-10,000 annual risk reductions over 10 years for a 46-year-old respondent. Prior to receiving questions about WTP to reduce latent risks, respondents are asked the probability they think they have of surviving to age 70. A variety of surveys have shown that individuals are reasonably good at estimating future survival probabilities (Hamermesh, 1985; Hurd and McGarry, 1996).

An extensive series of debriefing questions follows the risk valuation questions. These questions test respondents' understanding of concepts presented in the survey, acceptance of various elements of the risk reduction scenario and baseline risk, and the presence of other factors that could call the credibility of answers into question.

4.2. Adaptation of the survey instrument

The survey instrument used in this study is based closely on one first developed for use in the United States (Alberini et al., 2004) and more recently adapted for use in China (Krupnick et al., 2010). In adapting the survey for use in Mongolia, we had two sometimes conflicting goals. The primary goal was to produce a survey that Mongolian respondents would understand and find credible. The secondary goal was to maintain as much consistency as possible with versions of the survey that had

Fig. 4. Payment card elicitation of willingness to pay (Mongolian version).

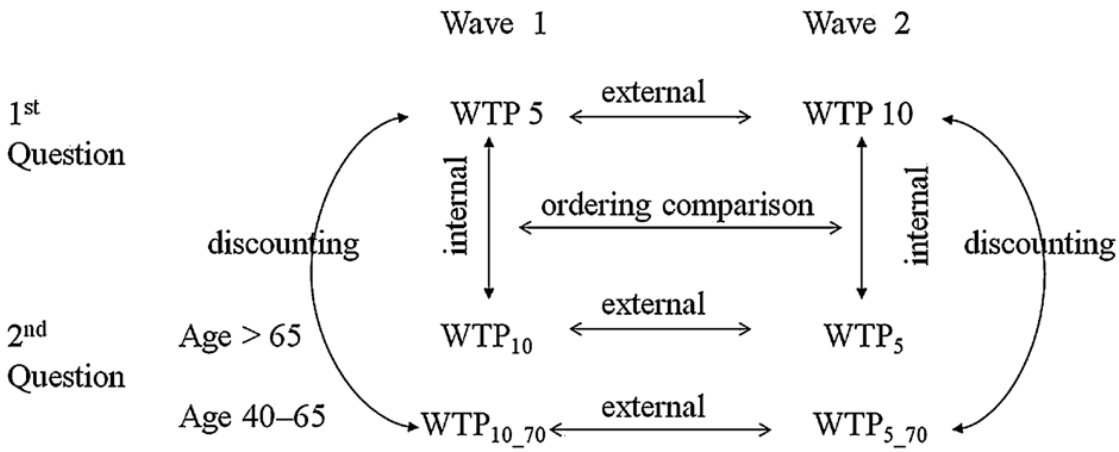

Fig. 5. Scope tests.

been administered successfully in other countries. This consistency enables results of this study to be used to test the reliability of benefits transfers among countries.

After translating the survey into Mongolian and back translating into English to verify the accuracy of the translation, a series of focus groups were held in Ulaanbaatar in fall 2009 to help adapt the survey. Several changes were needed to assure that the survey reflected Mongolian circumstances. These included descriptions of the structure of health care delivery, causes of mortality, gender-specific life expectancy, and activities individuals engage in to protect their health. Focus group participants found the use of a generic product to reduce mortality risks credible, but they also suggested that describing the product as being approved by United States and European, as well as Mongolian, health authorities would enhance its credibility. As in many developing countries, households in Ulaanbaatar often have multiple sources and types of income. The income question was revised to ask respondents to include all of these sources in reporting their household income. Focus groups also disclosed that there is a culture in Ulaanbaatar of not disclosing a small portion of one's income, which participants referred to as "secret income." As a result, income figures reported in this survey should be viewed as underestimating real household income. This would have the effect of inflating estimates of WTP as a percent of household income but it cannot be known by how much. Focus group participants recommended changing the reference to roulette wheels in the probability tutorial to a reference to a better known, but similar, wheel used on a popular Mongolian TV game show. Few other changes in risk communication were needed. A risk attitude question referring to air and train travel to measure aversion to safety hazards was changed to one related to walking across busy streets.

Focus groups were also used to test the appropriate size of risk changes. Participants in the focus groups found it difficult to understand or distinguish the magnitude of the smallest risk reduction, 1-in-10,000. Results from the original U.S. version of the survey indicated that U.S. respondents also had difficulty comprehending the 1-in-10,000 risk reduction (although the results passed an external scope test, i.e., WTP increased with the size of the risk reduction). The response of focus group participants as well as results from pretests of the survey indicated that Ulaanbaatar respondents did understand the meaning of 5-in-10,000 and 10-in-10,000 annual reductions in mortality risk and believed that the product offered in the survey would reduce mortality risks by these magnitudes.

4.3. Survey design and validity testing

The survey's design permits testing for consistency of responses with the underlying theoretical model and good survey design standards. The structure of the contemporaneous risk reduction questions allows tests of whether WTP is increasing with the size of the risk reduction by comparing responses between samples (external scope tests) (Fig. 5) and, for those 65 and above, within samples (internal scope test). The external scope test can also be performed for the latent risk reductions. The Mongolian survey design does not permit testing for the influence of the latent risk reduction question

being preceded by a contemporaneous risk reduction question.[3] An alternative form of the external scope test is to combine the subsamples and use multivariate regression analysis with a dummy variable to identify which sample got the larger risk reduction. This test is more difficult to pass than a split sample means test because of the *ceterus paribus* conditions imposed by the regression framework (although by doubling the number of observations, power is increased). Comparison of responses to contemporaneous and latent risk reduction questions allows tests for discounting. Finally, the study design allows for tests of whether contemporaneous question order affects responses for those over 65 years old (Fig. 5).

4.4. Interpretation of the payment card

The payment card elicitation approach (Fig. 4) presents respondents with a matrix of ordered numbers and asks them to pick one corresponding to their maximum willingness to pay for the risk reduction. The chosen "stated" number can be thought of as (a) an appropriate estimate of WTP, (b) the top end of an interval between the chosen number and the next lowest number (which we estimated using a Weibull distribution), or (c) the bottom end of an interval between this number and the next highest number (which we did not estimate). In order of conservativeness, what we will call the Weibull estimate is more conservative (produces a lower WTP estimate) than use of the stated value as an estimate. We also set WTP at the value in the matrix immediately below the chosen number, which in this context is similar to a Lower Turnbull estimate. This would be the most conservative of the three estimates used in our analysis.

The values presented in the payment card range from 0 to 500,000 tugrug per year for 10 years. Focus group results suggested that WTP in Ulaanbaatar would generally not exceed 500,000 tugrug per year. The range also includes the range of WTP values seen in other countries where this survey instrument was administered (using a purchasing power parity exchange rate conversion for comparisons). This helps assure that study results can be compared to those of previous studies. As shown in Fig. 4, numbers were set up in a matrix, rather than a list, to be compact and to mitigate visual starting point bias. Range intervals were designed to be roughly constant in percentage terms. Psychometric studies and experimental economics suggest this as a means of reducing response error (see Rowe et al. (1996) and Ready et al. (2001) for discussions of this literature). In the Mongolian survey, a box with a flashing cursor was placed to the side of matrix. The survey enumerator asked respondents which value in the matrix they would be willing to pay for the risk reduction.

4.5. Administration of the survey

The survey was administered in Ulaanbaatar, the capital of Mongolia, using random sampling stratified by neighborhood. Respondents came to a neighborhood center. The survey was administered on weekends to accommodate respondents' work schedules and was administered on laptop computers operated by enumerators who entered participants' responses. Enumerators received very explicit training regarding the level of interaction with respondents that was permissible to maintain consistency in administration across respondents. Participants were also given an inexpensive gift upon completion of the survey to thank them for participating after completing the survey. A sample of 629 respondents completed the survey.

5. Sample characteristics, estimation methods and results

5.1. Cleaning the sample

The survey ends with a number of debriefing questions that indicate whether respondents understood or accepted the scenarios presented in the survey. Table 2 provides descriptive statistics on

[3] We were limited in the number of survey versions that could be placed in the field. One additional version for each wave would have been needed to test this effect.

Table 2
Descriptive statistics for cleaning criteria variables.

Variable	Number obs.	% of total obs.
FLAG1[a]	6	0.95
OVER80[b]	1	0.16
FLAG6[c]	131	20.82
WTP/income $\geq$ 0.9	26	4.13

[a] FLAG1 is a categorical variable indicating that a respondent failed to pass one or more probability comprehension tests in the survey.
[b] OVER80 identifies respondents who were over 80 years of age.
[c] FLAG6 is a categorical variable indicating that a respondent answered that they did not understand probability well, i.e., less than 5 on a 10-point scale.

indicators that a respondent may not have understood aspects of the survey. Four of these indicators were used to clean the sample. Six respondents were dropped because they responded incorrectly to questions testing their understanding of the way probabilities of death were presented in the survey (FLAG1). Respondents over 80 are unlikely to survive to realize the benefits provided by the product. One respondent was over 80 and dropped, as were respondents over 80 in prior studies using this survey. A total of 131 respondents (or 20 percent of the full sample) said they did not understand probability well (FLAG6). WTP estimates did not pass scope when FLAG6 respondents were included in the sample, suggesting that more work may be necessary to overcome innumeracy when conducting this or similar surveys in developing countries. Finally, 26 respondents indicated that they were willing to pay more than 90 percent of their stated income for the indicated risk reduction. This indicates a failure to take the budget constraint seriously, a failure to understand the question being asked, or a protest bid. Use of this cleaning criterion was also important to WTP estimates passing scope tests. None of the cleaning variables were strongly correlated; correlation coefficients were all less than $|0.05|$.[4]

We applied three alternative cleaning approaches in our analysis. Cleaning approach A eliminates the respondent over 80 as well as respondents who failed tests of understanding of probability. Cleaning approach B eliminates these same respondents plus those who state they do not understand probability well. Cleaning approach C drops those in cleaning approach B plus respondents who expressed a willingness to pay that is 90 percent or more of their per capita income. In general, we report the results of cleaning approach A because it is the least restrictive. We do not report results for cleaning approach B, because they were generally quite similar to those for cleaning approach A. We report results for cleaning approach C because with this cleaning, WTP estimates behave in ways more consistent with the underlying theoretical model. We also found no statistically significant difference on major demographic variables between the sample remaining after cleaning approach C was applied and the sample as a whole (Table 3).

5.2. Sample characteristics

Table 4 reports descriptive statistics for the full sample. The study sample is more female than male (38 percent male). In Mongolia as a whole, 51 percent of the population is male (National Statistical Office of Mongolia, 2009). Ulaanbaatar residents account for 65 percent of the Mongolian population (National Statistical Office of Mongolia, 2009). Twenty-four percent of respondents have college

[4] A number of debriefing questions appeared unimportant in explaining responses to the survey and therefore were not used to clean the sample. We asked respondents whether they were confident that the product offered to reduce the risk was effective. All but 6 of 629 respondents said they were. None of these said that their lack of confidence affected their WTP. The variable indicating that the respondent was not confident in the effectiveness of the product was not statistically significant in explaining WTP. 593 of 629 respondents said they believed that the stated baseline health risk applied to them personally. Of those who did not believe it did, 15 believed their risk of dying was higher than stated, and 21 believed it was lower. Again, this was not a statistically significant effect.

Table 3
Comparison of demographic variables by cleaning and place of birth.

Variable definition	Whole sample Mean	Cleaning approach C[a] Mean	Comparison of means Wald tests (*p* value)
Socio-demographics			
Age	54.37	53.99	−0.61 (*0.54*)
Age distribution (%)			
40–49	38	39	0.32 (0.75)
50–59	31	31	1.05 (0.29)
60–69	21	21	0.48 (0.63)
70 and older	10	9	0.32 (0.75)
Gender (1 = male, 0 = female) (%)	38	40	0.71 (*0.48*)
Share of respondents with university degree (%)	24	25	0.66 (*0.51*)
Monthly income (tugrug)			
Household income	390,063	402,118	0.88 (*0.38*)
Per capita income	103,113	106,548	0.75 (*0.45*)
Observations	629	472	

[a] Cleaning approach C drops Flag1, Flag6, Over80, WTP/income ≥0.9.

degrees and 78 percent have completed high school. High school in Mongolia is compulsory and free of charge. Mean monthly household income in the sample, 390,063 tugrug (or 103,113 tugrug per capita), is roughly comparable to the national average household income of 399,096 tugrug per month, and much lower than the Ulaanbaatar average of 547,825 tugrug (National Statistical Office of Mongolia, February 2009). This difference may be due to the under-reporting of income noted above. Also, because the sample is skewed toward a younger population despite an effort to oversample in the older age categories, incomes may be lower than the average. Eighty-three percent of respondents expect to live to age 70. This may seem optimistic in light of the fact that 23 percent of respondents report having problems with bronchitis and 44 percent have heart disease. But in 2005, an average 60-year-old Mongolian woman could expect to live to age 78, a man to age 75 (United Nations, 2010). Life expectancy has risen rapidly in Mongolia and is projected to continue to increase in the coming years. While the average life expectancy at birth of Mongolians is currently 68, it was only 42 in 1950 and is projected to reach 73 by 2030 (United Nations, 2010).

Table 4
Descriptive statistics.

Variable definition	Whole sample			
	Mean	St. dev.	Min.	Max.
Socio-demographics				
Age	54.4	10.3	40	81
Age distribution (%)				
40–49	38			
50–59	31			
60–69	21			
70 and older	10			
Share of respondents with university education (%)	45	49	0	1
Health status (%)				
Respondent with bronchitis	23	42	0	1
Respondent with cancer	07	26	0	1
Respondent with heart disease	44	49	0	1
Chance to live to age 70	83	22	0	100
Monthly income (tugrug)				
Household income	390,063	223,358	150,000	850,000
Per capita income	103,113	75,680	12,500	42,5000
Number of observations	629			

Table 5
Stated value: external scope tests under alternative data cleaning approaches (respondents 40–65).

Cleaning criteria	Sample A (Flag1 and Over80)			Sample B (Flag1, Flag6, Over80)			Sample C (Flag1, Flag6, Over80, WTP/income ≥0.9)		
	n	Mean		*n*	Mean		*n*	Mean	
WTP5	303	265,402		232	256,362		221	221,067	
WTP10	319	286,802		260	297,434		251	256,665	
WTP5_70	238	130,299		189	132,474		184	132,376	
WTP10_70	270	206,394		223	213,361		210	178,902	
		t-Tests for significance of difference in means[a]							
			p			*p*			*p*
5 vs 10		−0.63	0.26		−1.04	0.15		−1.67	0.05
5_70 vs 10_70		−3.92	0.00		3.59	0.00		−2.76	0.00
5 vs 5_70		6.20	0.00		5.52	0.00		4.80	0.00
10 vs 10_70		2.36	0.01		2.08	0.02		3.79	0.00

Notes: WTP5 means WTP for an annual 5 in 10,000 risk reduction beginning immediately and occurring each year over a ten year period. WTP5_70 means the same thing, except the risk reduction would begin at age 70. WTP10 is WTP for an annual 10 in 10,000 risk reduction each year for ten years starting immediately. See Table 2 for definitions of cleaning criteria.

[a] Ho: wtp5 = wtp10; Ha: wtp5 < wtp10.

5.3. Estimation methodology

To estimate WTP for the assumption that the chosen value is the respondents' actual WTP and for the Lower Turnbull estimates, we use a Tobit model. For the interpretation of the stated WTP as lying within an interval (the Weibull model), the underlying econometric model is

$$\log WTP_i^* = X_i\beta + \varepsilon_i \tag{5}$$

where WTP^* is the underlying willingness to pay for a selected risk reduction; X denotes a vector of age, health, and other attributes; β is a vector of coefficients; and ε is an extreme value Type I error term. Effectively, Eq. (1) describes a survival time model based on the Weibull distribution. The log-likelihood function of the data is

$$\log L = \sum_{i-1}^{n} \left\{ F\left[\log \frac{WTP_i^H - X_i\beta}{\sigma}\right] - F\left[\log \frac{WTP_i^L - X_i\beta}{\sigma}\right] \right\} \tag{6}$$

where F is the Type I extreme value distribution with scale σ, WTP_i^H and WTP_i^L are upper and lower bounds for WTP, and X is a vector of age, health, and other attributes with β as the corresponding coefficients. σ is the scale parameter of ε, as well as the reciprocal of the shape parameter of the Weibull distribution describing WTP. The scale parameter for the Weibull distribution is $\exp(X\beta)$.

5.4. Results

Tables 5–7 present results for respondents aged 40–65. Analysis reported in Table 5 assumes that the values respondents choose from the payment card represent their true WTP. Analysis presented in Tables 6 and 7 assume a Weibull and a Lower Turnbull model, respectively. Under cleaning approach C, WTP responses pass external scope tests for both contemporaneous and latent risk reductions for all distributional assumptions (Tables 5–7). WTP for a contemporaneous 5-in-10,000 annual risk reduction ranged from 159,457 tugrug using a Lower Turnbull estimation to 221,067 assuming stated WTP is the respondents' true WTP. Results show positive and statistically significant discounting for all cleaning approaches and distributional assumptions (Tables 5–7). We could find no statistical evidence that individual WTP varied by age when other explanatory factors were allowed to vary.

Tables 8 and 9 present results for respondents over 65. The survey design allows for internal scope tests to be run only on those over 65. Table 8 treats stated WTP as respondents' true WTP. Table 9 assumes the true value lies within the stated and the next lower value, described by a Weibull distribution. WTP estimates pass internal scope for all cleaning approaches for both models (Tables 8 and 9).

Table 6
Weibull: external scope tests under alternative data cleaning approaches (respondents 40–65).

Cleaning criteria	Sample A (Flag1 and Over80)			Sample B (Flag1, Flag6, Over80)			Sample C (Flag1, Flag6, Over80, WTP/income ≥0.9)		
	n	Mean		*n*	Mean		*n*	Mean	
WTP5	299	200,192		229	202,175		218	182,624	
WTP10	317	213,851		258	222,279		249	211,664	
WTP5_70	210	127,291		168	129,622		164	129,622	
WTP10_70	248	176,732		204	180,612		192	165,945	
		t-Test for significance of difference in means[a]							
			p			*p*			*p*
5 vs 10		−16.18	0.00		−18.75	0.00		−28.78	0.00
5_70 vs 10_70		−54.79	0.00		−45.91	0.00		−34.26	0.00
5 vs 5_70		86.52	0.00		67.20	0.00		51.39	0.00
10 vs 10_70		41.07	0.00		37.77	0.00		42.72	0.00

[a] Ho: wtp5 = wtp10; Ha: wtp5 < wtp10.

Table 7
Lower Turnbull (lower bound value): external scope tests under alternative data cleaning approaches (respondents 40–65).

Cleaning criteria	Sample A (Flag1 and Over80)			Sample B (Flag1, Flag6, Over80)			Sample C (Flag1, Flag6, Over80, WTP/income ≥0.9)		
	n	Mean		*n*	Mean		*n*	Mean	
WTP5	303	169,274		232	173,772		221	159,457	
WTP10	319	185,721		260	194,481		251	185,717	
WTP5_70	238	101,933		189	105,053		184	104,647	
WTP10_70	170	143,315		223	145,538		210	136,857	
		t-Tests for significance of difference in means[a]							
			p			*p*			*p*
5 vs 10		−1.35	0.09		−1.49	0.07		−1.96	0.03
5_70 vs 10_70		−3.41	0.00		−3.17	0.001		−2.58	0.01
5 vs 5_70		5.84	0.00		2.04	0.02		4.36	0.00
10 vs 10_70		3.41	0.00		3.55	0.00		3.63	0.00

[a] Ho: wtp5 = wtp10; Ha: wtp5 < wtp10.

Table 8
Stated value: internal scope tests under alternative data cleaning approaches (respondents over 65 years old).

Cleaning criteria	Sample A (Flag1 and Over80)			Sample B (Flag1, Flag6, Over80)			Sample C (Flag1, Flag6, Over80, WTP/income ≥0.9)		
	n	Mean	*p*	*n*	Mean	*p*	*n*	Mean	*p*
First WTP5	49	217,551		37	205,000		36	204,027	
Second WTP10	49	325,673		37	322,378		36	321,333	
t-Test		−1.66	0.05		−1.61	0.06		−1.57	0.06
First WTP10	65	249,400		43	264,116		42	263,262	
Second WTP5	65	118,230		43	123,488		43	123,488	
t-Test		3.92	0.00		3.22	0.00		3.16	0.00

The influence of question order on responses can only be tested for respondents over 65. Ordering effects were found with cleaning approach A, both when treating the stated value as respondents' true WTP and when using a Lower Turnbull model (Table 10). When respondents who fit cleaning criteria A received a 10-in-1000 risk reduction over 10 years as the first question, they were willing to pay less for a 5-in-1000 risk reduction over 10 years than those who were asked the willingness to pay for a 5-in-1000 risk reduction over 10 years as their first question. Correspondingly, those who received the 5-in-1000 risk reduction over 10 years as the first question were willing to pay more for a 10-in-1000 risk reduction over 10 years than those presented with it as a first question. These results are evidence for a question ordering effect that would have implications for interpreting an internal

Table 9
Weibull: internal scope tests under alternative data cleaning approaches (respondents over 65 years old).

Cleaning criteria	Sample A (Flag1 and Over80)			Sample B (Flag1, Flag6, Over80)			Sample C (Flag1, Flag6, Over80, WTP/income ≥0.9)		
	n	Mean	*p*	*n*	Mean	*p*	*n*	Mean	*p*
First WTP5	47	175,138		36	169,258		35	167,552	
Second WTP10	48	246,530		37	242,583		36	239,628	
t-Test		−12.4	0.00		−10.43	0.00		−9.89	0.00
First WTP10	65	189,736		43	201,559		42	199,509	
Second WTP5	53	120,913		35	128,747		34	131,261	
t-Test		20.78	0.00		14.19	0.00		12.81	0.00

scope test (but not an external scope test). While no statistically significant ordering effect is found with cleaning approach C, this is probably due to the small sample size.

Tables 12 and 13 present results from regression analysis for contemporaneous and latent risk reductions pooling the entire dataset. We show results for cleaning methods A and C only. Results for cleaning approach B are quite similar to those for cleaning approach A. We show the results only for Weibull models. Results for the Lower Turnbull regressions were virtually identical to these and therefore are not presented. Models defining stated WTP as "true" WTP did not perform as well as those assuming a Weibull distribution.

We look first at contemporaneous risk reduction regression results (Table 11). Under both cleaning criteria A and C, the coefficient for income is highly significant and positive. Higher income respondents are willing to pay more to reduce mortality risk. With cleaning approach C, but not A, a cumulative risk reduction of 5-in-1000 over 10 years is significant and negative, again indicating the results pass scope, even controlling for observable sample characteristics. For cleaning approach A, we find an inverted-parabolic relationship between age and WTP, with a maximum WTP at 60 years of age. But, additional regressions run using a dummy variable for respondents age 60 and over, rather than age and age-squared, did not find a significant effect. A variable interacting gender with age or with age categories was also tried as an alternative to using age and gender separately, but was not significant. In total, this is at best weak evidence of an age effect on WTP in Ulaanbaatar. Existing health conditions do not influence WTP. Baseline risk and age have a correlation of 0.82. As in most past studies that use this survey instrument, baseline risk is also not a significant predictor of WTP when substituted for age and gender. When added to the Weibull regressions reported in Table 11, baseline risk was not itself significant. Its inclusion increased the significance of age and age-squared but had little effect on the sign or size of any other coefficients. In cleaning approach A, being born in Ulaanbaatar and having a university education results in higher WTP, independent of reported income. These variables could be standing in for unreported income. Those who are more risk averse are willing to pay more, although this affect is not significant with cleaning criteria C. Use of robust variance estimators does not change regression results, indicating that heteroskedasticity is not playing a strong role.

Table 12 shows results for regressions on WTP for a latent risk reduction. Income has a significant and positive influence on WTP. The coefficient on the 5-in-10,000 annual risk reduction relative to that of 10-in-10,000 is negative and significant, indicating that WTP estimates pass the scope test with both cleaning approaches A and B. The relationship between age and WTP is not significant. Again, baseline risk was not significant when substituted for age and gender and use of robust variance estimation did not affect results.

6. Discussion

6.1. Implications for Mongolia

In choosing a value of statistical life to use in Mongolian policy analysis, we consider construct validity (performance on split sample tests and regressions) and the appropriateness of the risk reduction type and magnitude to the policy setting. In other countries studied, a 5-in-10,000 risk reduction has

Table 10
Test of ordering effect with data cleaning approach A (respondents over age 65).[a]

	Stated value			Weibull			Lower Turnbull		
	n	Mean	*t*-stat (*p*)	*n*	Mean	*t*-stat (*p*)	*n*	Mean	*t*-stat (*p*)
WTP5 first question	49	217,551	2.44	47	175,138	13.52	49	148,775	−2.59
WTP5 second question	65	118,231	(0.02)	48	120,913	(0.00)	65	88,231	(0.01)
	Stated value			Weibull			Lower Turnbull		
	n	Mean	*t*-stat (*p*)	*n*	Mean	*t*-stat (*p*)	*n*	Mean	*t*-stat (*p*)
WTP10 first question	65	249,400	−1.38	65	189,736	30.07	65	159,153	−1.56
WTP10 second question	49	325,674	(0.17)	48	246,530	(0.00)	49	204,286	(0.12)

[a] 2-Tailed test.

Table 11
Weibull: construct validity of WTP for the current risk reduction.

Variable	Sample A[a]		Sample C[b]	
	Coef.	(S.E.)	Coef.	(S.E.)
Intercept	5.05***	(1.42)	2.19	(1.56)
Age	0.08*	(0.04)	0.07	(0.04)
Age square	−0.00*	(0.00)	−0.00*	(0.00)
University education (=1)	0.19**	(0.08)	0.12	(0.08)
Gender dummy (1 = male)	0.05	(0.08)	−0.02	(0.08)
Per capita monthly income (log form)	0.38***	(0.06)	0.60***	(0.06)
Born in Ulaanbaatar (=1)	0.19**	(0.08)	0.13	(0.08)
Risk averse (=1)	0.17**	(0.08)	0.02	(0.08)
Heart disease dummy	−0.08	(0.08)	−0.05	(0.08)
Bronchitis dummy	0.06	(0.09)	0.07	(0.09)
Cancer dummy	0.09	(0.15)	0.05	(0.15)
If the risk variable is 5 in 1000 reduction	−0.03	(0.08)	−0.13*	(0.07)
Scale parameter	1.08	(0.04)	1.26	(0.05)
Number of observations	616		468	

[a] Sample A cleaning approach: drop Flag1, Flag6 and >80 years old.
[b] Sample C cleaning approach: drop Flag1, Flag6, >80, and WTP/income ≥ 0.9.
* Significant at 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

been the most robust for use in assessing policies that reduce conventional air pollution and roughly matches the health risk reductions generally expected from air pollution policies. Given the extreme pollution conditions in Ulaanbaatar, it remains to be seen whether this over- or under-estimates reductions that can actually be achieved. Until better information is available on the effectiveness of air pollution policy in Ulaanbaatar, the 5-in-10,000 risk reduction is the best available estimate of expected effects.

Mongolia's relatively short life expectancy raises a question about the validity of including older respondents in a calculation of the VSL for Ulaanbaatar and about the validity of results for WTP

Table 12
Weibull: construct validity of WTP for the latent risk reduction.[a]

Variable	Sample A[a]		Sample C[b]	
	Coef.	(S.E.)	Coef.	(S.E.)
Intercept	3.15	(2.67)	1.63	(2.76)
Age	0.12	(0.09)	0.11	(0.09)
Age square	−0.00	(0.00)	−0.00	(0.00)
University education (=1)	0.11	(0.09)	0.03	(0.09)
Gender dummy (1 = male)	−0.001	(0.09)	−0.02	(0.10)
Per capita monthly income (log form)	0.38***	(0.08)	0.52***	(0.08)
Born in Ulaanbaatar (=1)	0.18**	(0.09)	0.05	(0.10)
Risk averse (=1)	0.12	(0.09)	−0.08	(0.10)
Heart disease dummy	−0.12	(0.09)	0.01	(0.09)
Bronchitis dummy	0.11	(0.11)	0.18	(0.12)
Cancer dummy	0.28	(0.17)	0.21	(0.18)
If the risk variable is 5 in 1000 reduction	−0.47***	(0.09)	−0.32*	(0.09)
Chance to survive to 70	0.005**	(0.002)	0.002	(0.002)
Scale parameter	1.06	(0.04)	1.16	(0.04)
Number of observations	458		356	

[a] Sample A cleaning approach: drop Flag1, Flag6 and >80 years old.
[b] Sample C cleaning approach: drop Flag1, Flag6, >80, and WTP/income ≥ 0.9.
* Significant at 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

for latent risks. Average life expectancy at birth in Mongolia is 68.1 years, compared to 79.9 in the United States or 83.1 in Japan (U.N. 2010). For those who survived to age 65 in 2005, life expectancy in Mongolia was 15 years for women and 12 for men. This compares to a life expectancy of 20 for 65-year-old women in the United States and 15 for U.S. men. In part for this reason and in part because current policy proposals focusing on reducing particulate matter pollution will primarily provide contemporaneous rather than latent health benefits, we recommend relying on estimates of WTP for contemporaneous risk reductions based on responses from respondents 65 and younger in evaluating current air pollution policy proposals in Ulaanbaatar. That said, investments in pollution control and sanitation in Ulaanbaatar should lead to increased life expectancy over time. Other changes that are part of the process of development that Mongolia expects, such as improved health care, increased educational levels, and economic growth, are expected to increase life expectancy during the lifetimes of survey respondents (United Nations, 2010). As a result, from a policy perspective, it should also be of interest to understand respondents' WTP to reduce their own annual mortality risk in their old age.

Critics often claim that stated preference studies overstate WTP because respondents do not actually have to expend their own money and do not receive actual benefits. As explained above, the Lower Turnbull estimate of WTP as applied in this analysis provides a more conservative (lower) estimate of willingness to pay than either the stated value or the Weibull-based estimate. In China, a Lower Turnbull estimate was used for policy analysis, although in other countries where this survey was administered the Weibull estimates were used. If policymakers in Mongolia are concerned about overstatement of WTP, a conservative approach would be to use the Lower Turnbull estimates in assessing policy.

Choice of sample cleaning criteria is based on model performance. In general, all cleaning approaches provided similar results in this study. However, on perhaps the most important criterion, whether WTP is increasing with scope, only cleaning approach C passed all scope tests for all modeling assumptions. Results for both cleaning approaches A and B fail to pass external scope when the stated WTP is assumed to be the true WTP and in the contemporaneous risk reduction regressions. Given these considerations, one justifiable interpretation of these results is that a Lower Turnbull estimate of WTP for a 5-in-10,000 risk reduction using cleaning approach C for respondents 65 and younger provides a conservative and reliable estimate of Ulaanbaatar residents' WTP for reductions in mortality risk that are likely to result from air pollution policy. Under these assumptions, this study estimates that individuals are willing to pay 159,000 tugrug for a 5-in-10,000 contemporaneous risk reduction (Table 7). This implies a VSL of 319 million tugrug. With reported monthly household income of 402,118 tugrug, WTP is 3.3 percent of annual household income. For international comparison purposes, using a purchasing power parity (PPP)-based exchange rate, the VSL is $493,000 ($U.S.), while on an official exchange rate basis, the VSL is $221,000 (U.S.).

6.2. Comparisons of results from other countries using the "same" survey

As already noted, the survey reported on above is very similar to surveys administered to respondents in other countries, including the United States, Canada, Japan, France, Italy, United Kingdom, and China. Key results from these surveys are compared in Table 13. This comparison is based on the risk changes and statistical assumptions that are common to all eight studies: a 5-in-10,000 annual risk reduction and the use of an estimate assuming they are distributed within an interval according to a Weibull distribution.[5] PPP exchange rates are used in this study to facilitate comparisons across countries because they provide a more realistic comparison of the trade-offs consumers face in making expenditures in different countries. PPP exchange rates are based on the price of a similar basket of consumer goods in different countries rather than currency exchange rates that may be set or influenced by government policy for macroeconomic reasons and may be affected by capital flows (which are influenced by many financial factors that are not closely related to consumer preferences).

[5] Note that for the Chinese and Mongolia studies, the interval is formed from a stated WTP and the next lower value on a payment card while for the other countries, the interval is defined from the double-bound dichotomous choice responses.

Table 13
Mean VSL by country and study for 5-in-10,000 annual risk change (conversion to $US using PPP and official exchange rates for Mongolia; PPP for all other countries).

	Mongolia Ulaanbaatar	China Shanghai, Juijiang, Nanning	Canada Hamilton, Ontario	U.S. Entire country	Japan Shizuoke	U.K. Bath	France Strasbourg	Italy 5 cities
WTP (current 5 in 10,000) as a % of average household income	3.3%	1.68%	1.00%	1.45%	0.81%	1.59%	7.71%	3.50%
Current VSL: 5/10,000 $ US (millions)	0.25[a] 0.57[b]	0.44	0.93	1.54	0.66	1.17	4.56	2.28
Scope test: ratio of VSLs for 10 vs 5 in 10,000 risk reduction	1.15	1.21	1.3	1.6	1.5	NA	?	NA
Latent VSL: 5/10,000 $US (millions)	0.18[a] 0.40[b]	0.39	0.53	0.69	0.48	0.51	1.25	0.87
Ratio of future to contemporaneous VSL 5 in 10,000 risk reduction	0.71	0.90	0.57	0.45	0.73	0.44	0.27	0.38

[a] Official exchange rate: 1447 tugrug/$ (2009) (World Bank, 2010).
[b] PPP exchange rate: 635 tugrug/$ (2009) (U.S. CIA, 2010).

Official exchange rates are also reported for Mongolia because of uncertainty in the PPP calculations for a small economy like Mongolia's.

While significant effort was made to make certain that the studies were as similar as possible, there are some unavoidable differences. WTP was elicited using a payment screen format in China, a payment card format in Mongolia and a dichotomous choice (DC) approach in the other six countries. The U.S. study used a nationally representative sample. All other studies reflect WTP in urban areas in each country.

The top row of Table 13 gives mean WTP for a 5-in-10,000 risk reduction as a percentage of respondents' average household income in the various country studies. Because Mongolia is considerably poorer than the other countries studied, we would expect a lower VSL, but not necessarily a lower WTP as a percentage of household income. The Ulaanbaatar study implies a VSL for a 5-in-10,000 contemporaneous risk reduction (from Table 6) of $565,000 U.S. using PPP rates and $252,000 using official exchange rates based on a Weibull model. These are in the low range of VSL estimates in studies using this survey instrument, most comparable to those of China (Table 13). For latent WTP, the corresponding VSLs are $401,000 and $179,000, respectively. Again, as expected, these values are on the lower range of those found in other countries. However, WTP is higher as a percentage of household income in the Mongolian study than in all other study countries except France and Italy. To some extent, this may reflect income underreporting in the Mongolian study. Sensitivity to scope for Mongolia, represented as the ratio of WTP for 10- to 5-in-10,000 risk reduction (1.15) is lower than, but close to those found in other studies, ranging from 1.21 for China to 1.6 in the U.S. study. Our qualitative metric for discounting, expressed as the ratio of WTP for future risk reductions divided by WTP for contemporaneous risk reductions, is in the high range of those found in other study countries. Mongolians (with a ratio of 0.71) appear to be as or less future-oriented than respondents in other Asian countries in the set, namely China (0.90) and Japan (0.73) and more future oriented than those Canada (0.57) and the U.S. (0.45).

7. Conclusion

This study presents the results of a new stated preference study of WTP to reduce mortality risk conducted by the Mongolian Ministry of Health and Resources for the Future, funded by the World Bank. To the best of our knowledge, it is the first such study to be conducted in Mongolia. The study estimates willingness to pay for mortality risk reduction in the Mongolian capital of Ulaanbaatar. The risk reduction scenarios include both a contemporaneous risk reduction and a future risk reduction because conventional air pollution causes both immediate and delayed increases in mortality risk. The study uses a survey instrument adapted from one that has been used in seven other countries, including China. This was done to permit international comparisons and, ultimately, to examine benefits transfer issues. The survey also uses a payment card to elicit willingness to pay. This was done because our administration of the survey in China and another stated preference study in Inner Mongolia found a payment card or "payment screen" was a useful means of overcoming excessive "yea-saying" encountered with use of a dichotomous choice payment vehicle in these studies.

The survey was fielded in Ulaanbaatar in January and February 2010. Several analyses were conducted: evaluation of respondents' understanding of the survey instrument, estimation of WTP, analysis of the robustness of WTP estimates to different approaches to data cleaning, tests for response to the scope of the risk reduction and to latency of the risk reduction, and regressions to check for construct validity.

The relationships between WTP for risk reductions of various sizes and WTP for contemporaneous and future risk reductions were generally in the direction expected from theory. These relationships were also statistically significant with appropriate sample cleaning and estimation procedures. We found a higher than usual percentage of respondents who exhibited some problem in understanding or accepting the survey; however, regression analysis checking for construct validity performed well.

Critics often claim that stated preference studies overstate WTP because respondents do not actually have to expend their own money and do not receive actual benefits. To address this tendency, we rely on the Lower Turnbull estimates rather than stated WTP or the Weibull-based estimates as a measure of true WTP. Given the appropriate performance of the instrument, a VSL of 319 million tugrug for

a contemporaneous 5-in-10,000 annual risk reduction is justified by this study for use in cost-benefit analyses. This value translates into a VSL in US$ of $221,000 on the basis of the official exchange rate or $493,000 on a PPP basis. These values should be interpreted as only valid for Ulaanbaatar, not outlying rural areas of Mongolia.

Irrespective of what the appropriate number is for a VSL used in a cost-benefit analysis, this study shows that people living in Ulaanbaatar place a high value relative to their income on reducing their risks of death. Given recent monitoring showing that Ulaanbaatar has perhaps the worst particulate air pollution of any city in the world coupled with substantial scientific research strongly associating this level of pollution with significant mortality and minimal abatement efforts taken to date, policymakers in Mongolia as well as international lenders, such as the World Bank, can feel secure that significant action to reduce such air pollution is warranted.

References

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

Alberini, A., Cropper, M., Krupnick, A., Simon, N., 2006a. Willingness to pay for mortality risk reductions: does latency matter? Journal of Risk and Uncertainty 32 (3), 231–245.

Alberini, A., Hunt, A., Markandya, A., 2006b. Willingness to pay to reduce mortality risks: evidence from a three-country contingent valuation study. Environmental and Resource Economics 33 (2), 251–264.

Aldy, J.E., Viscusi, W.K., 2007. Age difference in the value of statistical life: revealed preference evidence. Review of Environmental Economics and Policy 1 (2), 241–260.

Arthur, W.B., 1981. The economics of risks to life. American Economic Review 71 (1), 54–64.

Bhattacharya, S., Alberini, A., Cropper, M., 2007. The value of mortality risk reductions in Delhi, India. Journal of Risk and Uncertainty 34 (1), 21–47.

Champ, P., Bishop, R., 2006. Is willingness to pay for a public good sensitive to the elicitation format? Land Economics 82 (2), 162–173.

Cropper, M., Sussman, F., 1990. Valuing future risks to life. Journal of Environmental Economics and Management 19 (2), 160–174.

Finucane, M., Slovic, P., Mertz, C.K., Flynn, J., Satterfield, T., 2000. Gender, race, and perceived risk: the 'White Male' effect. Health, Risk and Society 2, 159–172.

Flynn, J., Slovic, P., Mertz, C.K., 1994. Gender, race, and perception of environmental health risks. Risk Analysis 14 (6), 1101–1108.

Hammitt, J., Zhou, Y., 2006. The economic value of air-pollution-related health risks in China: a contingent valuation study. Environmental and Resource Economics 33 (3), 399–423.

Hamermesh, D., 1985. Expectations, life expectancy, and economic behavior. Quarterly Journal of Economics 100 (2), 389–408.

Hurd, M., McGarry, K., 1996. The Predictive Validity of the Subjective Probabilities of Survival in the Health and Retirement Survey. Economics Department, State University of New York at Stony Brook, Stony Brook, NY.

Itaoka, K., Krupnick, A., Saito, A., Akai, A., Alberini, A., Cropper, M., Simon, N., 2005. Age, health, and the willingness to pay for mortality risk reductions: a contingent valuation survey in Japan. Discussion paper 05-34. Resources for the Future, Washington, DC.

Krupnick, A., Alberini, A., Cropper, M., Simon, N., O'Brien, B., Goeree, R., Heintzelman, M., 2002. Age, health, and the willingness to pay for mortality risk reductions: a contingent valuation study of Ontario residents. Journal of Risk and Uncertainty 24 (2), 161–186.

Krupnick, A., 2007. Mortality-risk valuation and age: stated preference evidence. Review of Environmental Economics and Policy 1 (2), 261–282.

Krupnick, A., Hoffmann, S., Qin, P., 2010. Willingness to pay for mortality risk reductions. Draft working paper.

Li, Y., Bai, M., Zhang, W., Yang, K., Wang, X., 2002. Analysis on the influence factors of residents' willingness to pay for improving air quality in Beijing. China Population, Resources and Environment 12 (6), 123–126.

National Research Council, 2008. Estimating Mortality Risk Reduction and Economic Benefits from Controlling Ozone Air Pollution. National Academies Press, Washington, DC.

National Statistical Office of Mongolia, 2009. Yearbook. Ulaanbaatar, Mongolia.

Ready, R., Navrud, A., Dubourg, W.R., 2001. How do respondents with uncertain willingness to pay answer contingent valuation questions? Land Economics 77 (3), 315–326.

Rowe, R., Schulze, W., Breffle, W., 1996. A test for payment card biases. Journal of Environmental Economics and Management 31 (2), 178–185.

Shepard, D., Zeckhauser, R., 1982. Life-cycle consumption and willingness to pay for increased survival. In: Jones-Lee, M.W. (Ed.), The Value of Life and Safety. North Holland, Amsterdam, the Netherlands.

Shepard, D., Zeckhauser, R., 1984. Survival versus consumption. Management Science 30, 423–439.

United Nations. UNdata: A World of Information. Life expectancy at birth and life expectancy at age *x*. http://data.un.org (accessed 20.10.10).

U.S. CIA (U.S. Central Intelligence Agency), 2010. The World Factbook. https://www.cia.gov/library/publications/the-world-factbook/geos/mg.html (accessed 24.11.10).

Wang, H., Mullahy, J., Chen, D., Wang, L., Peng, R., 2001. Willingness to pay for reducing the risk of death by improving air quality: a contingent valuation study in Chongqing, China. In: Paper Presented at the Third International Health Economic Association Conference, July 22–25, York, UK.

WHO (World Health Organization), 2006. WHO Air Quality Guidelines for Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide: Global Update 2005 Summary of Risk Assessment. WHO/SDE/PHE/OEH/06.02. World Health Organization, Geneva, Switzerland.

World Bank, 2011. Mongolia: Air Quality Analysis of Ulaanbaatar: Improving Air Quality to Reduce Health Impacts. World Bank Discussion Paper Series 66082.

World Bank, 2010. World Development Indicators. World Bank Press, Washington, DC, p. 279.

Yaari, M., 1965. Uncertain lifetime, life insurance and the theory of the consumer. Review of Economic Studies 32 (2), 137–150.

Zhang, M., Fan, J., Hu, C., Zhang, B., 2004. Assessment of total economic value of improving atmospheric quality of Lanzhou. Journal of Arid Land Resources and Environment 18 (3), 28–32 (in Chinese).

Zhang, X., 2002. Valuing Mortality Risk Reductions using the Contingent Valuation Methods: Evidence from a Survey of Beijing Residents in 1999. Centre for Environment and Development, Chinese Academy of Social Sciences, Beijing, China.

Xu, Z., Cheng, G., Zhang, Z., Su, Z., Johan, L., 2003. Applying contingent valuation in China to measure the total economic value of restoring ecosystem services in Ejina Region. Ecological Economics 44 (2–3), 345–358.