



# R package DCchoice for dichotomous choice contingent valuation: a contribution to open scientific software and its impact

Hideo Aizaki<sup>1</sup> · Tomoaki Nakatani<sup>2</sup> · Kazuo Sato<sup>3</sup> · James Fogarty<sup>4</sup>

Received: 31 January 2022 / Revised: 23 May 2022 / Accepted: 24 June 2022 / Published online: 26 July 2022  
© The Author(s) 2022

## Abstract

The R package DCchoice is designed to mitigate programing-related barriers to the application of dichotomous choice contingent valuation (DCCV) methods in empirical studies. Since its release in 2014, DCchoice has been updated. This paper introduces the current version of DCchoice which supports single-, one-and-one-half-, and double-bounded DCCVs, with and without a spike. Additionally, the willingness-to-pay and its confidence intervals can be calculated for a representative respondent as well as for a user-defined specific respondent using the current version. The associated web tutorial and R Commander plug-in for basic usage of DCchoice are also available. DCchoice has advanced DCCV applications in various fields.

**Keywords** Contingent valuation · Dichotomous choice · R package · Stated preference method

## 1 Introduction

Computer languages and/or the software products developed using such languages are a critical component of the infrastructure that supports statistical and data sciences and related fields (Donoho, 2017). Software development is increasingly recognized as a significant scientific activity in its own right (Howison et al., 2015). Free and/or open

---

✉ Hideo Aizaki  
azk-r@spa.nifty.com

<sup>1</sup> Research Faculty of Agriculture, Hokkaido University, Sapporo, Hokkaido, Japan

<sup>2</sup> Graduate School of Agricultural and Life Sciences, The University of Tokyo, Bunkyo, Tokyo, Japan

<sup>3</sup> College of Agriculture, Food and Environment Sciences, Rakuno Gakuen University, Ebetsu, Hokkaido, Japan

<sup>4</sup> Agricultural and Resource Economics, School of Agriculture and Environment, The University of Western Australia, Perth, WA, Australia

source software (F/OSS) occupies an especially important role within the statistical and data science field. F/OSS enhances the reproducibility of research results, allows innovations in techniques and approach to be quickly validated and shared, and provides the basis for educating the next generation of researchers and analysts across many fields (Carmichael & Marron, 2018). There are various languages and/or software products released under F/OSS licenses, but R (R Core Team, 2021) is the most popular dedicated statistical computer language.<sup>1</sup> R is used widely across all science disciplines, and also in the humanities when data analysis constitutes a major activity (for the evolution of R, see Chambers, 2020). The popularity of R is fundamentally linked to the ease of installation, and its expandability, both of which are facilitated by the Comprehensive R Archive Network (CRAN). CRAN is the official R site for distributing R and its add-on packages. R can be expanded by installing an add-on package, which is a collection of functions, dataset, and help documents for a specific topic. CRAN connects package users and developers effectively: users can easily install add-on packages to their computer from CRAN by executing a single command, and developers can publish their packages to their (potential) users via CRAN (for the features and roles of CRAN in detail, see, Hornik, 2012; Mora-Cantalops et al., 2020). As of 23 May 2022, 18,591 packages were available on CRAN.

DCchoice (Nakatani et al., 2021) is an add-on R package that provides functions for analyzing responses to contingent valuation (CV) questions available from CRAN. CV involves asking people about their willingness-to-pay (WTP) for a good or service, and CV has been used extensively to value non-market goods and services (Carson, 2012). Applications of CV methods are varied, and this diversity can be seen across examples considering: the benefits of water quality improvement (Whitehead, 1995, 2015); consumer values for food labeling information (Loureiro & Hine, 2004); citizen value of vaccines (Gracia & Cerda, 2020); and patient value of time (Verbooy et al., 2018).

CV methods can be implemented in a number of ways (Carson & Hanemann, 2005), but the most popular approach is the dichotomous choice contingent valuation (DCCV) method. The classical DCCV implementation is the single-bounded (SB) DCCV method (Bishop & Heberlein, 1979), and this approach asks respondents whether they would be willing to pay a ‘bid’ amount of money for a change in a good or service. Varying bid values across respondents reveals the WTP for the change.

The first version of DCchoice was released from R-Forge in 2014, and since 2015 revised versions have been distributed from CRAN. DCchoice was developed to mitigate programming-related barriers associated with implementing DCCV methods for empirical studies. Both parametric and non-parametric methods have been used to analyze the responses to DCCV questions (Carson & Hanemann, 2005). Parametric approaches generally fit discrete choice models to the dataset using a maximum likelihood estimation method. SB-DCCV survey data can be analyzed with standard logit or probit model specifications. However, when analyzing the responses to other format questions users typically need to write code chunks to: (a) define the log-likelihood function corresponding to the model to be fitted, and (b) search parameter values

<sup>1</sup> Current popularity metrics for languages can be found at <https://www.tiobe.com/tiobe-index/> and <https://pypi.github.io/PYPL.htm>, and while there are more popular languages than R, for example, Python, these tend to be general application languages and not dedicated data science and statistical analysis languages.

that maximize the log-likelihood function under the responses using an optimization function. The non-parametric estimation approach also needs support from software packages because, for example, the proportion of YES for each bid should be adjusted when the relationship between bids and the proportion of YES is a non-monotonic sequence. Additional lines of code are also needed to calculate the mean/median WTP and associated confidence intervals from the fitted models, even when using the SB-DCCV model. Writing the required code to estimate DCCV models is thus challenging, and the need for this technical skill is a barrier to implementing DCCV.

An early version of DCchoice is detailed in Aizaki et al. (2014), but over time the features of DCchoice have been substantially updated in response to user requests; the available models have been expanded; and the functionality regarding the calculation of WTPs extended. This paper provides an overview of DCCV methods (Sect. 2); describes the current version of DCchoice, and documents the differences between the initial and current versions of the package (Sect. 3); illustrates how to use DCchoice (Sect. 4); provides evidence on the impact of DCchoice on CV applications (Sect. 5); and then presents some concluding comments (Sect. 6).

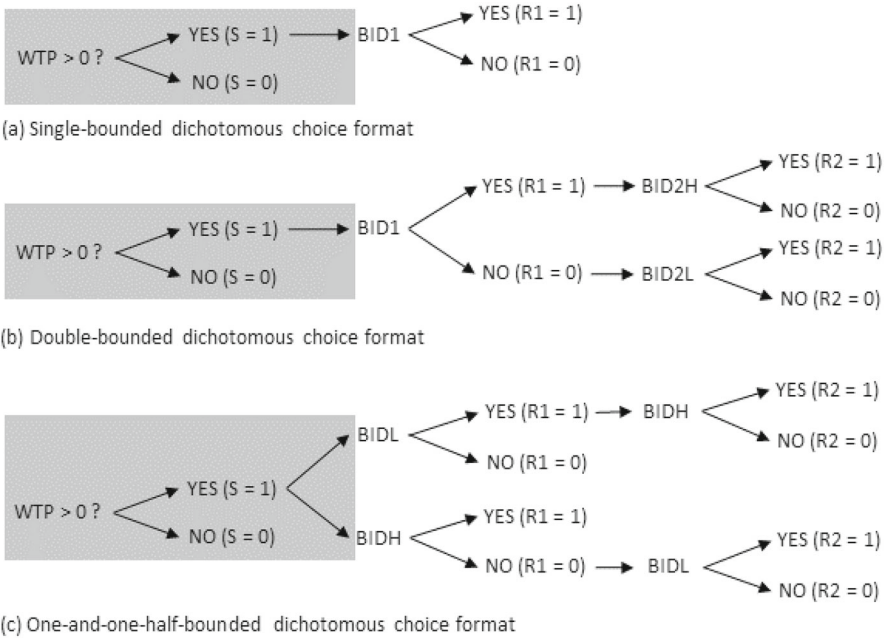
## 2 Outline of DCCV formats

In addition to the SB-DCCV format, DCCV has two other major variants: the double-bounded (DB) DCCV format (Carson, 1985; Hanemann, 1985); and the one-and-one-half-bounded (OOHB) DCCV format (Cooper et al., 2002). Each model can also be implemented either with or without a spike at zero (Kriström, 1997). A model with a spike at zero allows for respondents to have zero WTP for the good because they are not in the relevant ‘market’. A spike model is analogous to the corner solution outcome in private good markets, where even when the price is zero, the good is not purchased.

The SB-DCCV asks respondents whether they would be willing to pay a bid for a change in a good or service (Fig. 1a). The bid is randomly assigned from a list of bid values defined in advance. The relationship between the bids (BID1) and their responses (proportion of YES [ $R1 = 1$ ] responses) can be modeled statistically. The model reveals the WTP for a change in the good or service.

The DB-DCCV involves respondents answering a DCCV question in SB format twice (Fig. 1b). Although the bid shown to respondents in the first stage (BID1) is assigned randomly from the list of bids in the same way as the SB-DCCV, the bid in the second stage is determined by the response to the first stage. Specifically, the second bid value is higher (BID2H) than the first when the first response is YES ( $R1 = 1$ ) they are willing-to-pay the amount, and lower (BID2L) than the first, when the first response is NO ( $R1 = 0$ ). Thus, there are four possible patterns of responses to the DB-DCCV question: YES–YES ( $R1 = R2 = 1$ ), YES–NO ( $R1 = 1, R2 = 0$ ), NO–YES ( $R1 = 0, R2 = 1$ ), and NO–NO ( $R1 = R2 = 0$ ).

With the OOHB-DCCV format (Fig. 1c), after answering the first stage of the DCCV question, only respondents who satisfy a condition are requested to answer the second-stage question in the SB-DCCV format. The steps of the OOHB question are as follows: (1) a list of bid ranges [ $BIDL_j, BIDL_j$ ] ( $j = 1, 2, \dots, J$ ), where  $BIDL_j < BIDL_j$ ,



**Fig. 1** Steps in single-bounded-, double-bounded-, and one-and-one-half-bounded dichotomous choice contingent valuation questions with and without a spike question (gray background)

is designed; (2) one of the bid ranges is randomly assigned to the respondent; (3) the respondent is asked whether they would pay the amount (the first stage) randomly selected from the two bids; and (4) they are asked to answer the second stage if they satisfy either condition—(a) their answer in the first stage is YES ( $R1 = 1$ ) when the lower bid (BIDL) is presented in the first stage; and (b) their answer in the first stage is NO ( $R1 = 0$ ) when the higher bid (BIDH) is presented in the first stage. There are then six possible responses to the OOHB–DCCV question: NO ( $R1 = 0$ ), YES–NO ( $R1 = 1, R2 = 0$ ), and YES–YES ( $R1 = R2 = 1$ ) when the lower bid is shown in the first stage; and YES ( $R1 = 1$ ), NO–YES ( $R1 = 0, R2 = 1$ ), and NO–NO ( $R1 = R2 = 0$ ) when the higher bid is shown in the first stage.

The spike model implementation assumes a non-zero probability of a zero WTP for a good or service, and a zero probability of a negative WTP (Kriström, 1997). To implement this model variation an additional question asking respondents whether their WTP is positive ( $S = 1$ ) or not ( $S = 0$ ) is used. Only respondents who have a positive WTP are asked to answer the DCCV question. The spike model assumption has been considered in SB-, DB-, and OOHB–DCCV studies (Kriström, 1997; Yoo & Kwak, 2002; Kwak et al., 2013).

### 3 Software design

#### 3.1 Roles of DCchoice

DCchoice provides functions for preparing a dataset, fitting models, and calculating the confidence intervals of WTP estimates (Fig. 2). A raw dataset, including responses to the DCCV question, where rows corresponding to respondents (a respondent per row) and columns corresponding to variables or questions, is prepared using a spreadsheet software package. The raw dataset then is read into R as a data frame. If the format of the raw dataset is a contingency-table of bids and responses to the DCCV question, a function in DCchoice is used to convert the data into a data frame in respondent per row format.

There are six functions for parametric estimation: the three main formats SB-, DB-, and OOHB-DCCV, each with and without a spike. The function for parametric SB-DCCV estimation without a spike uses the function `glm` in the package `stats`, while the remaining five functions depend on the function `optim` in `stats`. Functions for non-parametric estimation approaches for SB-, DB-, and OOHB-DCCV responses follow the Kaplan–Meier–Turnbull (KMT) method (Carson & Hanemann, 2005), which is implemented using the function `icfit` in package `interval` (Fay & Shaw, 2010) internally. A function for Krström’s non-parametric estimation method for SB-DCCV data (Krström, 1990) is also provided.

Outputs from the parametric and non-parametric estimation functions have corresponding object classes such as "sbchoice" and "turnbull.sb". The `print`, `summary`, and `plot` methods for these object classes are provided. Applying the `summary` method to the output from the estimation functions returns WTP estimates.

With parametric estimation the confidence intervals for WTPs are calculated using the functions `krCI` and `bootCI`, where `krCI` implements Krinsky and Robb’s parametric bootstrap approach (Krinsky & Robb, 1986, 1990) and `bootCI` implements the non-parametric bootstrap approach. Both approaches simulate a large number of parameter vectors, construct empirical distributions of the WTP from the replicated

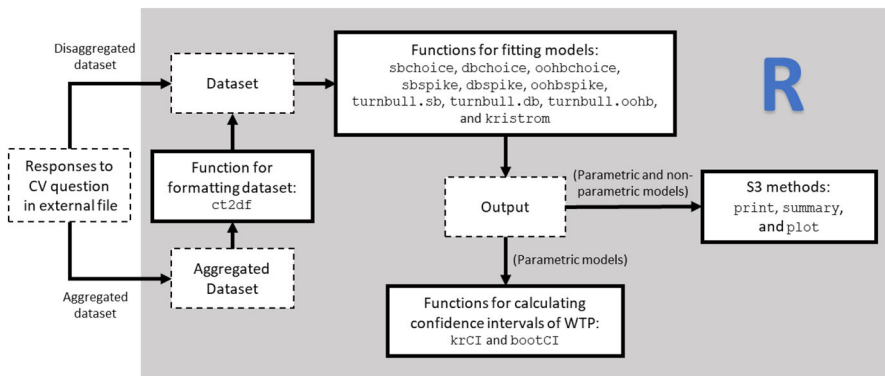


Fig. 2 Workflow of analyzing DCCV data with DCchoice

parameter vectors, and find the confidence intervals of the WTP according to the corresponding percentiles of the empirical distribution. Specifically, the Krinsky and Robb approach replicates parameter vectors from a multivariate normal distribution with the estimated parameters, and variance and covariance matrix of the estimated parameters, while the non-parametric bootstrap approach creates datasets by resampling the original dataset and fits the model to the bootstrapped datasets, resulting in replicated parameter vectors.

### 3.2 DCchoice: from initial to current version

Table 1 provides a reconciliation of the functions in the initial and current versions of DCchoice. The function `ct2df`, which converts a dataset that is in contingency-table format to a dataset in one respondent per row format was unavailable in the initial

**Table 1** Comparison of the initial and current versions of DCchoice

Functionality	Initial version	Current version	Improvement
Preparing dataset			
Formatting dataset	None	<code>ct2df</code>	New function
Fitting models			
Single-bounded dichotomous choice models			
Parametric model with a spike	None	<code>sbspike</code>	New function
Parametric model without a spike	<code>sbchoice</code>	<code>sbchoice</code>	None
Non-parametric model	<code>turnbull.sb</code>	<code>turnbull.sb</code>	Confidence intervals
	<code>kristrom</code>	<code>kristrom</code>	None
Double-bounded dichotomous choice models			
Parametric model with a spike	None	<code>dbspike</code>	New function
Parametric model without a spike	<code>dbchoice</code>	<code>dbchoice</code>	None
Non-parametric model	<code>turnbull.sb</code>	<code>turnbull.sb</code>	Confidence intervals
One-and-one-half bounded dichotomous choice models			
Parametric model with a spike	None	<code>oohbspike</code>	New function
Parametric model without a spike	None	<code>oohbchoice</code>	New function
Non-parametric model	None	<code>turnbull.oohb</code>	New function
Calculating confidence intervals of willingness-to-pay (WTP)			
Parametric approach	<code>krCI</code>	<code>krCI</code>	WTP for a user-defined individual
Non-parametric approach	<code>bootCI</code>	<code>bootCI</code>	WTP for a user-defined individual

version of DCchoice. This function allows complete case study examples derived from papers showing only the bid values and the response patterns in a contingency-table format (e.g., Lim et al., 2014) to be prepared, and so is helpful for creating DCCV examples for teaching.

The initial version of DCchoice had functions for fitting parametric SB- and DB-DCCV models (without a spike), and non-parametric models. The current version has additional functions for fitting OOHB-DCCV parametric and non-parametric models, and spike SB-, DB-, and OOHB-DCCV parametric models. This is a material increase in package functionality. In the current version of DCchoice, confidence intervals for WTP estimates from non-parametric KMT models can also be calculated. The overall improvement in the types of models that can be fitted is one of the main user benefits available due to updating DCchoice.

In the initial version of DCchoice, functions for calculating the WTP and associated confidence intervals assumed only a representative individual, and used mean values of covariates in the estimated model when solving for key values. The current version of DCchoice can calculate the WTP and associated confidence intervals for a user-defined individual, expressed via combinations of different values of the covariates (see Sects. 3.3 and 4 for details). This functionality enables the comparison of WTPs for a target good or service among various types of respondents. As such comparisons are highly valuable in empirical CV studies, this change represents a substantial increase in package functionality.

### 3.3 Main functionalities of DCchoice

This section details the functions for fitting parametric models and calculating the confidence intervals of WTPs, which is typically the primary focus in any DCCV study.

As detailed in Table 2, the ordinal (non-spike) and spike SB-, DB-, and OOHB-DCCV functions have common arguments. Under the general situation of applying DCCV methods, these models can be fitted by setting only three arguments: `dist`,

**Table 2** Arguments of functions for parametric estimation models

Argument	Description
<code>formula</code>	A symbolic description of the model formula
<code>data</code>	A data frame containing the variables in the model formula
<code>subset</code>	An optional vector specifying a subset of observations
<code>na.action</code>	A function which indicates what should happen when the data contains NAs
<code>dist</code>	A character string setting the error distribution in the model, which takes "logistic", "normal", "log-logistic", "log-normal", or "weibull". This is invalid for spike model functions
<code>par</code>	A vector of initial parameters over which the optimization is carried out. This is invalid for the SB-DCCV function

data, and formula. For non-spike models, five error distributions are available: logistic, log-logistic, normal, log-normal, and Weibull. For spike models, the error distribution is limited to the logistic distribution, and so the argument `dist` is not available for spike models. Most empirical studies using spike models assume a logistic distribution.

A data frame containing the variables used for fitting models is assigned to the argument `data`. The model formula, which is assigned to the argument `formula`, must be expressed according to the class "Formula" (Zeileis & Croissant, 2010) that extends the base class "formula". For the non-spike model functions `sbchoice`, `dbchoice`, and `oohbchoice`, the model formula consists of three parts. The first part is the left-hand side of the tilde (`~`) and contains the response variable(s). For the SB format, the response variable is a single variable, such as `R1` in Fig. 1a. For the DB and OOHb formats, there are two response variables, corresponding to the two stages of the DCCV question, such as `R1` and `R2` in Fig. 1bc. The right-hand side of the tilde is divided into two parts with a vertical bar (`|`): the part before the vertical bar contains covariates such as respondent characteristic variables, and the part after the vertical bar contains a single bid variable such as `BID1` for the SB (Fig. 1a), two bid variables such as `BID1` and `BID2` for the DB (Fig. 1b), or `BIDL` and `BIDH` for the OOHb (Fig. 1c). For example, when the covariates are respondent age and gender, the model formula for `sbchoice`, `dbchoice`, and `oohbchoice` are given, respectively, as:

$$\begin{aligned} R1 &\sim \text{AGE} + \text{FEM} | \text{BID1} \\ R1 + R2 &\sim \text{AGE} + \text{FEM} | \text{BID1} + \text{BID2} \\ R1 + R2 &\sim \text{AGE} + \text{FEM} | \text{BIDL} + \text{BIDH}, \end{aligned}$$

where `AGE` is a variable detailing the respondent's age; and `FEM` is a dummy variable taking the value 1 for female respondents and 0 otherwise. When there are no covariates, the corresponding part contains only a value of 1, as follows:

$$\begin{aligned} R1 &\sim 1 | \text{BID1} \\ R1 + R2 &\sim 1 | \text{BID1} + \text{BID2} \\ R1 + R2 &\sim 1 | \text{BIDL} + \text{BIDH}. \end{aligned}$$

For spike models, the part to the left of the tilde is divided into two parts with a vertical bar: the part before the vertical bar contains the response variable(s), and the part after the vertical bar contains a dummy variable showing the response to the spike question regarding whether their WTP is positive or not (i.e., the variable `S` in Fig. 1). For the spike models, the model formula for the models with the covariates mentioned above is modified as follows:

$$\begin{aligned} R1 | S &\sim \text{AGE} + \text{FEM} | \text{BID1} \\ R1 + R2 | S &\sim \text{AGE} + \text{FEM} | \text{BID1} + \text{BID2} \\ R1 + R2 | S &\sim \text{AGE} + \text{FEM} | \text{BIDL} + \text{BIDH}. \end{aligned}$$



**Table 3** Arguments of functions `krCI` and `bootCI`

Argument	Description
<code>obj</code>	An object of S3 class "sbchoice", "dbchoice", "oohbchoice", or "spike"
<code>nsim</code>	An integer value specifying the number of draws of the parameters
<code>nboot</code>	An integer value specifying the number of bootstrap resampling
<code>CI</code>	A value specifying a percentile of the confidence intervals to be estimated
<code>individual</code>	A data frame containing covariates that show an individual of which to estimate WTP

The functions `krCI` and `bootCI` for calculating the confidence intervals of WTPs have common arguments, except for `nsim` and `nboot` (Table 3). The effects of the two arguments are, however, similar: the value assigned to the argument corresponds to the number of simulated WTPs. The output from the two functions contains vectors of simulated WTPs (the length of the vectors corresponds to the value assigned to the argument `nsim` or `nboot`) and a matrix of point estimates and confidence intervals of WTPs. The corresponding `print` method displays only the point estimates and confidence intervals of the WTPs. A significant feature of each function is the calculation of WTPs and confidence intervals for a specific individual, when a data frame of covariates containing values corresponding to the individual is assigned to the argument `individual` (see the example below).

## 4 Simple example

This section demonstrates how WTP for improving the environmental conditions in a location can be estimated with each spike model. We assume that respondents are selected randomly from the residents of the location, and DCCV questions in three formats, SB, DB, and OOHb, are used. The datasets are modified versions of `CarsonSB`, `CarsonDB`, and `oohbsyn` that are available with `DCchoice`. Code chunks for creating the datasets, model fitting, and estimating WTPs are available in the online supplementary file. Additional examples are available via a free web tutorial (Aizaki & Fogarty, 2019) and the help in `DCchoice` (Nakatani et al., 2021). The variables used in the example are the same as those defined in the previous sections, however, bid variables are typed in lowercase letters, such as `bid1`.

Figure 3 displays the code chunks for the SB spike model and the outputs. After fitting the model (line 1), executing the `summary` method with the output from `sbspike` (line 2) returns the estimates and the corresponding statistics, summary statistics of the model, and point estimates of mean and median WTPs for a representative respondent (lines 3–31). The function `spikeCoef` calculates and returns a spike coefficient using the output from the spike model function (lines 32–34). Note, a value of one minus the spike coefficient corresponds closely to the y-axis intercept (Kriström, 1997). Applying the function `krCI` to the output from the spike

```

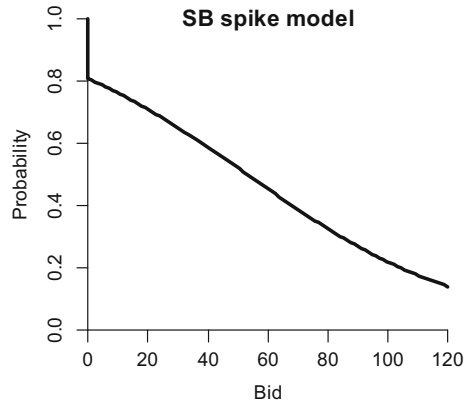
1: > SBspk.out <- sbspike(R1 | S ~ AGE + FEM | bid1, data = SBspk.dat)
2: > summary(SBspk.out)
3:
4: Call:
5: sbspike(formula = R1 | S ~ AGE + FEM | bid1, data = SBspk.dat)
6:
7: Formula:
8: R1 | S ~ AGE + FEM | bid1
9:
10: Coefficients:
11:             Estimate Std. Error z value Pr(>|z|)
12: (Intercept)  1.300584   0.214176   6.0725 <2e-16 ***
13: AGE          0.002729   0.004351   0.6271  0.5306
14: FEM          0.015955   0.124983   0.1277  0.8984
15: bid1        -0.026998   0.001178  -22.9151 <2e-16 ***
16: ---
17: Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
18:
19: Distribution: spike logistic
20: Number of Obs.: 1043
21: Log-likelihood: -1112.620001
22:
23: LR statistic: 0.404 on 2 DF, p-value: 0.817
24: AIC: 2233.240002 , BIC: 2253.039428
25:
26: Iterations: 65 9
27: Convergence: TRUE
28:
29: WTP estimates:
30: Mean : 60.8989
31: Median: 52.94869
32: > spikeCoef(SBspk.out)
33:             Estimate Std. Error z value Pr(>|z|)
34: Spike 0.1931711 0.01132701 17.05402 0
35: > krCI(SBspk.out)
36: the Krinsky and Robb simulated confidence intervals
37:             Estimate LB UB
38: Mean      60.899 56.410 66.123
39: Median    52.949 47.918 58.165
40: > krCI(SBspk.out, individual = data.frame(AGE = 30, FEM = 1))
41: the Krinsky and Robb simulated confidence intervals
42:             Estimate LB UB
43: Mean      59.971 52.533 67.228
44: Median    51.795 43.126 59.581
45: > krCI(SBspk.out, individual = data.frame(AGE = 60, FEM = 0))
46: the Krinsky and Robb simulated confidence intervals
47:             Estimate LB UB
48: Mean      61.941 55.614 69.230
49: Median    54.236 46.428 62.459
50: > plot(SBspk.out, main = "SB spike model")

```

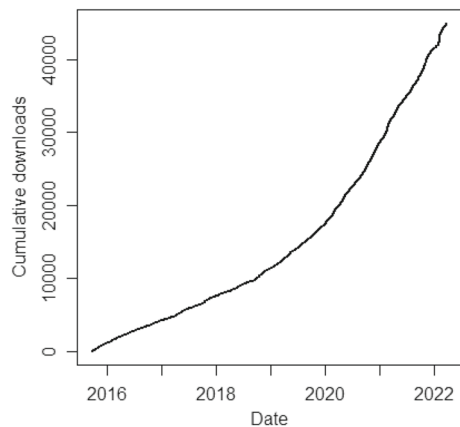
**Fig. 3** Fitting a SB spike model, calculating a spike coefficient, estimating mean and median WTPs, and plotting the fitted model

model function under three different settings of the arguments calculates the mean and median WTPs and their confidence intervals for the representative respondent (lines 35–39), female respondent of 30 years old (lines 40–44), and male respondent of 60 years old (lines 45–49), respectively. The `plot` method draws the probability of paying the assigned bid for the representative respondent (line 50), ranging from 0 to the maximum bid value (Fig. 4). For the results of the DB and OOHb spike models, execute the supplementary file in R.

**Fig. 4** Probability of paying the bid predicted from the fitted SB spike model



**Fig. 5** Cumulative downloads of DCchoice via the RStudio CRAN mirror until March 31, 2022. Data obtained using the package cranlogs (Csardi, 2019)



## 5 Impact

In response to user requests, DCchoice has been substantially updated and improved since its first release in 2014. The major updates are summarized as follows: (1) addition of a function for fitting OOHb-DCCV models; (2) addition of functions for fitting SB, DB, and OOHb spike models; and (3) the ability to calculate WTP and confidence intervals for respondents with different characteristics. There is an additional package *mded* (Aizaki, 2015) that is available to test whether a WTP differs from another on the basis of empirical distributions of the WTPs (Poe et al., 1997, 2005) that are provided using *krCI* and *bootCI* in DCchoice (Lorber et al., 2021). Accordingly, using DCchoice with other CRAN packages makes R a free platform for implementing comprehensive DCCV analysis. A CRAN package *RcmdrPlugin.DCCV* (Aizaki, 2021), which is a plug-in package for R Commander (Fox, 2005, 2017; Fox & Bouchet-Valat, 2020), has been developed to enable the use of some parametric functions in DCchoice without writing R code chunks. DCchoice also serves as a base package for developing functions for fitting new models (Qiu et al., 2020) and calculating the confidence

intervals of WTPs more quickly (Howard, 2018). Furthermore, a free web tutorial for DCchoice (Aizaki & Fogarty, 2019) is provided under the Creative Commons license. Thus, DCchoice plays a core role in the activities that induce new material for DCCV applications. The number of DCchoice downloads from the RStudio CRAN mirror continues to increase (Fig. 5) and the package is increasingly used in published research (Lin et al., 2017; Lorber et al., 2021; Morello et al., 2019; Mostafa, 2016; Tokunaga et al., 2020; Whitehead, 2018), education (Polomé, 2020), and practice (Turpie et al., 2017). These applications indicate the impact of DCchoice in various fields.

## 6 Conclusion

DCchoice is easily installed from CRAN. The functions in DCchoice are designed to be accessible to novice R users, via an R Commander plug-in. For more advanced applications, the model formula differs only slightly from that used in linear regression models, such as the formula structure used for the `lm` function in the stats package, and is consistent with many other extension packages. DCchoice has a wide range of applicability in various fields and is updated regularly. A web tutorial also serves to provide additional user support, and assist with the adoption of DCCV methods.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s42081-022-00171-1>.

**Funding** This study was supported by JSPS KAKENHI Grant Numbers JP20K06251 and JP21K18227.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aizaki, H. (2015). mded: Measuring the difference between two empirical distributions. R package version 0.1-2. <https://CRAN.R-project.org/package=mded>
- Aizaki, H. (2021). RcmdrPlugin.DCCV: R Commander plug-in for dichotomous choice contingent valuation. R package version 0.1-1. <https://CRAN.R-project.org/package=RcmdrPlugin.DCCV>
- Aizaki, H., & Fogarty, J. (2019). An illustrative example of contingent valuation. In NMVR Team (Ed.), *Non-market valuation with R*. Retrieved November 24, 2021, from <http://lab.agr.hokudai.ac.jp/nmvr/>
- Aizaki, H., Nakatani, T., & Sato, K. (2014). *Stated preference methods using R*. CRC Press.
- Bishop, R. C., & Heberlein, T. A. (1979). Measuring values of extramarket goods: Are indirect measures biased? *American Journal of Agricultural Economics*, 61(5), 926–930. <https://doi.org/10.2307/3180348>

- Carmichael, I., & Marron, J. S. (2018). Data science vs. statistics: Two cultures? *Japanese Journal of Statistics and Data Science*, 1, 117–138. <https://doi.org/10.1007/s42081-018-0009-3>
- Carson, R. T. (1985). *Three essays on contingent valuation*. Dissertation, University of California Berkeley.
- Carson, R. T. (2012). *Contingent valuation: A comprehensive bibliography and history*. Edward Elgar.
- Carson, R. T., & Hanemann, W. M. (2005). Contingent valuation. In K.-G. Mäler, & J. R. Vincent (Eds.), *Handbook of environmental economics* (Vol. 2, pp. 821–936). Elsevier.
- Chambers, J. M. (2020). S, R, and data science. *Proceedings of the ACM on Programming Languages*, 4(HOPL), 1–17. <https://doi.org/10.1145/3386334>.
- Cooper, J. C., Hanemann, M., & Signorello, G. (2002). One-and-one-half-bound dichotomous choice contingent valuation. *The Review of Economics and Statistics*, 84, 742–750. <https://doi.org/10.1162/003465302760556549>
- Csardi, G. (2019). cranlogs: Download logs from the ‘RStudio’ ‘CRAN’ mirror. R package version 2.1.1. <https://CRAN.R-project.org/package=cranlogs>
- Donoho, D. (2017). 50 years of data science. *Journal of Computational and Graphical Statistics*, 26(4), 745–766. <https://doi.org/10.1080/10618600.2017.1384734>
- Fay, M. P., & Shaw, P. A. (2010). Exact and asymptotic weighted logrank tests for interval censored data: The interval R package. *Journal of Statistical Software*, 36(2), 1–34. <https://doi.org/10.18637/jss.v036.i02>
- Fox, J. (2005). The R Commander: A basic-statistics graphical user interface to R. *Journal of Statistical Software*, 14(9), 1–42. <https://doi.org/10.18637/jss.v014.i09>
- Fox, J. (2017). *Using the R Commander: A point-and-click interface for R*. Chapman and Hall/CRC Press.
- Fox, J., & Bouchet-Valat, M. (2020). Rcmdr: R Commander. R package version 2.7-1. <https://CRAN.R-project.org/package=Rcmdr>
- Gracia, L. Y., & Cerda, A. A. (2020). Contingent assessment of the COVID-19 vaccine. *Vaccine*, 38, 5424–5429. <https://doi.org/10.1016/j.vaccine.2020.06.068>
- Hanemann, W. M. (1985). Some issues in continuous-and discrete-response contingent valuation studies. *Northeastern Journal of Agricultural Economics*, 14, 5–13. <https://doi.org/10.1017/S0899367X00000702>
- Hornik, K. (2012). The comprehensive R archive network. *WIREs Computational Statistics*, 4, 394–398. <https://doi.org/10.1002/wics.1212>
- Howard, J. (2018). *Parallelized implementation of bootCI for DCchoice*. Retrieved November 24, 2021, from <https://jameshoward.us/2018/11/11/parallelized-implementation-of-bootci-for-dcchoice/>
- Howison, J., Deelman, E., McLennan, M. J., Ferreira de Silva, R., & Herbsleb, J. D. (2015). Understanding the scientific software ecosystem and its impact: Current and future measures. *Research Evaluation*, 24, 454–470. <https://doi.org/10.1093/reseval/rvv014>
- Krinsky, I., & Robb, A. L. (1986). On approximating the statistical properties of elasticities. *The Review of Economics and Statistics*, 68, 715–719. <https://doi.org/10.2307/1924536>
- Krinsky, I., & Robb, A. L. (1990). On approximating the statistical properties of elasticities: A correction. *The Review of Economics and Statistics*, 72, 189–190. <https://doi.org/10.2307/2109761>
- Krström, B. (1990). A non-parametric approach to the estimation of welfare measures in discrete response valuation studies. *Land Economics*, 66(2), 135–139. <https://doi.org/10.2307/3146363>
- Krström, B. (1997). Spike models in contingent valuation. *American Journal of Agricultural Economics*, 79, 1013–1023. <https://doi.org/10.2307/1244440>
- Kwak, S.-J., Yoo, S.-H., & Kim, C.-S. (2013). Measuring the willingness to pay for tap water quality improvements: Results of a contingent valuation survey in Pusan. *Water*, 5, 1638–1652. <https://doi.org/10.3390/w5041638>
- Lim, K.-M., Lim, S.-Y., & Yoo, S.-H. (2014). Estimating the economic value of residential electricity use in the Republic of Korea using contingent valuation. *Energy*, 64, 601–606. <https://doi.org/10.1016/j.energy.2013.11.016>
- Lin, Y., Wijedasa, L. S., & Chisholm, R. A. (2017). Singapore’s willingness to pay for mitigation of transboundary forest-fire haze from Indonesia. *Environmental Research Letters*, 12, 024017. <https://doi.org/10.1088/1748-9326/aa5cf6>
- Lorber, C., Dittrich, R., Jones, S., & Junge, A. (2021). Is hiking worth it? A contingent valuation case study of Multnomah Falls Oregon. *Forest Policy and Economics*, 128, 102471. <https://doi.org/10.1016/j.forpol.2021.102471>
- Loureiro, M. L., & Hine, S. (2004). Preferences and willingness to pay for GM labeling policies. *Food Policy*, 29, 467–483. <https://doi.org/10.1016/j.foodpol.2004.07.001>

- Mora-Cantalops, M., Sicilia, M. -Á., Gracia-Barriocanal, E., & Sánchez-Alonso, S. (2020). Evolution and prospects of the Comprehensive R Archive Network (CRAN) package ecosystem. *Journal of Software: Evolution and Process*, 32, e2270. <https://doi.org/10.1002/smr.2270>
- Morello, T., Martino, S., Duarte, A. F., Anderson, L., Davis, K. J., Silva, S., & Bateman, I. J. (2019). Fire, tractors, and health in the Amazon: A cost-benefit analysis of fire policy. *Land Economics*, 95(3), 409–434. <https://doi.org/10.3368/le.95.3.409>
- Mostafa, M. M. (2016). Egyptian consumers' willingness to pay for carbon-labeled products: A contingent valuation analysis of socio-economic factors. *Journal of Cleaner Production*, 135, 821–828. <https://doi.org/10.1016/j.jclepro.2016.06.168>
- Nakatani, T., Aizaki, H., & Sato, K. (2021). DCchoice: An R package for analyzing dichotomous choice contingent valuation data. R package version 0.1.0. <https://CRAN.R-project.org/package=DCchoice>
- Poe, G. L., Giraud, K. L., & Loomis, J. B. (2005). Computational methods for measuring the difference of empirical distributions. *American Journal of Agricultural Economics*, 87, 353–365. <https://doi.org/10.1111/j.1467-8276.2005.00727.x>
- Poe, G. L., Welsh, M. P., & Champ, P. A. (1997). Measuring the difference in mean willingness to pay when dichotomous choice contingent valuation responses are not independent. *Land Economics*, 73, 255–267. <https://doi.org/10.2307/3147286>
- Polomé, P. (2020). *Research in Applied Econometrics 2020–21*. Retrieved November 24, 2021, from [http://risques-environnement.universite-lyon.fr/IMG/pdf/rae\\_2\\_cv-2.pdf](http://risques-environnement.universite-lyon.fr/IMG/pdf/rae_2_cv-2.pdf)
- Qiu, R. T. R., Park, J., Li, S. N., & Song, H. (2020). Social costs of tourism during the COVID-19 pandemic. *Annals of Tourism Research*, 84, 102994. <https://doi.org/10.1016/j.annals.2020.102994>
- R Core Team. (2021). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org>
- Tokunaga, K., Sugino, H., Nomura, H., & Michida, Y. (2020). Norms and the willingness to pay for coastal ecosystem restoration: A case of the Tokyo Bay intertidal flats. *Ecological Economics*, 169, 106423. <https://doi.org/10.1016/j.ecolecon.2019.106423>
- Turpie, J., Brick, K., Letley, G., & Maclaren, C. (2017). Potential for the use of a Payment-for-Ecosystem Services system in Namibia's Communal Conservancies. In Ministry of Environment and Tourism (Ed.), *Namibia's national TEEB study: The development of strategies to maintain and enhance the protection of ecosystem services in Namibia's state, communal and freehold lands, Vol III*. Report prepared by Anchor Environmental Consultants and Namibia Nature Foundation for the GIZ, on behalf of Namibia's Department of Environmental Affairs (pp. 1–101). Retrieved November 24, 2021, from <https://resmob.org/wp-content/uploads/2019/03/TEEB-Study-Vol-III-Towards-a-system-of-Payment-for-Ecosystem-services-in-Namibias-Communal-Conservancies-FINAL-WEBSITE-VERSION.pdf>
- Verbooy, K., Hoefman, R., van Exel, J., & Brouwer, W. (2018). Time is money: Investigating the value of leisure time and unpaid work. *Value in Health*, 21, 1428–1436. <https://doi.org/10.1016/j.jval.2018.04.1828>
- Whitehead, J. C. (1995). Willingness to pay for quality improvements: Comparative statics and interpretation of contingent valuation results. *Land Economics*, 71(2), 207–215. <https://doi.org/10.2307/3146501>
- Whitehead, J. C. (2015). Albemarle-Pamlico Sounds revealed and stated preference data. *Data in Brief*, 3, 90–94. <https://doi.org/10.1016/j.dib.2015.01.006>
- Whitehead, J. C. (2018). A comment on “Three reasons to use annual payments in contingent valuation.” *Journal of Environmental Economics and Management*, 88, 486–488. <https://doi.org/10.1016/j.jeem.2016.09.004>
- Yoo, S.-H., & Kwak, S.-J. (2002). Using a spike model to deal with zero response data from double bounded dichotomous choice contingent valuation surveys. *Applied Economics Letters*, 9, 929–932. <https://doi.org/10.1080/13504850210139378>
- Zeileis, A., & Croissant, Y. (2010). Extended model formulas in R: Multiple parts and multiple responses. *Journal of Statistical Software*, 34(1), 1–13. <https://doi.org/10.18637/jss.v034.i01>