IDEC DP^2 Series Vol. 7 No. 3

A Nonparametric Welfare Analysis via a Randomized Conjoint Field Experiment: an Application to Water Quality Improvement and the Floating Settlements on Inlay Lake, Myanmar

Su Thet Hninn
Department of Trade, Ministry of Commerce, Myanmar

Keisuke Kawata
Institute of Social Science
the University of Tokyo

Shinji Kaneko
Graduate School for International Development and Cooperation (IDEC)
Hiroshima University

Yuichiro Yoshida
Graduate School for International Development and Cooperation (IDEC)
Hiroshima University

Department of Development Policy
Division of Development Science
Graduate School for International
Development and Cooperation (IDEC)
Hiroshima University
1-5-1 Kagamiyama, Higashi-hiroshima
7398529 Japan

A Nonparametric Welfare Analysis via a Randomized Conjoint Field Experiment: an Application to Water Quality Improvement and the Floating Settlements on Inlay Lake, Myanmar

Su Thet Hninn, Shinji Kaneko, Keisuke Kawata, Yuichiro Yoshida*

October 20, 2017

Abstract

This study proposes a new approach to survey-based empirical welfare analysis, which combines a new design of the conjoint experiment of Hainmueller et al. (2014) and a non-parametric rational choice model. We focus on the welfare impact of a multi-attribute policy and report the identification result of the marginal component effect on the distribution of willingness-to-pay. As an illustration, the paper evaluates a water improvement policy package for the not previously analyzed floating settlements on Inlay Lake, Myanmar. Our estimation result shows that the average surplus gain from the lake water quality improvement is at least as large as 5.9% of the average annual per capita income of those on the lake. Moreover, attributes such as toilet provision have a clear welfare effect.

JEL Codes: Q53, Q56, Q58

Keywords: conjoint experiment, non-parametric welfare analysis, willingness-to-pay

*Hninn: Department of Trade, Ministry of Commerce, Myanmar (e-mail: yukihninn@gmail.com); Kaneko: Graduate School for International Development and Cooperation (IDEC), Hiroshima University, 1-5-1 Kagamiyama, Higashi Hiroshima, Hiroshima, JAPAN 7398529 (e-mail: kshinji@hiroshima-u.ac.jp); Kawata: Institute of Social Science, the University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, JAPAN 7398529 (email: keisukekawata@iss.u-tokyo.ac.jp); Yoshida: Graduate School for International Development and Cooperation (IDEC), Hiroshima University, 1-5-1 Kagamiyama, Higashi Hiroshima, Hiroshima, JAPAN 7398529 (email: yuichiro@hiroshima-u.ac.jp).

1 Introduction

The estimation of policy preferences is a central challenge in various policy studies including those in political science, economics, environmental science, and development studies. A difficulty in the estimation is multi-dimensionality: a policy typically has multiple attributes, for instance, an environmental policy consists of its target area, period, and cost. To evaluate a multi-attribute policy, the present paper connects a newly developed survey design and commonly used measurement in policy evaluation, namely, Hainmueller et al. (2014)'s design of the conjoint experiment and willingness-to-pay (WTP).

Recently, a conjoint experiment and the associated analytical methods offered by Hainmueller, Hopkins, and Yamamoto (2014) (referred as an HHY conjoint here after) has become a popular survey experiment design to estimate preferences for multi-attribute policy[1]. Their approach non-parametrically estimates the causal effect of individual policy attributes on observable choice outcomes (e.g., rating and/or ranking among alternative policies). The advantage of the HHY conjoint is its design-based inference; while the causal inference in traditional conjoint experiments is based on explicit decision-making models (e.g., the random utility model; see Train 2000), HHY conjoint inference is not. It is instead, based on the survey design. All causal quantities of interest are defined and identified based on Rubin's potential outcome framework (Neyman 1923; Rubin 1974), and their estimation result is then free from bias due to specification error in the decision-making model[2].

However, design-based approaches have had the common limitation that the approach cannot estimate the causal effect on "unobservable" outcomes, of which, probably the most relevant from policy implementation perspective is WTP. WTP is the maximum amount that people are willing to pay for policy implementation, which is a common measurement in policy evaluation among not only academic researchers but also policy practitioners (see Mishan and Quah 2007). Conventionally, WTP is defined over an explicitly-specified decision-making model, while the original inference of HHY conjoint does not assume any explicit specification. Conventional conjoint analyses are based on specific choice models that require parametric assumptions on the preference distribution. For wexample, the conventional random utility model assumes that the preference for observable characteristics is additive separable, the preference for unobservable characteristics follows the type-I extreme value distribution, and the preference parame-

[1] An increasing number of studies use the new randomized conjoint analysis to examine policy preferences. These include studies on international environmental agreements (Bechtel and Scheve 2013, Gampfer, Bernauer, and Kachi 2014, Bernauer and Gampfer 2015) and migration policies (Hainmueller and Hopkins 2015).

[2] Hainmueller et al. (2015) provide evidence that the HHY conjoint performs well at replicating real-world behavior by comparison with natural experiments.

ters of observable characteristics follow the normal distribution[3]. The estimation results then potentially depend on these parametric assumptions.

Our approach is based on a rational choice model but without any functional form assumption on the preference distribution. Without sacrificing the virtue of the HHY conjoint –of not depending on any specific decision-making model– the present paper introduces a new approach to draw welfare implications from the HHY conjoint with minimal behavioral assumptions. We define the marginal component effect on the WTP distribution (MCE-WTP) as a causal effect of each policy attribute on the marginal WTP distribution and show that the MCE-WTP can be non-parametrically recovered from the average marginal component effect introduced in HHY conjoint.[4]

This paper first theoretically identifies that the WTP distribution of a policy alternative defined by a set of given values of attributes (referred to as the conditional WTP distribution, the definition of which is provided below) can be recovered. Therefore, we can potentially estimate the exact relationship between the WTP distribution and a multi-attribute policy. Similarly, we show that a causal effect of a change in each policy attribute on conditional WTP distributions can also be theoretically recovered. However, there are two problems in directly applying these identification results in practice. The first problem is that in many cases, the sample size may not be insufficient to empirically recover the distributions. The second problem is the limited variation in the "individual policy burden" of the conjoint experiment.

This paper thus introduces a more practical causal quantity, MCE-WTP, that allows for the evaluation of the causal impact of each attribute on the marginal WTP distribution given a reasonable sample size, where the distributional effects of attributes other than the financial burden and the attribute of interest are marginalized. Here, if the survey design has enough and continuous variation in the monetary burden of policy implementation (and the sample size is large enough), the MCE-WTP is globally recovered from the conjoint data, and the MCE-WTP is then point-identified. However, the policy burden often has only limited discrete variation. For instance, in our illustration, the individual burden has only four levels, namely, 5,000, 10,000, 15,000, and 20,000 kyats[5]. We therefore show that, even in the case of limited variation in the level of policy burden, the MCE-WTP distribution is partially estimated. Furthermore, we also demonstrate the estimation of a lower bound of the average marginal WTP, as the recovery of the marginal WTP distribution is less restrictive to sampling size than that of the MCE-WTP.

[3]Furthermore, a conventional welfare analysis based on a logit model assumes that the underlying utility function is log-linear in the numerare (see, for example, Small and Rosen, 1981).

[4]Throughout the paper, the term "marginal" is used in the sense of marginal probability distribution, such that the marginal variable integrates away all other "conditional" variables into the margin of the table.

[5]The kyat is a currency unit in Myanmar.

As an application of our approach, we then estimate the WTP of floating settlements on Inlay lake in Myanmar. Although its main purpose is to illustrate our methodology, the results obtained nevertheless hold some empirical relevance. While those residents face serious environmental problems, such as water quality degradation, there are no previous studies to estimate the preference for a water improvement policy. Moreover, our non-parametric approach is especially relevant, as there is insufficient empirical evidence that provides prior knowledge on the preference structure to justify any specific assumption concerning the form of the utility function, with such studies in these disadvantaged areas in the developing hemisphere being extremely limited. Our estimation results show that the lower bound of the average marginal WTP for the representative water improvement policy is already as high as 5.9% of respondents' income, on average. Moreover, the estimated WTP distribution reveals heterogeneous preferences for the policy across settlements. Regarding the causal effect of each individual attribute, we find that the toilet provision and joint implementation with an NGO are two important attributes that positively influence WTP.

The remainder of the paper is organized as follows. In the next section, we present our non-parametric method for WTP identification after an overview of the literature on WTP estimation. Section 3 applies the methodology to the empirical case of water quality improvement policy for Inlay Lake, where we discuss the current situation and potential policy as empirical background; this is followed by the survey design, estimation strategy and results. Finally, Section 4 concludes the paper.

2 Non-Parametric WTP Estimation Methodology

This section shows the new approach to draw out the welfare implications directly from the data generated by the HHY conjoint experiment (Hainmueller, et al. 2014), without relying on the strong assumptions that conventional methods such as those based on random utility models require. As mentioned above, the major strength of HHY conjoint is that it can non-parametrically elicit the causal effect of individual attributes on the choice outcome, which conforms to Rubin's potential outcome framework. It is hence considered more robust than other discrete choice experiment methods that specify the decision-making process using the random utility model. Extending the literature on WTP measurement without sacrificing the novel strength of the HHY conjoint approach is the primary purpose of this paper. We begin this section by reviewing the methodological developments of WTP measurement and then present the non-parametric method of WTP identification.

2.1 Methodological Background

WTP is a measure of social welfare, which is often used in the field of social science including political science (among the recent works, see for example Bodea and LeBas 2016 and Fairbrother 2017) as well as policy evalution and policy-making decisions (see for example Arrow et, al., 1993 and Alberini and Cooper 2000 while Competition Commission 2010 reviews methodology for the practitioners). For instance, in the typical cost-benefit analysis, a policy implementation is justified if the average WTP is larger than the its average costs.

Two main WTP measurement approaches are revealed and stated preferences. While the revealed preference approach observes actual decisions made by individuals on the realized policy, the stated preference approach collects data on decision-making behavior in a hypothetical situation. The revealed preference approach is often argued to be more valid than the stated preference approach, as it is based on actual behavior. When a policy has yet to be implemented, however, we cannot observe the actual behavior that can identify people's preferences regarding the policy. The preferences for a multi-attribute policy that can potentially have different levels in each attribute are especially difficult to identify with the revealed preference approach.

There are two popular approaches when using stated preferences: the contingent valuation method (CVM) and discrete choice experiments (DCEs) (Carson and Louviere 2011 provide a nomenclature for stated preference approaches). In the CVM, respondents are directly asked about their WTP to implement a hypothetical policy.[6] The CVM is straightforward and became the dominant stated-preference approach for measuring WTP in the 1990s. However, the CVM is not appropriate for estimating preferences over different multi-attribute policy alternatives because hypothetical policies presented to respondents have common attribute levels in a typical CVM survey design.

In DCEs (including conjoint experiments), WTP is estimated based on hypothetical choice experiments in which respondents are asked about their preferred choice alternative with multiple attributes. A main advantage of DCEs is that they allow the researcher to estimate the effect of changes in individual policy attributes on the WTP distribution, which may be more informative from a practical policy perspective (Hanley et al., 2001). Therefore, DCEs have become increasingly popular, especially since 2000 (see Birol and Koundouri 2008; Carson and Czajkowski 2014). In DCEs, WTP is estimated based on a common theoretical decision model known as the random utility model developed by Lancaster (1966)

[6] CVM is proposed by Bowen (1943), and Ciriacy-Wantrup (1947) and Davis (1963) is the first empirical application. Carson and Hanemann (2005) and Whittington, Adamowicz, and Lloyd-Smith (2017) survey recent discussions.

and McFadden (1974). More recently, Wiswall and Zafar (2016) develop an approach to estimate the distribution of WTP without assuming any restricted forms of the coefficient distribution.[7] Although methodological advances have greatly improved the validity of estimation results, these decision models retain some untestable assumptions: the additive separability of utility functions and error terms that follow the type-I extreme-value distribution are still required. Insofar as these decision models describe the true decision-making process, statistical relationships between choices and attributes are considered causal, and the WTP can be recovered. However, these models are typically expressed by a highly parametric specification with a particular functional form of the utility function, which is a problematic assumption, particularly in the context of development studies where a priori information about preferences is limited, or even non-existent.

To reduce the likely bias due to a miss-specified decision-making process, recent studies propose a WTP estimation approach based on non-parametric decision models, which is referred to as non-parametric welfare analysis (Hausman and Newey 2017). Bhattacharya (2015) reports point identification results for WTP (and consumer surplus) of a price change in the discrete choice setting.[8] In this approach, the preference distribution is not parametrically specified; instead, only weak assumptions are required. We adapt the idea of non-parametric welfare analysis developed by Bhattacharya to the WTP estimation of policy implementation by introducing multi-attribute alternatives into his model and report the identification result of the MCE-WTP.

Moreover, our strategy for identifying causal effects on welfare is in line with the sufficient statistics approach in public economics (Chetty 2009). The sufficient statistics approach is a middle ground between design-based (reduced-form) and model-based (structural estimation) approaches. In this approach, the theoretical model formally connects the causal welfare effect to estimable causal effects, and those causal effects are estimated using the design-based approach. We also first show the theoretical relationship between the MCE-WTP and the average marginal component effect and then estimate both quantities using the strategy as HHY conjoint experiment in Hainmueller et al. (2014).

2.2 Conceptual Framework of Non-Parametric Welfare Analysis

This section provides a general framework with a non-parametric choice model as the randomized conjoint experiment, which allows us to obtain the welfare implications of a policy from the choice data. A key

[7]The cost of relaxing the assumption on the coefficient distribution is increased cognitive burden for the respondents, due to asking about their "belief" in the long-run in terms of probability.

[8]Hausman and Newey (2016) demonstrate a set identification result in a continuous choice setting.

statistical quantity is the approval rating, which indicates that a policy is preferred to the status quo and can then be used to recover the distribution of individual WTP without any strong assumptions on the preference distribution.

2.2.1 Potential outcome framework

Before introducing the identification of WTP, we begin by reviewing the discussion of the potential outcome framework of HHY conjoint experiment. In each of K choice tasks presented to respondent i, she/he chooses a preference ranking among J hypothetical policies and the status quo. An individual choice indicator, denoted by, say, Y_{ijk}, takes value one if an individual i prefers to implement jth policy in the kth choice task instead of the status quo, zero otherwise. Additionally, the policy burden of a policy C_{ijk} and other non-pecuniary attributes is $\boldsymbol{A}_{ijk} = \{A_{jk1}, ..., A_{jkL}\}$, where L is the number of attributes, are also observed.

Let $y_i(c, \boldsymbol{a})$ be a potential outcome of a binary choice indicator for a policy $\{c, \boldsymbol{a}\}$ where c is the individual burden to implement the policy, and $\boldsymbol{a}$ is a vector of attributes. The potential outcome $y_i(c, \boldsymbol{a})$ is equal to one if respondent i's stated preference rank of the policy is higher than that of the status quo, while $y_i(c, \boldsymbol{a})$ is equal to zero if the preference rank is lower than that of the status quo. Let us define the potential policy approval rating as

$$q(c, \boldsymbol{a}) = E[y_i(c, \boldsymbol{a})],$$

which is the share of individuals who prefer policy $\{c, \boldsymbol{a}\}$ to the status quo.

The *average conditional component effect* (ACCE) of an attribute l is defined by

$$\pi_l(a_1, a_0 | c, \boldsymbol{a}_{-l}) \equiv q(c, a_1, \boldsymbol{a}_{-l}) - q(c, a_0, \boldsymbol{a}_{-l}), \tag{1}$$

where $\boldsymbol{a}_{-l}$ is a vector of non-pecuniary attributes excluding attribute l. The ACCE can be interpreted as the increase in the policy approval rating if the value of an attribute l is changed from a_0 to a_1 given the values of the individual burden and other attributes.

Here we assume *Independence* and *Randomization* as follows.

Independence *We assume that an individual choice indicator Y_{ijk} depends only on respondent i's pref-*

erence and the attributes of the policy j. Formally, for any i, j, k, k',

$$Y_{ijk} = Y_{ijk'} \text{ if } \{C_{ijk}, \boldsymbol{A}_{ijk}\} = \{C_{ijk'}, \boldsymbol{A}_{ijk'}\}$$

Independence requires that the preference indicator between a policy and the staus quo must take same value if the attributes of the policy are identical. Under *Independence,* the round of the choice task and the order of the policy has no impact on the choice indicator; Y_{ijk}, C_{ijk} and A_{ijk} can then simply be preferences denoted as Y_{ij}, C_{ij} and A_{ij}.

Randomization $\{C_{ij}, \boldsymbol{A}_{ij}\}$ is randomized, as (i) $\{y_i(c, \boldsymbol{a})\}_{c,\boldsymbol{a}}$ is statistically independent of $\{C_{ij}, \boldsymbol{A}_{ij}\}$, (ii) C_{ij} is independent of $\left\{\boldsymbol{A}_{ij}, \boldsymbol{A}_{i[-j]}\right\}$, A_{ijl} is independent of $\left\{C_{ij}, \boldsymbol{A}_{ij[-l]}\right\}$, and (iii) it is fully supported such that $0 < p(c, \boldsymbol{a}) \equiv \Pr(C_{ij} = c, \boldsymbol{A}_{ij} = \boldsymbol{a})$ for all $c \in \Phi_C$ and $\boldsymbol{a} \in \Phi_{\boldsymbol{a}}$, where Φ_C and $\Phi_{\boldsymbol{a}}$ are the supports of c and $\boldsymbol{a}$ in the conjoint experiment.

Randomization stipulates the assignment rule of policy attributes and allows us to identify the approval probability and the ACCE as

$$\hat{q}(c, \boldsymbol{a}) = \overline{Y}_{ij}|_{C_{ij}=c, \boldsymbol{A}_{ij}=\boldsymbol{a}}, \tag{2}$$

and

$$\hat{\pi}_l(a_1, a_0|c, \boldsymbol{a}_{-l}) = \overline{Y}_{ij}|_{C_{ij}=c, A_{ijl}=a_1, \boldsymbol{A}_{ij[-l]}=\boldsymbol{a}_{-l}} - \overline{Y}_{ij}|_{C_{ij}=c, A_{ijl}=a_0, \boldsymbol{A}_{ij[-l]}=\boldsymbol{a}_{-l}}, \tag{3}$$

respectively, where $\overline{Y}_{ij}|_{C_{ij}=c, \boldsymbol{A}_{ij}=\boldsymbol{a}}$ is a conditional sub-sample average of the observed choice indicator, and $\boldsymbol{A}_{ij[-l]}$ is a vector of non-pecuniary attributes that excludes the l-th attribute.

The identification of ACCE in (2) and (3) is difficult when the number of attributes and their levels is large, as the numbers of observations that belong to the conditioning set on the right-hand side of the equations become small. HHY conjoint experiment then introduce a marginalized causal quantity, the *average marginal component effect* (AMCE). Under *Randomization*, the AMCE and its identification become

$$\begin{aligned}
\pi_l\left(a_1, a_0\right) &= \sum_{c, \boldsymbol{a_{-l}}} \pi_l(a_1, a_0|c, \boldsymbol{a}_{-l}) \times p\left(c, \boldsymbol{a}_{-l}\right) \\
\hat{\pi}_l\left(a_1, a_0\right) &= \overline{Y}_{ij}|_{C_{ij}=c, A_{ijl}=a_1} - \overline{Y}_{ij}|_{C_{ij}=c, A_{ijl}=a_0}
\end{aligned}$$

where $p\left(c, \boldsymbol{a}_{-l}\right)$ is a joint probability distribution of the individual burden and other attributes $\boldsymbol{a}_{-l}$. The AMCE can be interpreted as the increase in the approval rating if the value of an attribute l is changed from a_0 to a_1, averaged over all possible values of the individual burden and other attributes given their joint distribution.

2.2.2 Rational choice framework

We next define and identify the component effect on the WTP distribution based on a non-parametric choice model. The following discussion reveals that the identified ACCE equation in (3) is a sufficient statistic to recover the component effect on the WTP distribution.

Preference The utility of an individual i in the absence of policy, termed the status quo utility, is denoted by v_i. The utility of an individual i under a policy $\{c, \boldsymbol{a}\}$ is denoted by $u_i(c, \boldsymbol{a})$. We assume that the utility function has the following properties.

Monotonicity For any $\boldsymbol{a} \in \Phi_{\boldsymbol{a}}$, $u_i(c, \boldsymbol{a})$ is strictly monotonically decreasing in $c \in R_+$.

Continuity For any $\boldsymbol{a} \in \Phi_{\boldsymbol{a}}$, $u_i(c, \boldsymbol{a})$ is a continuous function of $c \in R_+$.

Boundary For any $\boldsymbol{a} \in \Phi_{\boldsymbol{a}}$, $\lim_{c \to \infty} u_i(c, \boldsymbol{a}) < v_i$, and $u_i(0, \boldsymbol{a}) \geq v_i$.

Rationality For any $c \in \Phi_c$ and $\boldsymbol{a} \in \Phi_{\boldsymbol{a}}$, $u_i(c, \boldsymbol{a}) \geq v_i$ if and only if $y_i(c, \boldsymbol{a}) = 1$.

Monotonicity implies that no individuals prefer a higher individual policy burden, and *Continuity* is a technical assumption to ensure the existence of the WTP. *Boundary* requires that (i) individuals do not prefer to implement a policy with infinitely high cost and (ii) must prefer to implement a policy with zero burden[9]. This assumption stems from the fact that all levels of attributes offered to the respondents will only improve the situation relative to the status quo and is necessary to ensure non-negativity. Finally, *Rationality* requires consistency between real and stated preferences in the conjoint experiment. The

[9]Although it is not noted, we abbreviate the term "weakly" here and below.

assumption allows us to connect the approval rating with preferences as follows:

$$q(c, \boldsymbol{a}) = \Pr[u_i(c, \boldsymbol{a}) \geq v_i]. \tag{4}$$

Equation (4) shows that the approval rating or, equivalently, the share of individuals who prefer to implement a policy rather than to maintain the status quo is the probability that an individual prefers the policy over status quo.

One advantage of this approach is that the four rather unrestrictive assumptions above are all that we need to identify the welfare impact. In other words, no parametric assumptions on the utility function are needed, and the result is then free from bias due to model specification error. Traditional DCEs elicit WTP based on the random utility model (See Train 2009) that, unlike our approach, assumes additively separable utility to allow for preference heterogeneity [10].

2.2.3 WTP distribution

Let us denote by $\Omega_i(\boldsymbol{a})$ an individual's WTP, under which an individual i is indifferent between implementing a policy $\{\Omega_i(\boldsymbol{a}), \boldsymbol{a}\}$ and the status quo. Formally, the individual WTP is defined as

$$u_i(\Omega_i(\boldsymbol{a}), \boldsymbol{a}) = v_i.$$

WTP depends on the policy attributes $\boldsymbol{a}$, and *Monotonicity*, *Continuity*, and *Boundary* ensure the existence of non-negative WTP for any $\boldsymbol{a}$.

Individual WTP will vary across individuals due to the heterogeneity in preferences and cannot be

[10] Typically, it is specified as

$$u_i = \alpha_{ci} \times I_c + \sum \alpha_{li} \times I_l + \epsilon_{ij},$$

where I_{cj} and I_{lij} are vectors of dummy variables for each level of the financial burden and attribute l, respectively, while α_{ci} and α_{li} are vectors of the corresponding coefficients, and ϵ_{ij} captures the preference for unobservable attributes of policy j. The values of α_{ci} and α_{li} are different across individuals and are assumed to be normally distributed.

estimated, but its distribution is estimable. We define the conditional WTP distribution as

$$F(X|\boldsymbol{a}) = \Pr[\Omega_i(\boldsymbol{a}) \leq X],$$

which represents the share of individuals whose WTP is equal to X or lower, given other attributes. *Monotonicity* then entails that $v_i = u_i(\Omega_i(\boldsymbol{a}), \boldsymbol{a}) \geq u_i(X, \boldsymbol{a})$ if and only if $X \geq \Omega_i(\boldsymbol{a})$ and, thus, rewrites the above as

$$F(X|\boldsymbol{a}) = \Pr[v_i \geq u_i(X, \boldsymbol{a})]. \tag{5}$$

This implies that the share of individuals whose WTP is no more than X is equal to the share of individuals who do not prefer to implement the policy $\{X, \boldsymbol{a}\}$. Combining (4) and (5) yields

$$F(X|\boldsymbol{a}) = 1 - q(X, \boldsymbol{a}),$$

which shows that the approval rating is sufficient to recover the WTP distribution. Incorporating (2) into the above equation obtains an identification result of the WTP distribution, say $\hat{F}$, as

$$\hat{F}(X|\boldsymbol{a}) = 1 - \overline{Y}_{ij}|_{C_{ij}=X, \boldsymbol{A}_{ij}=\boldsymbol{a}}. \tag{6}$$

The WTP distribution can then be identified solely by the conditional sub-sample average of the observed choice indicator, however, it again faces the sample-size problem.

A more practical welfare quantity is the marginal WTP distribution, which is defined as

$$F(X) = \sum_{\boldsymbol{a}} F(X|\boldsymbol{a}) \times p(\boldsymbol{a}), \tag{7}$$

where $p(\boldsymbol{a})$ is a joint probability distribution of attributes. The marginal WTP distribution is interpreted as the share of individuals having WTP of X or lower, averaged over all the possible values of attributes given their joint probability distribution $p(\boldsymbol{a})$. Combining this with (6) and (7) yields an identification result:

$$\hat{F}(X) = 1 - \overline{Y}_{ij}|_{C_{ij}=X}. \tag{8}$$

The marginal WTP distribution identifies the welfare implications of a "representative" policy. Specifically, by allowing for heterogeneous preferences, our approach can obtain the distributional implication of WTP without imposing any restriction on the WTP distribution. In fact, our conjoint analysis reveals significant heterogeneity in WTP (see Section 5).

Here, let us define $F^{-1}(F)$ as the inverse function of $F(X)$. Then, the average WTP is defined as $\mu = \int_0^1 F^{-1}(\phi)\, d\phi$, which we can also identify under the following assumption.

Full support The individual burden has a continuous distribution over the positive real space, or equivalently, $\Phi_c = R_+$.

Under *Full support*, we can identify $E[Y_{ij}|C_{ij} = X]$ for any $X \in R_+$, and the average marginal WTP is then point-identified as

$$\hat{\mu} = \int_0^1 \hat{F}^{-1}(\phi)\, d\phi. \tag{9}$$

This is also an important welfare quantity because the compensation principle argues that the implementation of a policy is justified if the average costs of such a policy are lower than its benefits.

2.2.4 Component effect on WTP distribution

While the marginal WTP distribution given in (8) yields the "representative" WTP distribution, the component effect on the WTP distribution evaluates the causal effect of each attribute on welfare. First, the *conditional component effect on the WTP distribution* (CCE-WTP) is defined as

$$\Pi_l(X|a_1, a_0, \boldsymbol{a}_{-l}) = F(X|a_1, \boldsymbol{a}_{-l}) - F(X|a_0, \boldsymbol{a}_{-l}),$$

where $\boldsymbol{a}_{-l}$ is a vector of attributes excluding the l-th attribute. The CCE-WTP can be interpreted as the increase in the share of WTP less than X if the value of an attribute l is changed, say from a_0 to a_1, given other attributes $\boldsymbol{a}_{-l}$.

The identification result of the conditional WTP distribution in (6) allows us to identify the CCE-WTP as

$$\hat{\Pi}_l(X|a_1, a_0, \boldsymbol{a}_{-l}) = \overline{Y}_{ij}|_{C_{ij}=X, A_{lij}=a_0, \boldsymbol{A}_{-lij}=\boldsymbol{a}_{-l}} - \overline{Y}_{ij}|_{C_{ij}=X, A_{lij}=a_1, \boldsymbol{A}_{-lij}=\boldsymbol{a}_{-l}}, \tag{10}$$

because $F(X|a, \boldsymbol{a}_{-l}) = 1 - E\left[Y_{ij}|C_{ij} = X, A_{ijl} = a, \boldsymbol{A}_{ij[-l]} = \boldsymbol{a}_{-l}\right]$. The above equation shows that $\hat{\Pi}_l\left(X|a_1, a_0, \boldsymbol{a}_{-l}\right)$ is simply identified by the difference-in-means estimators, but its use in statistical inference is also difficult due to sample size problems.

Similar to the above, the *marginal component effect on the WTP distribution* (MCE-WTP) is defined as

$$\Pi_l(X|a_1, a_0) = \sum_{\boldsymbol{a}_{-l}} \Pi_l\left(X|a_1, a_0, \boldsymbol{a}_{-l}\right) \times p(\boldsymbol{a}_{-l}).$$

The MCE-WTP can be interpreted as the increase in the share of WTP less than X if the value of an attribute l is changed from a_0 to a_1, averaged over all the possible values of other attributes given their joint distribution $p\left(\boldsymbol{a}_{-l}\right)$. From equation (10), the MCE-WTP is then empirically identified as

$$\hat{\Pi}_l(X|a_1, a_0) = \overline{Y}_{ij}|_{C_{ij}=X, A_{ijl}=a_0} - \overline{Y}_{ij}|_{C_{ij}=X, A_{ijl}=a_1}. \tag{11}$$

Equation (11) shows that the MCE-WTP is estimated by the average value of Y_{ij} conditional only on the individual burden and the level of the attribute of interest.

2.2.5 Partial identification: a bounds analysis

An important note is that *Full support* is often not well observed because the proposed policy burden in the conjoint experiment may have only a discrete variation. For instance, as shown below in the empirical application, our conjoint data provide only estimators of the approval rating at four points, namely, $C_{ij} = 5$, 10, 15, and 20 thousand kyats. This section presents the bounds estimation methods, which can be applied for those cases with discrete variation in the proposed policy burden.

We rewrite the average WTP as

$$\mu = \sum_{i=0}^{k} \int_{F(c_i)}^{F(c_{i+1})} F^{-1}\left(\phi\right) d\phi$$

where c_i is the ith threshold value for c, k is the number of threshold values, and $c_0 = 0$ and $c_{k+1} = \infty$ such that $F\left(c_0\right) = 0$ and $F\left(c_{k+1}\right) = 1$. Using the mean value theorem, the above equation can be further rewritten as

$$\mu = \sum_{i=0}^{k} \tilde{c}_i \left[F\left(c_{i+1}\right) - F\left(c_i\right)\right]$$

for some $\tilde{c}_i \in [c_i, c_{i+1}]$ for any $i = 0, \ldots, k$. Since F is monotonically increasing in c, the terms in the square brackets are jointly positive, and therefore, letting $\tilde{c}_i$ simply be c_i yields the lower bound of the average marginal welfare gain, say $\underline{\mu}$, as

$$\underline{\mu} = \sum_{i=0}^{k} c_i \left[F\left(c_{i+1}\right) - F\left(c_i\right) \right].$$

Similarly, the lower bound of the conditional average welfare gain $\left.\underline{\mu}\right|_{A_l=a_l}$ is also obtained as

$$\left.\underline{\mu}\right|_{A_l=a_l} = \sum_{i=0}^{k} c_i \left[F\left(\left. c_{i+1} \right| A_l = a_l \right) - F\left(\left. c_i \right| A_l = a_l \right) \right].$$

We can thus identify these statistics in the empirical application as follows:

$$\begin{aligned} \hat{\underline{\mu}} &= \sum_{i=0}^{k} c_i \left[\hat{F}\left(c_{i+1}\right) - \hat{F}\left(c_i\right) \right] \\ \left.\hat{\underline{\mu}}\right|_{A_l=a_l} &= \sum_{i=0}^{k} c_i \left[\hat{F}\left(\left. c_{i+1} \right| A_l = a_l \right) - \hat{F}\left(\left. c_i \right| A_l = a_l \right) \right], \end{aligned}$$

which we obtain using the point estimate of the WTP distribution at each c_i. In the following section, we illustrate this identification of the average welfare gains for the case of water quality improvement policy for Inlay lake.

3 Empirical Application

3.1 Empirical Motivation: Water Quality of Inlay Lake

We address the case of water quality improvement policy for Inlay lake in Myanmar as an empirical application of our method. Inlay Lake is the second-largest wetland in Myanmar (Su and Jassby 2000) and is located in Nyaung Shwe Township, Taunggyi District of Southern Shan State. The lake is well known by domestic and international tourists for its rich cultural heritage and biological diversity. For example, the Inlay Lake Wildlife Sanctuary, established in 1985, became an ASEAN Heritage Park in 2003 and a part of the World Network of Biosphere Reserves of UNESCO in 2015.

The population of Nyaung Shwe Township increased steadily from 77,000 to 189,000 between 1973 and 2014, with an average annual growth rate of approximately 2.2% (Su and Jassby 2000, Ministry of Immigration and Population 2015). This long-term steady demographic pressure combined with intensifying economic activities on and around the lake have resulted in water quality degradation (*i.e.*, eutrophication) of the lake. Our findings indicate that, as part of the population of Nyaung Shwe Township, there are 17 floating settlements comprising 2,284 households and 13,794 people living on the lake at the time of our survey in 2015. These residents intensively exploit environmental services, and thus, their livelihoods are highly dependent on the quality and quantity of the lake water. However, their livelihood activities seriously pollute the lake water.

For example, as one of the main economic activities, tomato farming on floating gardens represents two-thirds of regional agricultural production (Butkus and Myint 2001).[11] Consequently, 32.4% of the open surface water area, or 46.7 km^2, was lost between 1935 and 2000 due partly to the development of the floating gardens (Sidle *et al.* 2007). More than 15 years ago, Butkus and Myint (2001) reported that the use of chemical fertilizers and pesticides to improve the productivity of tomato farming had already reached levels of overuse. Therefore, the excessive use of fertilizers and pesticides for tomato production is considered one of the causes of eutrophication, with another being human waste disposal. Although scientific evidence through regular monitoring and/or ad hoc research is minimal, Akaishi *et al.* (2006) found that the concentrations of PO4-P, NO2-N, and NO3-N are relatively high in Inlay Lake's water. They also found *E.coli* or coliform bacteria in the surface water of the lake, which can cause diarrhea if the water is not treated before drinking. While most houses dispose of excretions directly into the lake, some village people still use lake water as drinking water.

To improve the water quality and the environment of the lake, various stakeholders, including state government departments, the local government, international organizations (*e.g.*, donor agencies and the United Nations), and international NGOs, have pursued conservation through programs and projects. Moreover, local NGOs known as community-based organizations (CBOs) have been established. There are currently approximately 20 local CBOs working on community development activities, including en-

[11] Floating gardens are large blocks of organic-rich soil brought from the wetlands around the lake (Sidle et al. 2007). These must cope with decayed grasses, reeds, marsh plants, and aquatic plants excavated from the lake bottom.

vironmental conservation in and around Inlay Lake. However, there has been little improvement thus far in the quality of the lake water (Norwegian Government and UN-Habitat 2014). This is because most activities are public awareness and training campaigns, which are small and independent, and these are occasionally joined by government efforts to promote environmental conservation of the lake. Therefore, large-scale and comprehensive countermeasures are required, including improvements in sanitary and environmental infrastructure and regulations.

In this study, we focus on the floating settlements on the lake, as their livelihoods are extremely lake water-dependent. Anecdotal evidence indicates widespread recognition that the quantity and quality of the lake water have deteriorated over the last decade, which has increased the awareness and concern for locals' livelihoods. Although individual sources of income vary across the floating settlements, such as traditional fishing, tomato production on the floating gardens, and tourism (*e.g.*, restaurants and hotels, boat drivers, sales and production of handcrafts, textile dyeing and silver manufacturing), most of them directly depend on, or are affected by, the lake water's quality and quantity. Moreover, compared with their income sources, the livelihoods of the residents depend more homogeneously on the lake water. For example, boats, with and without engines, are the only means of transportation, even when visiting a neighboring house. In addition, the residents still use lake water for daily activities such as bathing and washing their clothes and dishes. However, the recent pollution of the lake water has meant that drinking water and water for cooking have to be secured from alternative sources such as pipes from spring water reservoirs or groundwater wells from the surrounding land, delivery tanks or rainwater. The source of water depends on people's location and wealth. From our survey, we found that only 11.62% of the residents in our sample area drink the lake water directly, although most of the residents drank the lake water ten years ago. The lake is also a sink for human waste and garbage, although some residents take their garbage to surrounding land areas for burning.

It is thus highly likely that most residents of the floating settlements are concerned about the quality of the lake water; however, scientific evidence on the subject is absent. Therefore, an important empirical task here is to investigate whether the residents of floating settlements' WTP for lake water quality improvements exceeds the amount that justifies policy intervention. Important evidence on the WTP for water quality improvements in different areas using DCEs is provided in many studies, such as Smith and Desvousges (1986), Boatable (1993), and Huber, Viscusi, and Bell (2000). More recently, Bell, Huber, and Viscusi (2009) and Viscusi, Huber, and Bell (2008, 2012) also report the WTP for water quality improvements. However, most empirical applications are concentrated in developed counties, where a certain amount of prior knowledge in the literature allows for reasonable expectations and assumptions about preferences, and thus, the use of DCEs with random utility models may be justifiable. The methodology proposed in this study is believed to be particularly relevant for the cases in which the prior knowledge and information in the literature are insufficient to impose behavioral assumptions. Thus, the welfare analysis of water quality improvements for floating settlements on Inlay Lake in Myanmar is our empirical objective in this study.

3.2 Survey Design

Scenario

The water flowing into Inlay Lake through several river channels varies in terms of volume and quality. When there is heavy rain in the upstream watershed, some rivers supply fresh water, while others supply high-sediment water, which have different impacts on the quality of the lake water. Moreover, the spatial distribution of floating houses and gardens, as sources of pollutants, contributes to the substantial variation in the quality of the water in different locations. Despite this hydrological complexity, there is insufficient information on the changes and variations in water quality to be able to generalize water quality over time and space. Therefore, the scenario described in the experimental survey employs a narrative of water quality improvement that requires no scientific information on the part of the respondents. The following is the scenario given to the respondents before conducting the conjoint experiments.

"We would like to propose several different public policy programs for improving the water quality of Inlay Lake. We assume that all of the proposed programs will provided the same level of water quality improvement, that is, the improvement in quality will be sufficient to ensure that water from anywhere in the lake can always be used for cooking but may not be good for drinking. It should be also noted that any money collected is fully and properly used for achieving the aforementioned water quality improvement goal. The project is primarily implemented by the local government using this collected money" Since the status quo of the water quality varies across respondents, the concept of improvement also varies. However, these heterogeneous variations across respondents are captured by individual preference parameters.

Attributes

Each alternative is a proposed policy program and is characterized by seven attributes. The first is a toilet system (*TOILET*), with three levels. In the first level, the local government provides a toilet and a collection tank for each house. The second level adds a collection service by the government to the services provided at the first level. The third level assumes that no toilet system is provided. The second attribute is a garbage collection service (*GARBAGE*), with two levels: once-a-week collection and no service. In the pilot survey, villagers were rather indifferent among collection frequencies, owing to a lack of familiarity with the service, although they are eager to have the service itself. Thus, the first level is once a week, being the most-preferred option in the pilot survey, and the second level is no garbage collection service. The third attribute is a collective public wastewater treatment facility for dyeing and silver gilt (*WASTEWATER*), introduced and operated by the government, with two levels: with and without the service. The original scenario in the pilot survey had three levels, differentiating between a mandatory connection and a voluntary connection. However, there was no significant difference between the two connection types in the pilot survey. The fourth attribute, consisting of three levels, is government regulation and guidance on optimal fertilizer and pesticide inputs for tomato production (*FERTILIZER*). Its first level assumes that government services are provided only for fertilizer input, the second level assumes that they are only provided for pesticide inputs, while the third level provides for both.

In addition to the first four attributes defining different services offered by the local government, three more attributes specify the characteristics of the proposed policy. The first is the project period (*PERIOD*), with three levels that specify the duration wherein the predetermined common narrative target is achieved, being 5, 10, and 20 years. The next attribute is the implementing organizations (*ORGANIZATION*), which has two levels: the government alone and the government in conjunction with local NGOs. In the pilot survey, villagers revealed a strong preference for joint implementation involving the government and local NGOs. As mentioned above, in the study area, third parties such as NGOs and CBOs play an active role in many public policy programs, augmenting government services.

The last attribute is payment (*PAYMENT*), which has three levels. The selection of payment vehicles and the range of levels were examined carefully in the pilot survey. The pilot survey used three levels of monthly payments during the given project period, from 3,000 to 15,000 kyats, where 3,000 kyats is the hourly wage for unskilled labor in the area. The result of the pilot survey revealed that even 15,000 kyats would not strongly discourage the willingness to support the policy programs; thus, we modified the range to 5,000 kyats to 20,000 kyats, in four levels. Note that cash donation to Buddhist temples is common practice in the study communities. Thus, we suppose that credit liquidity is sufficient to employ cash payments in this study. Moreover, our sample shows that the average household donation per month is slightly more than 20,000 kyats, which accounts for 6.7% of household income.

In our conjoint experiment, each choice set has two policy alternatives in addition to the status quo that are then ranked[12]. The seven attributes and their respective levels yield 864 policy alternatives in total, two of which are randomly paired[13]. Each respondent is required to make a ranking decision three times for three different choice sets [14].

Sampling and Survey

In collaboration with the local government and a local NGO, we successfully collected an up-to-date list of settlements and all names of the household heads of Nyaung Shwe Township. We define floating settlements (villages) as those where all residents reside above the lake, and identify 17 floating settlements comprising 2,284 households and 13,794 people, as mentioned above. From this total population, we randomly select

[12] Making a ranked choice imposes a greater burden on respondents than making a dichotomous choice; however, we confirmed in our pilot survey that respondents clearly understood the scenario and survey design.

[13] The order of attributes given in the choice set is fixed, despite the literature suggesting that the order be changed randomly. However, because this causes confusion and difficulty for respondents, we prioritized reducing the burden of repeated choice decisions.

[14] In the main survey, we used printed copies of choice sets. In the pilot survey, we used a computer. The reason for the change is that computer screens occasionally made it difficult to communicate with respondents in their dark houses. For practical reasons, due to cost and time constraints, we prepared 144 versions of printed copies consisting of randomly defined 3-choice sets, which were also randomly assigned to each of the 327 respondents.

327 households.[15] Since there is relatively substantial variation in population size across settlements, we apply stratified random sampling by settlement, with sampling rates of between 13.5% and 15.0% across the 17 settlements. One of authors visited and met all selected households and conducted the survey within one month. No respondents were missing from the original list of selected households. The other co-authors participated either in the preliminary survey or the pilot survey.[16] Considering the various implementation challenges in developing countries, note that this ideal random sampling at the household level and the complete survey should minimize any sampling bias.

The data collected from the main survey can be divided into two categories. The first category is the choice preferences for the proposed policy programs. Here, we can further derive our main outcome variable from the ranking information in the data, namely a choice indicator, with the appropriate coding. The choice indicator compares the ranking between the status quo and the other two alternatives. Here, 1 is assigned to any policy alternative with a higher ranking than that of the status quo, and 0 is assigned otherwise. Note that if the ranking of the status quo is the highest or the lowest, both of the alternative policy programs receive a 0 or 1.

The second category consists of household survey data, comprising six subcategories: (1) a household roster, which identifies basic information (e.g., ethnicity, religion, age, gender, education, and job) and the mutual relations of all family members; (2) sources of household income for the last 12 months, with possible inputs including wage and working hours, production quantities, selling prices, necessary costs, and so forth; (3) living conditions, including water supply, toilet, garbage disposal, and health conditions; (4) possession and purchasing history of major durable goods; (5) financial conditions, such as debt, savings, and donations for the previous 12 months; and (6) environmental awareness, which includes perceptions of recent changes in the quality of the lake water, as well as three major causes from a given list and the respondent's level of understanding of the scenario used in the conjoint experiment.

The characteristics of our data sampled from the 327 households are as follows. Slightly more than half

[15]Khin (2011) notes that there are 35 floating settlements with 20,000 people. However, half of these settlements include residents whose houses are not floating on the lake. We target only those settlements where all residents resided on the lake at the time of our survey.

[16]The preliminary survey was conducted for six days from October 8 to 13, 2014, to collect 30 samples, mainly by means of interviews. The pilot survey was conducted for 10 days from October 7 to 16, 2015, collecting 50 samples for the conjoint field experiment. The main survey was conducted between December 29, 2015, and January 25, 2016.

of all respondents are female (51.38%), and the average age is 44.7 years. Nearly 60% of the household heads completed up to a primary level of education (59.33%), 21.10% completed up to middle school, and 3.98% and 1.83% are high school and college/university graduates, respectively, whereas 13.76% are illiterate. Family size ranges from 1 to 11, with an average size of 4.7. The total annual household income ranges from 360,000 kyats (300 USD) to 21.6 million kyats (18,000 USD), with an average of 3.6 million kyats (3,000 USD). Moreover, the average per capita income is 840,000 kyats (700 USD), with a range of between 158,000 kyats (131 USD) and 5.4 million kyats (4,500 USD). Further details on the characteristics of sample households can be found in the Appendix.

3.3 Estimation Strategies

3.3.1 Bounds estimation

Our conjoint data can provide only estimators of the approval ratings at $C_{ij} = 5000,\ 10,000,\ 15,000,$ and $20,000$ kyats, which makes it difficult to fully recover the WTP distribution. We then estimate the lower bound of the average WTP. First, the identification result in equation (8) allows us to estimate five probability intervals of the marginal WTP distribution as

$$\hat{F}(5,000) = 1 - \overline{Y}_{ij}|_{C_{ij}=5000},$$

$$\hat{F}(10,000) - \hat{F}(5,000) = \overline{Y}_{ij}|_{C_{ij}=5000} - \overline{Y}_{ij}|_{C_{ij}=10000},$$

$$\hat{F}(15,000) - \hat{F}(10,000) = \overline{Y}_{ij}|_{C_{ij}=10000} - \overline{Y}_{ij}|_{C_{ij}=15000},$$

$$\hat{F}(20,000) - \hat{F}(15,000) = \overline{Y}_{ij}|_{C_{ij}=15000} - \overline{Y}_{ij}|_{C_{ij}=20000},$$

$$1 - \hat{F}(20,000) = \overline{Y}_{ij}|_{C_{ij}=20000}. \qquad (12)$$

Equation (12) means that the "histogram" of the WTP distribution can be estimated from the conjoint experiment.

Next, the average marginal WTP can be rewritten in our context as

$$\mu = \int_{F(0)}^{F(5{,}000)} F^{-1}(\phi)\, d\phi + \int_{F(5{,}000)}^{F(10{,}000)} F^{-1}(\phi)\, d\phi + \int_{F(15{,}000)}^{F(20{,}000)} F^{-1}(\phi)\, d\phi + \int_{F(20{,}000)}^{F(\infty)} F^{-1}(\phi)\, d\phi.$$

From the above, the average marginal WTP is further modified as

$$\mu = \tilde{X}_{0,5000} F(5000) + \tilde{X}_{5000,10000} \left[F(10000) - F(5000)\right]$$

$$+ \tilde{X}_{10000,15000} \left[F(15000) - F(10000)\right] + \tilde{X}_{15000,20000} \left[F(20000) - F(15000)\right]$$

$$+ \tilde{X}_{20000,\infty} \left[1 - F(20000)\right],$$

for some $\tilde{X}_{x,x'} \in [x, x']$ for any $x, x' \in \{0, 5000, 10000, 15000, 20000\}$, and $\tilde{X}_{20000,\infty} \geq 20000$. To obtain the lower bound of μ, we replace $\tilde{X}_{x,x'}$ simply with x, which yields the following equation

$$\begin{aligned} \underline{\mu} \quad &= \quad 5{,}000 \times \left[F(10000) - F(5000)\right] + 10{,}000 \times \left[F(10000) - F(15000)\right] \\ &\quad + 15{,}000 \times \left[F(20000) - F(15000)\right] + 20{,}000 \times \left[1 - F(20000)\right]. \end{aligned} \tag{13}$$

Equation (12) can simply identify what is above the lower bound. Note that in our conjoint design, the lower bound is estimable because *Boundary* defines the lower bound of individual WTP as 0. However, our rational choice model does not provide the upper bound of individual WTP, and the upper bound cannot be identified. This "non-identification" result of the upper bound is, on the one hand, a limitation of our approach. On the other hand, however, the lower bound provides a sufficient cost-benefit condition for efficiency improvement: the policy can improve social efficiency if the estimated lower bound exceeds the expected cost of implementing the policy.

Finally, similar to the marginal WTP distribution, equation (11) implies that our conjoint data can also estimate five probability intervals of the MCE-WTP as

$$\hat{\Pi}_l(0 \leq WTP_i \leq 5000 | a_1, a_0) = -\overline{Y}_{ij}|_{C_{ij}=5000, A_{ijl}=a_1} + \overline{Y}_{ij}|_{C_{ij}=5000, A_{ijl}=a_0},$$

$$\hat{\Pi}_l(5000 \leq WTP_i \leq 10000 | a_1, a_0) = \overline{Y}_{ij}|_{C_{ij}=5000, A_{ijl}=a_1} - \overline{Y}_{ij}|_{C_{ij}=10000, A_{ijl}=a_1}$$

$$-\overline{Y}_{ij}|_{C_{ij}=5000, A_{ijl}=a_0} + \overline{Y}_{ij}|_{C_{ij}=10000, A_{ijl}=a_0},$$

$$\hat{\Pi}_l(10000 \leq WTP_i \leq 15000|a_1, a_0) = \overline{Y}_{ij}|_{C_{ij}=10000, A_{ijl}=a_1} - \overline{Y}_{ij}|_{C_{ij}=15000, A_{ijl}=a_1}$$

$$-\overline{Y}_{ij}|_{C_{ij}=10000, A_{ijl}=a_0} + \overline{Y}_{ij}|_{C_{ij}=15000, A_{ijl}=a_0},$$

$$\hat{\Pi}_l(15000 \leq WTP_i \leq 20000|a_1, a_0) = \overline{Y}_{ij}|_{C_{ij}=15000, A_{ijl}=a_1} - \overline{Y}_{ij}|_{C_{ij}=20000, A_{ijl}=a_1}$$

$$-\overline{Y}_{ij}|_{C_{ij}=15000, A_{ijl}=a_0} + \overline{Y}_{ij}|_{C_{ij}=20000, A_{ijl}=a_0},$$

and finally,

$$\hat{\Pi}_l(20000 \leq WTP_i|a_1, a_0) = \overline{Y}_{ij}|_{C_{ij}=20000, A_{ijl}=a_1} - \overline{Y}_{ij}|_{C_{ij}=20000, A_{ijl}=a_0}. \tag{14}$$

The first and the last equations are obtained from

$$Y_{ij}|_{C_{ij}=0, A_{ijl}=a_1} = Y_{ij}|_{C_{ij}=0, A_{ijl}=a_0} = 1$$

and

$$Y_{ij}|_{C_{ij}\to\infty, A_{ijl}=a_1} = Y_{ij}|_{C_{ij}\to\infty, A_{ijl}=a_0} = 0$$

as well as the *Boundary* assumption.

Equation (14) shows that the component effect on some selected points on the WTP distribution is estimable from the conjoint data with a discrete variation in the policy burden.

3.3.2 Estimation of approval rating

The data from the conjoint choice experiment allow us to estimate the causal effects of attributes on the approval rating. A la HHY conjoint, we estimate the following population model:

$$E\left[Y_{ij}|C_{ij} = c, A_{ij} = a\right] = \beta_0 + \beta_c I_{ijc} + \sum_{l=1}^{6} \beta_l I_{ijl}, \tag{15}$$

where I_{ijc} and I_{ijl} are vectors of dummy variables composed of each level of the financial burden and attribute l, respectively, while β_c and β_l are vectors of the corresponding coefficients.
Note that the unit of analysis in the regression is each alternative in each task of each respondent. Therefore, although the respondents are sampled randomly from the population, the observed choice outcomes within a respondent may be correlated, which may affect the statistical inference results. For example, respondents have unobservable characteristics that affect their answer in every task, which generates correlation among the choice outcomes within a respondent. To avoid the bias from such correlation in the error

terms, we use the cluster robust standard errors at the respondent level in all regressions, as suggested by Hainmueller, Hopkins, and Yamamoto (2014).

To recover the marginal distribution of WTP, equation (8) requires us to estimate the approval rating conditional on the financial burden, which is obtained from the estimation result of equation (15). However, the identification result of the MCE-WTP (equation 14) requires estimators of the approval rating conditional on the policy burden and an interest attribute l, $E[Y_{ij}|C_{ij} = c, A_{ijl} = a]$. To obtain those values, we estimate the population model with interaction:

$$E[Y_{ij}|C_{ij} = c, A_{ijl} = a] = \beta_0 + \beta_c I_{ijc} + \beta_l I_{ijl} + \beta_I \left(I_{ijc} \times I_{ijl}\right),$$

where β_I^E is a vector of the coefficients of the interaction term between the individual burden and attribute l.

3.4 Estimation Results

This section first presents the estimated AMCE, which captures the average marginal causal effects on the choice probabilities. We then show the estimated welfare quantities, the marginal distribution of WTP and the MCE-WTP.

3.4.1 Average marginal component effect on approval rating

Figures 1 reports the estimated coefficients and the 95% confidence intervals in population model (15). Each solid circle in the figure represents a point estimator, while the horizontal bars represent the 95% and 99% confidence intervals.

[Figure 1 around here]

With regard to the approval rating, Figure 1 shows that *TOILET* and *WASTEWATER* have positive and statistically significant effects, while no significant effects of *FERTILIZER* are observed. Furthermore, *ORGANIZATION* also has positive effects at the 10% significance level. *PERIOD* and *PAYMENT* also tend to have the expected estimated effects: a lower burden and a shorter targeting period have positive causal effects on the approval ratings. Note that the point estimators of individual burdens are not monotonically increasing because the estimated coefficient of 10,000 kyats is lower than the coefficient of 15,000 kyats. However, the results do not imply that the approval rating decreases if the individual burden decreases from 15,000 kyats to 10,000 kyats because the difference between these coefficients is not statistically significant.

An important advantage of a randomized conjoint design is that it allows us to derive causal interpretations for all estimated coefficients, which implies that we can compare the economic significance of the attributes. Among the additional services, the causal effect of *TOILET* is especially high, with a size roughly the same as the effect of reducing the financial burden from 20,000 kyats to 15,000 kyats or reducing the target period from 20 years to 5 years. Thus, Figure 1 demonstrates the high demand for a toilet service, which implies that a water improvement project that includes a toilet service would have more support from the local communities. Additionally, a collective wastewater treatment facility and joint implementation with local NGO are effective ways of securing residents' support.

3.4.2 Results of the welfare estimation

[Figure 2 around here]

We first report the estimated marginal distribution of WTP (equation 8) in Figure 2. The figure reports the estimated histogram with five bins, which shows that more than 70% of people have WTP of more than 20,000 kyats in each month. Because the median monthly income in the area is 270,000 kyats, most people's WTP is larger than 7.4% of median income. Additionally, the shares of people in other ranges are approximately 10% except for between 10,000 kyats and 15,000 kyats, which shows the diversity of individual WTP values due to heterogeneous policy preferences and/or policy attributes.

Note that the share of people whose WTP values are between 10,000 and 15,000 kyats is estimated to be negative, which indicates that the *Monotonicity* assumption may be violated. However, this estimated share is not statistical significant, and the concern that the assumption has been violated is not serious.

Equation (13) demonstrates that the lower bound of the average WTP can be estimated with the estimated marginal distribution of individual WTP. Using the estimators shown in Figure 2, the estimated

Table 1: List of the changes in attributes

Original level a_0	New level a_1
No toilet system	Toilet system
Toilet system	Toilet system with collection service
No garbage collection service	Garbage collection service once per week
No wastewater treatment service	Wastewater treatment service
Guidance on fertilizer input	Guidance on pesticide input
Guidance on fertilizer input	Guidance on both fertilizer and pesticide input
Project period: 20 years	Project period: 10 years
Project period: 10 years	Project period: 5 years
Implemented by government	Implemented by government and local NGOs

lower bound of the average marginal welfare gain is 16,090 kyats. Because the median monthly household income in the area is 270,000 kyats, the estimated average WTP is at least 5.9% of the median income. This paper thus demonstrates strong aggregate demand for water improvement policy through the high average WTP for the policy.

[Figures 3 to 8 around here]

Finally, Figures 3 to 8 report the estimated MCE-WTP (equation 14). We estimate the causal effects of following 9 scenarios of attribute changes (See Table 1).

Figure 3 shows the clear effect of toilet provision; providing a toilet system significantly increases the share of workers whose WTP values are larger than 20,000 kyats but decreases the share of people whose WTP values are between 0 and 5,000 kyats. The figure further reports that also providing a collection service does not have clear effects on the WTP distribution. Therefore, our estimates consistently demonstrate the importance of implementing a water improvement policy that includes a toilet service; toilet provision increases not only the approval rating but also improves the WTP distribution.

From Figures 4 to 8, with some exceptions, we cannot observe clear effects of the other attributes. An exception is the garbage collection service; the share of people whose WTP values exceed 20,000 kyats is significantly increased. Providing guidance on pesticide input decreases the share of the lowest bin but increases the share of people whose WTP is between 5,000 and 10,000 at the 5% significance level. Reducing the project period from 10 to 5 years increases the highest bin and decreases the bin range [15,000, 20,000]. Finally, joint implementation with local NGOs decreases the share of people whose WTP is between 5,000 and 10,000.

4 Conclusion

This study introduced a new approach to non-parametrically derive welfare implications from conjoint experiment data obtained by Hainmueller, Hopkins, and Yamamoto (2014). Using only weak assumptions on the preference distribution, the method enables us to estimate the causal effects of individual attributes on the WTP distribution. Our approach can apply to policy evaluations based on WTP. This paper applied this approach to the evaluation of a local environmental improvement project, and evaluations of such projects are otherwise difficult to estimate. Adopting this method to Inlay Lake, we conclude that the water quality improvement policy generates an average marginal surplus gain of at least 193,082 kyats per annum, which is 22.9% of the average annual per capita income in the area. Additionally, the analysis of the conditional average surplus gain shows that including the toilet service increases the lower bound of the surplus gain by 14%. The analyses of the choice probabilities and the surplus gain demonstrate the importance of providing a toilet service when implementing a water quality improvement policy. We believe that our approaches are useful in policy research to improve social welfare. However, our study has an important limitation in terms of the external validity of the randomized conjoint analysis. Although Hainmueller, Hangartner, and Yamamoto (2015) provide evidence of the validity of such an approach in developed countries, no studies focus on developing countries. Therefore, additional studies are needed to test the external validity of the conjoint experiments in developing countries.

Acknowledgments

We gratefully acknowledge financial support for the study by the Hiroshima University TAOYAKA Program for creating a flexible, enduring, peaceful-society, funded by the Program for Leading Graduate Schools, Ministry of Education, Culture, Sports, Science and Technology and the Sumitomo Foundation #153421. In addition, the field experiment was in part financially supported by the Japanese Grant Aid for Human Resource Development Scholarship (JDS) Program. We are also particularly thankful to Teppei Yamamoto and participants at the Polmeth XXXXIV for their very useful insights. All errors are our own.

References

[1] Arrow, K., Solow, R., Portney, P. R., Leamer, E. E., Radner, R., & Schuman, H. (1993). Report of the NOAA panel on contingent valuation. Federal register, 58(10), 4601-4614.

[2] Akaishi, F., Satake, M., Otaki, M., & Tominaga, N. (2006). Surface water quality and information about the environment surrounding Inle Lake in Myanmar. Limnology, 7(1), 57-62.

[3] Alberini, A., & Cooper, J. (2000). Applications of the contingent valuation method in developing countries: A survey (Vol. 146). Food & Agriculture Org..

[4] Bhattacharya, D. (2015). Nonparametric welfare analysis for discrete choice. Econometrica, 83(2), 617-649.

[5] Bechtel, M. M., & Scheve, K. F. (2013). Mass support for global climate agreements depends on institutional design. Proceedings of the National Academy of Sciences, 110(34), 13763-13768.

[6] Bernauer, T., & Gampfer, R. (2015). How robust is public support for unilateral climate policy?. Environmental Science & Policy, 54, 316-330.

[7] Birol, E., & Koundouri, P. (Eds.). (2008). Choice experiments informing environmental policy: a European perspective. Edward Elgar Publishing.

[8] Bodea, C., & LeBas, A. (2016). The origins of voluntary compliance: attitudes toward taxation in urban Nigeria. British Journal of Political Science, 46(1), 215-238.

[9] Bowen, H. R. (1943). The interpretation of voting in the allocation of economic resources. The Quarterly Journal of Economics, 58(1), 27-48.

[10] Butkus, S., & Myint, S. (2001). Pesticide use limits for protection of human health in Inle Lake (Myanmar) Watershed. Living Earth Institute, Olympia, Washington, USA). www. living-earth. org/inle. html.

[11] Carson, R., & Czajkowski, M. (2014). The discrete choice experiment approach to environmental contingent valuation (Doctoral dissertation, Edward Elgar Publishing).

[12] Carson, R. T., & Hanemann, W. M. (2005). Contingent valuation. Handbook of environmental economics, 2, 821-936.

[13] Carson, R. T., & Louviere, J. J. (2011). A common nomenclature for stated preference elicitation approaches. Environmental and Resource Economics, 49(4), 539-559.

[14] Chetty, R. (2009). Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods. Annu. Rev. Econ., 1(1), 451-488.

[15] Ciriacy-Wantrup, S. V. (1947). Capital returns from soil-conservation practices. Journal of farm economics, 29(4), 1181-1196.

[16] Competition Commission. (2010). Review of stated preference and willingness to pay methods. London, UK: Competition Commission.

[17] Davis, R. K. (1963). The value of outdoor recreation: an economic study of Maine woods. Unpublished Ph. D. dissertation, Harvard University.

[18] Durand-Morat, A., Wailes, E. J., & Nayga, R. M. (2015). Challenges of Conducting Contingent Valuation Studies in Developing Countries. American Journal of Agricultural Economics, aav068.

[19] Fairbrother, M. (2017). When Will People Pay to Pollute? Environmental Taxes, Political Trust and Experimental Evidence from Britain. British Journal of Political Science, 1-22.

[20] Gampfer, R., Bernauer, T., & Kachi, A. (2014). Obtaining public support for North-South climate funding: Evidence from conjoint experiments in donor countries. Global Environmental Change, 29, 118-126.

[21] Green, P. E., & Rao, V. R. (1971). Conjoint measurement for quantifying judgmental data. Journal of Marketing research, 355-363.

[22] Greenstone, M., & Jack, B. K. (2015). Envirodevonomics: A Research Agenda for an Emerging Field. Journal of Economic Literature, 53(1), 5-42.

[23] Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. Proceedings of the National Academy of Sciences, 112(8), 2395-2400.

[24] Hainmueller, J., & Hopkins, D. J. (2015). The hidden American immigration consensus: A conjoint analysis of attitudes toward immigrants. American Journal of Political Science, 59(3), 529-548.

[25] Hainmueller, J., Hopkins, D. J., & Yamamoto, T. (2014). Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments. Political Analysis, 22(1), 1-30.

[26] Hausman, J. A., & Newey, W. K. (2017). Nonparametric welfare analysis. Annual Review of Economics, 9(1).

[27] Hanley, N., Mourato, S., & Wright, R. E. (2001). Choice modeling approaches: a superior alternative for environmental valuatioin?. Journal of economic surveys, 15(3), 435-462.

[28] Hensher, D.A., Rose, J.M., Greene, W.H., 2005. Applied Choice Analysis. Cambridge University Press, Cambridge.

[29] Holmes, T. P., & Adamowicz, W. L. (2003). Attribute-based methods. In A primer on nonmarket valuation (pp. 171-219). Springer Netherlands.

[30] Perrings, C. (2014). Environment and development economics 20 years on. Environment and Development Economics, 19(03), 333-366.

[31] Khin, L. S. (2011) Development of clean water and sanitation in Inle Lake, Myanmar. EE2 Seminar: Water and Environment in Asia's Developing Communities Singapore International Water Week 2011

Co-located Event, Suntec Singapore International Convention and Exhibition Centre, 6 July 2011, Singapore.

[32] Lancaster, K. J. (1966). A new approach to consumer theory. The journal of political economy, 132-157.

[33] Mcfadden, D. (1974). Conditional logit analysis of qualitative choice behavior. Frontiers in Econometrics, 105-142.

[34] Ministry of Immigration and Population (2015) 2014 Myanmar Population and Housing Census, Union Report: Census Report Volume (2), Ministry of Immigration and Population, The Republic of the Union of Myanmar.

[35] Neyman, J (1923). On the application of probability theory to agricultural experiments: Essay on principles, section 9. (translated in 1990). Statistical Science 5, 465–80.

[36] Norwegian Government and UN-Habitat (2014). Inlay Lake Long Term Restoration and Conservation Plan. Ministry of Environmental Conservation and Forestry, Myanmar. http://unhabitat.org.mm/wp-content/uploads/2015/03/Long-Term-Restoration-Conservation-Plan-Inle-Lake.pdf (accessed on March 10, 2016)

[37] Perrings, C. (2014). Environment and development economics 20 years on. Environment and Development Economics, 19(03), 333-366.

[38] Raghavarao, D., J. B. Wiley, and P. Chitturi. 2011. Choice-based conjoint analysis: Models and designs. Boca Raton, FL: CRC Press.

[39] Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. Journal of educational Psychology, 66(5), 688.

[40] Sidle, R. C., Ziegler, A. D., & Vogler, J. B. (2007). Contemporary changes in open water surface area of Lake Inle, Myanmar. Sustainability Science, 2(1), 55-65.

[41] Small, K. A., & Rosen, H. S. (1981). Applied welfare economics with discrete choice models. *Econometrica*, 49(1), 105-130.

[42] Su, M., & Jassby, A. D. (2000). Inle: A large Myanmar lake in transition. Lakes & Reservoirs: Research & Management, 5(1), 49-54.

[43] Train, K. E. (2009). Discrete choice methods with simulation. Cambridge university press.

[44] Whittington, D. (1998). Administering contingent valuation surveys in developing countries. World development, 26(1), 21-30.

[45] Whittington, D. (2002). Improving the performance of contingent valuation studies in developing countries. Environmental and resource economics, 22(1), 323-367.

[46] Whittington, D., Adamowicz, W., & Lloyd-Smith, P. (2017). Asking Willingness-to-Accept Questions in Stated Preference Surveys: A Review and Research Agenda. Annual Review of Resource Economics, 9(1).

[47] Wiswall, M., & Zafar, B. Preference for the Workplace, Human Capital, and Gender (No. w22173). forthcoming, Quarterly Journal of Economics.

A. Appendix: Details of respondents' characteristics

A.1. Primary and secondary jobs

[Table A1]

In total, 43.4% of household heads are engaged in farming, typically tomato production (see Table A1). In the Inlay area, farmers cultivate tomatoes for 6 months, starting in March or April, and take on other jobs for the rest of the year, such as fishing, carpentry, and seasonal jobs. Although tomatoes can be harvested after 40 days from shedding seeds, farmers divide their floating gardens into blocks, and seeding and harvesting are scheduled to be repeated every 6 to 14 weeks. Thus, farmers are usually engaged in tomato farming for six months. Fishing is more common as a secondary job (77 household heads) than as a primary job (44 household heads).

A.2. Composition of household income by source

[Table A2]

In addition to the job structure of household heads, Table A2 summarizes the household income by source. The economic significance of farming is the largest, representing 36.13% of the total. Although fishing was replaced by tomato farming, approximately 15% of income is still generated by fishing. Relatively few villagers engage in small business, but such activities generate relatively large profits (approximately 10%). In contrast, 85 households (26%) generate income from local cheroot cigarette businesses as their main supplementary income source. However, this represents only 5% of income, despite it being a common activity.

A.3. Perception of recent changes on water quality

The respondents were asked to give their perceptions of recent (the last five years) changes in the water quality of Inlay Lake. The results show that 48% of respondents perceive that the water quality of the lake is significantly worse, while 4.9% perceive no change or that it has improved.

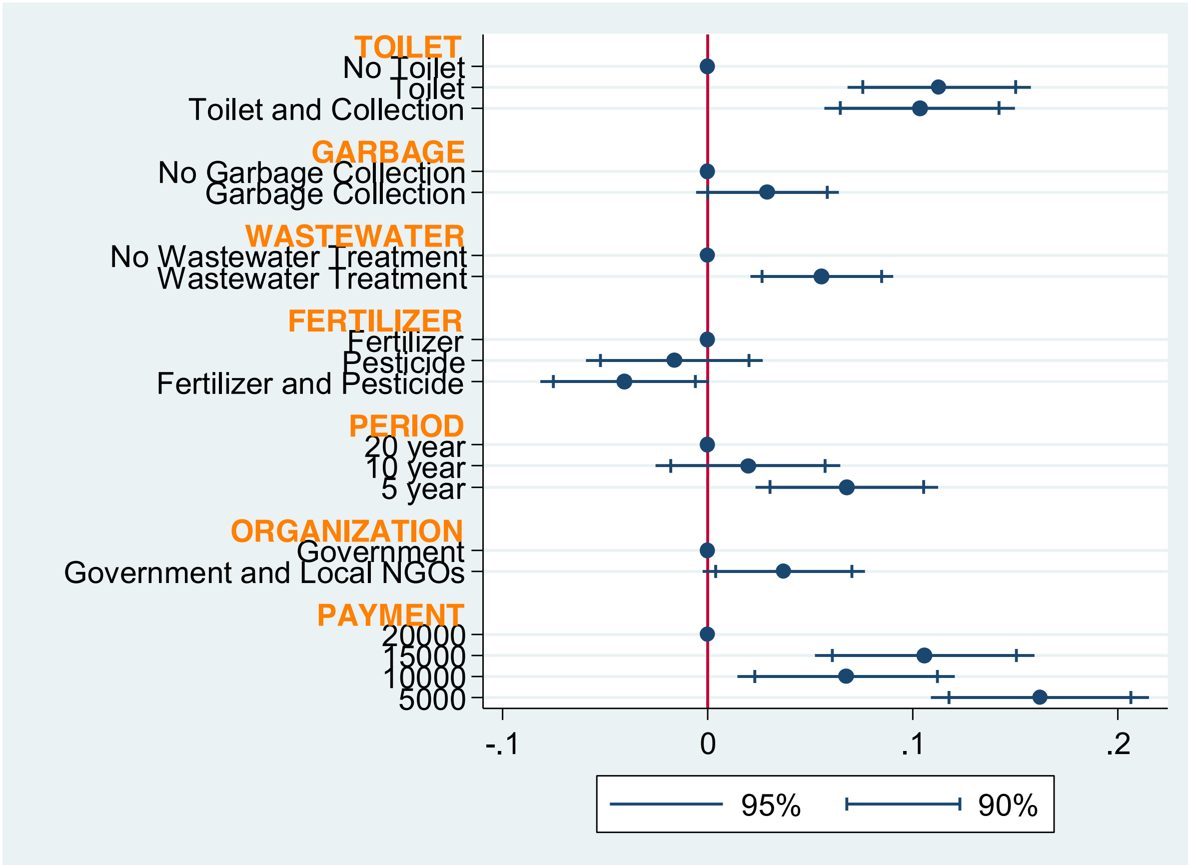

Fig 1. Estimated effect of policy attributes on approval rating. Each solid circle shows estimated marginal effects of individual policy. The estimation is based on the OLS regression with clustered standard error; horizontal bars represent 90% and 95% confidence levels. Note that circle without bars denote the reference level of each attribute.

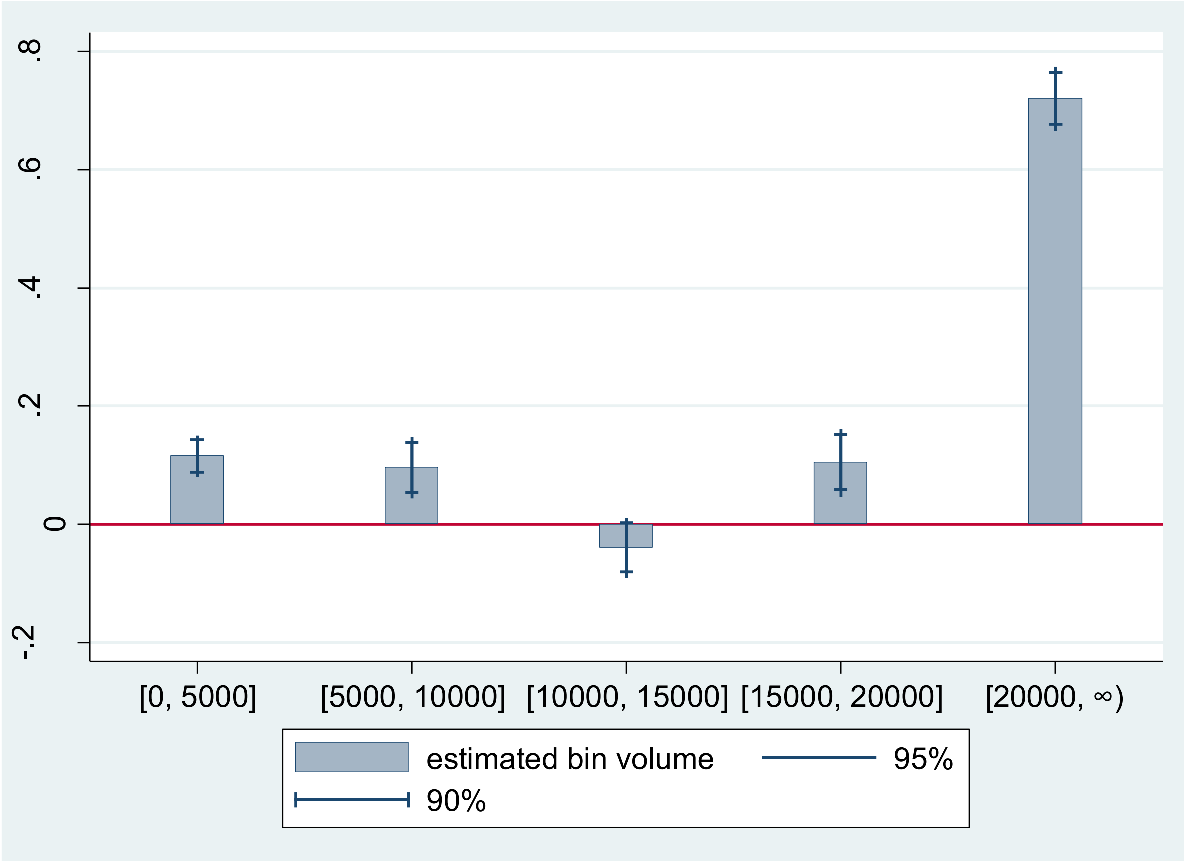

Fig 2. Estimated distribution of WTP. This histogram shows the estimated marginal distribution of willingness-to-pay. The estimated histogram is based on the OLS regression with clustered standard error; vertical bars represent the 90% and 95% confidence interval.

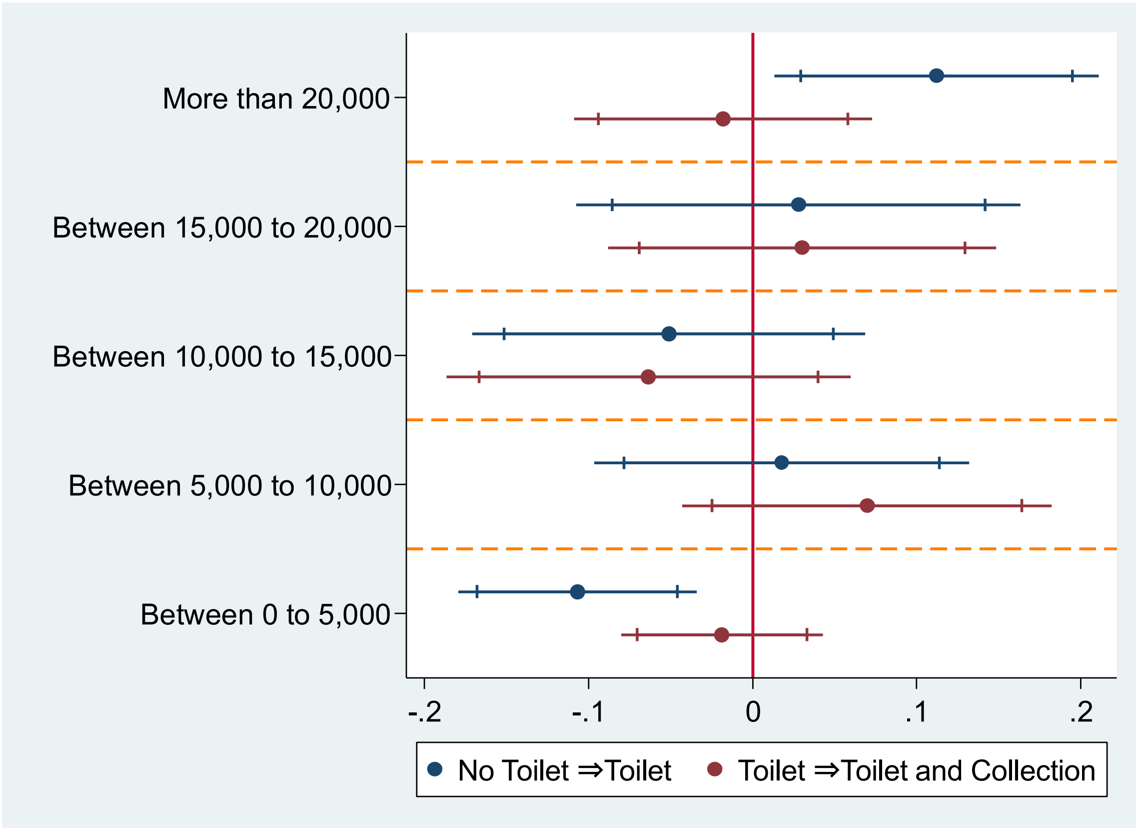

Fig 3. Estimated effect of TOILET attribute on the WTP distribution. Each circle shows estimated marginal effects of changing level of TOILET attribute on each bin of the histogram. The estimates are based on the OLS regression with clustered standard error; horizontal bars represent the 90% and 95% confidence interval.

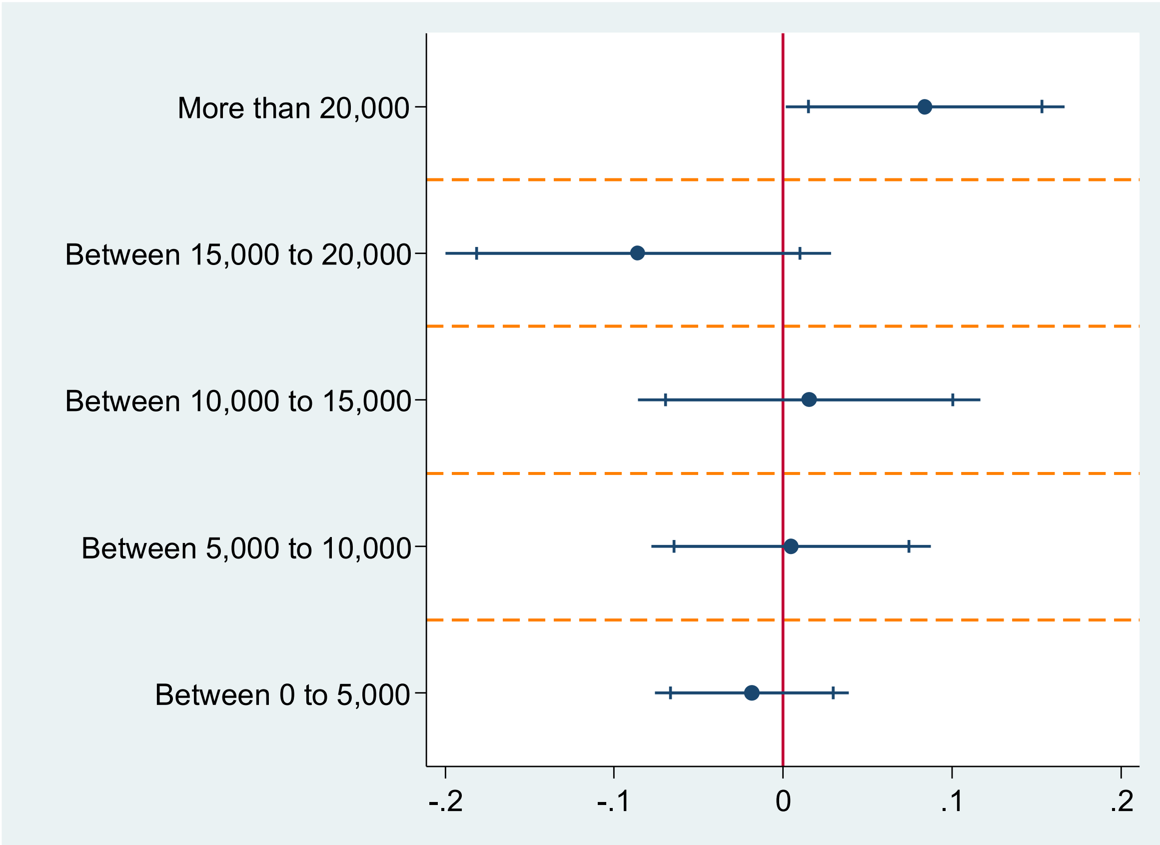

Fig 4. Estimated effect of GARBAGE attribute on the WTP distribution. Each circle shows estimated marginal effects of changing level of GARBAGE attribute on each bin of the histogram. The estimates are based on the OLS regression with clustered standard error; horizontal bars represent the 90% and 95% confidence interval.

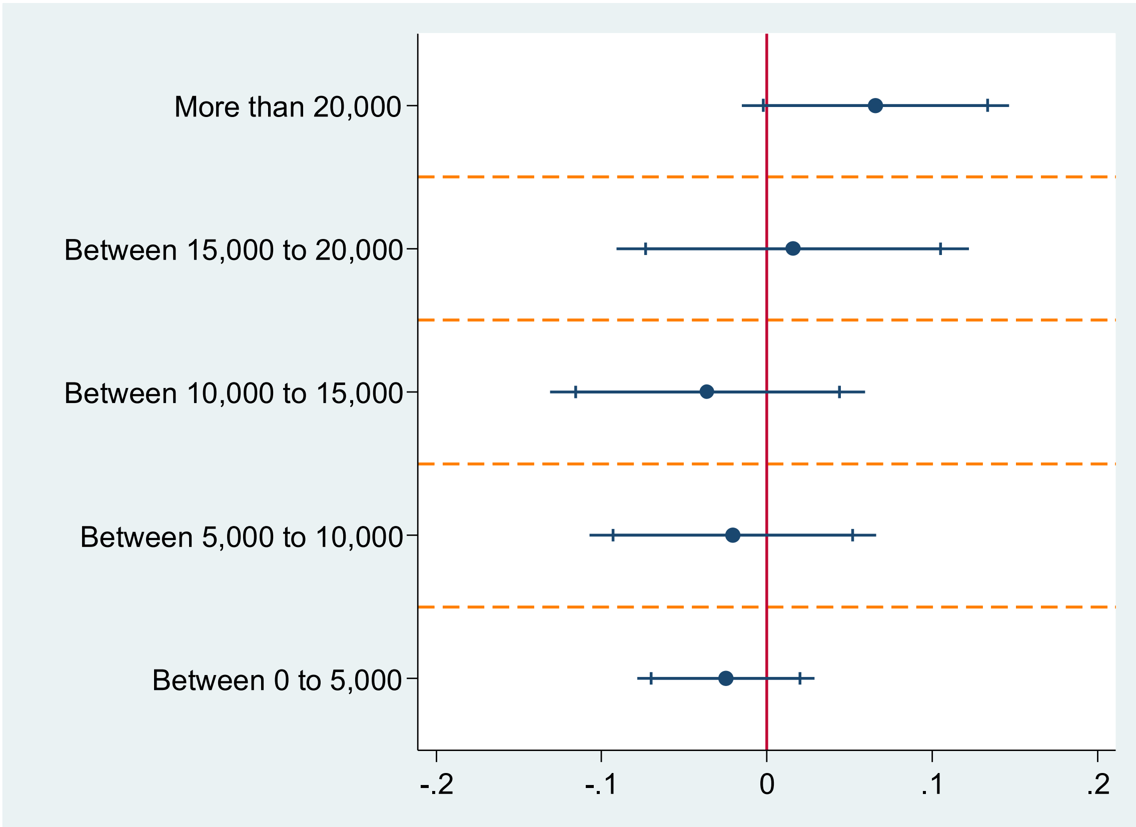

Fig 5. Estimated effect of WASTEWATER attribute on the WTP distribution. Each circle shows estimated marginal effects of changing level of WASTERWATER attribute on each bin of the histogram. The estimates are based on the OLS regression with clustered standard error; horizontal bars represent the 90% and 95% confidence interval.

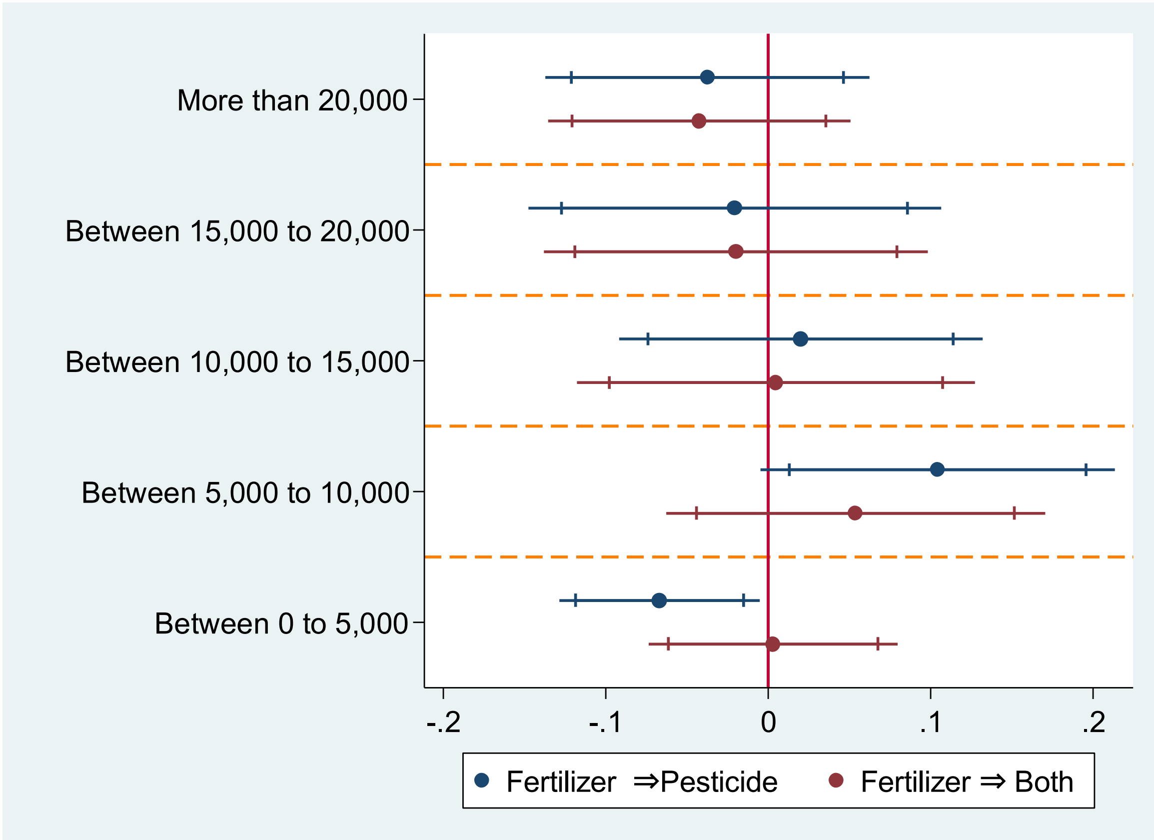

Fig 6. Estimated effect of FERTILIZER attribute on the WTP distribution. Each circle shows estimated marginal effects of changing level of FERTILIZER attribute on each bin of the histogram. The estimates are based on the OLS regression with clustered standard error; horizontal bars represent the 90% and 95% confidence interval.

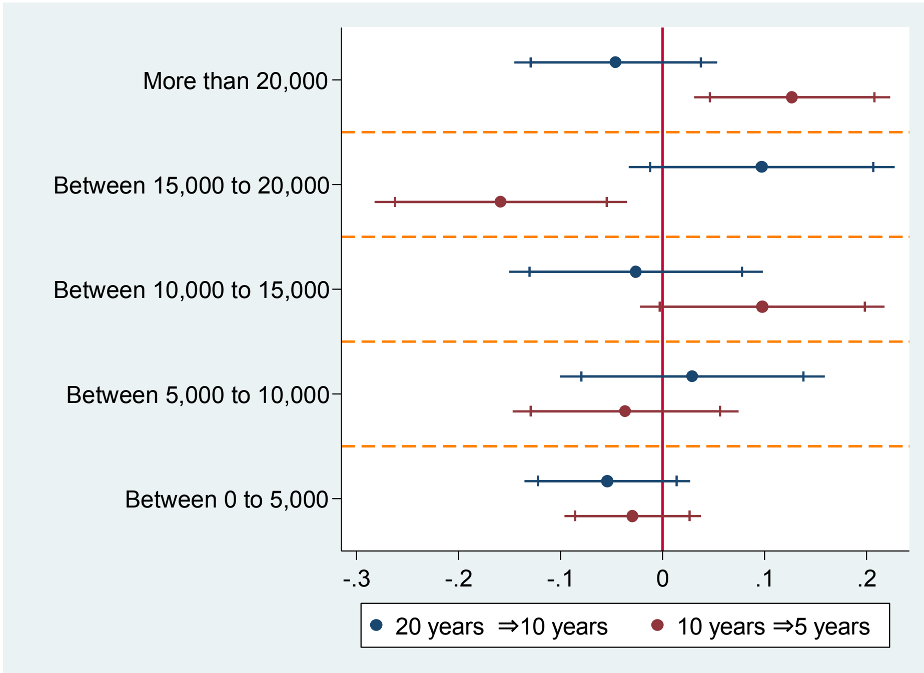

Fig 7. Estimated effect of PERIOD attribute on the WTP distribution. Each circle shows estimated marginal effects of changing level of PERIOD attribute on each bin of the histogram. The estimates are based on the OLS regression with clustered standard error; horizontal bars represent the 90% and 95% confidence interval.

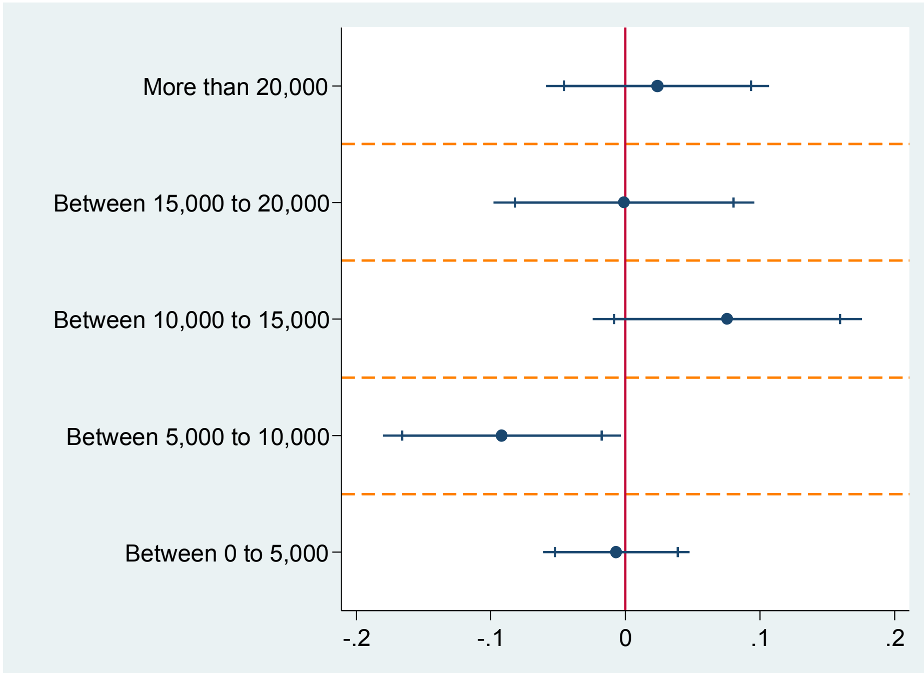

Fig 8. Estimated effect of ORGANIZATION attribute on the WTP distribution. Each circle shows estimated marginal effects of changing level of ORGANIZATION attribute on each bin of the histogram. The estimates are based on the OLS regression with clustered standard error; horizontal bars represent the 90% and 95% confidence interval.

Table A1. Primary and secondary jobs

Primary	Number of household heads	Composition (%)	Secondary	Number of household heads	Composition (%)
Farmer	142	43.4%	Fisher	77	23.5%
Fisher	44	13.5%	Wage worker for farming	22	6.7%
Wage worker for farming	27	8.3%	Shopkeeper	18	5.5%
Carpenter	26	8.0%	Farmer	7	2.1%
Local cheroot	10	3.1%	Boat driver	5	1.5%
Small business	8	2.4%	Wage work for tailoring	5	1.5%
Shopkeeper	7	2.1%	Small business	3	0.9%
Trader/Broker	6	1.8%			
Various others	26	8.0%	Various others	44	13.5%
No primary jobs	31	9.5%	No secondary jobs	146	44.6%
Total	**327**	**100.0%**		**327**	**100.0%**

Table A2. Distribution of total household income by source.

Source	Composition (%)
Farming	36.13
Fishing	15.30
Small business	10.01
Carpenter	6.09
Wage work for farming	6.39
Local cheroot	5.22
Trader/Broker	4.08
Wage work for tailoring	4.05
Boat driver	2.00
Various others	10.73
Total	**100.00**