

**Watson V, Becker F, de Bekker-Grob E.**

**[Discrete choice experiment response rates: a meta-analysis.](#)**

***Health Economics 2016***

**DOI: <http://dx.doi.org/10.1002/hec.3354>**

**Copyright:**

This is the peer reviewed version of the following article: Watson V, Becker F, de Bekker-Grob E. [Discrete choice experiment response rates: a meta-analysis.](#) *Health Economics* 2016, which has been published in final form at <http://dx.doi.org/10.1002/hec.3354>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

**DOI link to article:**

<http://dx.doi.org/10.1002/hec.3354>

**Date deposited:**

19/05/2016

**Embargo release date:**

27 April 2017

## **Discrete choice experiment response rates: A meta-analysis**

*Keywords:* discrete choice experiments, response rates, survey design, meta regression analysis

## **Introduction**

Discrete choice experiments (DCEs) are used to elicit patients', the public's or healthcare professionals' preferences for health and healthcare (de Bekker-Grob et al, 2012). DCE surveys are prone to coverage, sampling, non-response, and measurement errors (Groves, 1989). Much DCE research focusses on minimising measurement error. For example, by reducing hypothetical bias (Özdemir et al, 2009), improving question formats (Lancsar et al, 2013) or understanding how DCE study design affects responses (Bech et al, 2011; Ryan and Wordsworth, 2000). Research on non-response error in healthcare DCEs is lacking.

Non-response error occurs when sampled individuals who respond to the survey differ from those who do not respond. Maximising response rates may minimise non-response error. In this study, we use meta regression analysis (MRA) to test how study design and other factors within the researchers' control affect DCE survey response rates. We analyse a dataset of DCE studies in healthcare published between 2001 and 2011. In doing so, we combine evidence from across studies to understand the factors that affect DCE response rates.

## **Methods**

MRA is a statistical method that combines the results of different independent studies (Glass, 1976; Huque 1988). MRA overcomes a single study's limitations by considering heterogeneity between study results and provides insight into the variation of findings. MRA has been used to measure time preference rates (Asenso-Boadi et al, 2008) and

value statistical lives (Johnson et al, 1997). MRA has been used to pool willingness to pay estimates across stated preference studies that valued similar goods (Lindhjem and Navrud, 2008; Barrio and Loureiro, 2008). List and Gallet (2001), Little and Berrens (2004) and Murphy et al (2005) use MRA to investigate how study design affects the magnitude of hypothetical bias across contingent valuation studies. Our paper is similar in spirit to these studies.

The first stage of a MRA identifies the studies to be included as observations in the data set. We used an existing systematic review for 2001 to 2008 (de Bekker-Grob et al, 2012) combined with a new systematic review to the end of 2011. Our systematic review used the same methods as Ryan and Gerard (2003) and de Bekker-Grob et al (2012) (for more details see the online appendix). Following de Bekker-Grob et al (2012) we included studies in which respondents completed choice-based DCE tasks<sup>1</sup> that were published in English as a full-text article.

Some studies published more than one journal article from the same data set. For example, a study may publish a paper that reports the results in a clinical journal and a separate paper that reports the results of methodological work or a new econometric estimator in a (health) economics journal. In such cases, we included the study only once in our data set and took the first published paper as the publication date. Some publications report the results from more than one study. For example, for our purposes the same questionnaire administered to different populations (patients and/or the general public and/or healthcare professionals) is classed as different studies. Similarly, slightly different questionnaires administered to subgroups of the same population are considered different studies. An example of this is Bech et al (2011) who compared

---

<sup>1</sup> We exclude ranking and rating conjoint tasks and adaptive conjoint analysis.

response rate across three subgroups who each received a DCE with 5, 9 or 17 choice sets, respectively. In those cases, we included the studies as separate observations, but controlled for the interdependence of these studies in our analysis (see analysis section).

We motivate our MRA of response rates using Dillman et al (2009)'s social exchange theory of survey response. This assumes that individuals respond to a survey when the perceived benefits of responding outweigh the costs. We extracted information from the eligible studies about variables related to the costs and perceived benefits of response and used these to explain survey response rates in the MRA. If information was not included in the article, we emailed the corresponding author to request it.

The perceived benefits of responding differ depending on the sample who are asked to respond. For example, the perceived benefits to patients of responding to a DCE that elicits treatment preferences for their illness may include future treatment improvements for themselves and/or others with the same illness. We extracted data on the sample population, and we hypothesised patients/carers and healthcare professionals will perceive higher benefits from responding to healthcare DCEs than the general public. Therefore, surveys of patients or healthcare professional should attain higher response rates. The topic of the survey will also affect the benefits of response. For example, the severity or prevalence of the health condition that is the survey topic may affect response rates. We extracted data on the health condition that is the topic of the survey. Given the perception that individuals place more importance on treating cancer than other diseases (Linley and Hughes, 2013), we hypothesised that DCE surveys about cancer-related topics will have higher response rates than other studies.

The perceived costs of survey response depend on the survey's design and the cognitive burden of completing it (Presser et al, 2004). Previous research has suggested that the

cognitive burden of responding to a DCE survey increases with: the number of attributes describing the good, number of choice set alternatives, the number of choice sets to be answered<sup>2</sup> (Mazzota and Opaluch, 1995; Swait and Adamowicz, 2001; DeShazo and Fermo, 2002), and the inclusion of a cost attribute (Pedersen et al, 2011; Bryan et al, 2001). Evidence from the decision theory, contingent valuation and DCE literatures shows that individuals have difficulty to understand risks and probabilities (Hammit and Graham, 1999; Lloyd, 2001). Opt-out alternatives were introduced to DCEs to make choices more realistic and therefore, easier to answer (Hanley et al, 2003).

We extracted data on eight variables affecting cognitive burden: number of attributes; number of choice set alternatives; the number of choice sets to be answered; includes an opt-out alternative; includes a cost attribute; includes a risk attribute; elicits time preferences; the same hypothetical alternative appears in all choice sets (constant comparator experimental design). We hypothesised that the number of attributes, alternatives and choice sets, and the inclusion of risk or cost attributes will decrease survey response rates, and that including an opt-out will increase response rates. We also extracted data on general study characteristics that may affect response rates: whether non-respondents were sent reminders; the survey mode; the country in which the data were collected and the publication year as a proxy for the date the study was carried out.

For the survey mode, we distinguished between self-complete and interviewer administered surveys. Within self-complete surveys we distinguished between postal surveys, online surveys and those completed by patients in a clinic waiting room or

---

<sup>2</sup> Although, Bech et al (2011) find the number of choice sets has no effect on response rates to a DCE.

other central location. The explanatory variables may affect response rates across survey modes differently for many reasons. For instance, interview administered questionnaires provide respondents with support when they complete questionnaires. At the outset of online surveys the task ahead of respondents is often unclear. A separate MRA should be estimated for each mode.

### *Analysis*

The dependent variable in the MRA is the response rate from  $i$  studies,  $y_i$ . The explanatory variables are grouped into three categories – those related to cognitive burden,  $X_{COG}$ , those related to benefit of response,  $X_{BEN}$ , and other study characteristics,  $X_{STU}$ . The estimation model can be expressed as:

$$y_i = \alpha + X_{COGi}\beta_{COG} + X_{BENi}\beta_{BEN} + X_{STUi}\beta_{STU} + \varepsilon_i \quad (1)$$

$\alpha$  is a constant term, the  $\beta$  vectors are the coefficients to be estimated for the explanatory variables, and  $\varepsilon$  is the residuals. The vectors of explanatory variables comprise of dummy variables that summarise the study characteristics (see Table 1). We estimated equation (1) using a linear regression estimated with ordinary least squares (OLS). Each study counts equally in the data. The error terms will be correlated for those studies that use the same questionnaire for data collection in different populations or marginally different questionnaires for sub-samples of the same population. We estimate equation (1) using a cluster robust estimator of variance that relaxes the assumption that the error term is independent across all studies. We conduct a series of robustness tests that are reported in the online appendix.

## Results

The data set available for the MRA included 114 unique studies for 2001-2008. The systematic literature review for 2009-2011 identified 371 papers of which 306 met the inclusion criteria. From these 144 unique studies were identified. For the period 2001-2011, we have a total of 258 studies. When we excluded studies with missing data on key variables of interest, such as the response rate and use of reminders, the sample size for analysis reduced to 132 studies. However, across all modes except postal surveys the sample sizes are too small to estimate the MRA (Figure 1).

Sample sizes for all modes except postal studies are small because fewer studies use these modes and the modes have a high proportion of missing response rate data. Our correspondence with authors revealed that response rates were often missing because researchers (ourselves included) did not record the number of invitations to participate made at central locations or sent to potential online respondents. Restricting analysis to postal surveys reduced the sample size to 64 (Figure 1). The characteristics of all postal studies (including those with missing values) are presented in Table 1. The response rate is missing in 13% of eligible postal surveys and information on the use of reminders is missing in 17%.

Table 2 presents the results from the MRA. The model has a good fit with an  $R^2=0.61$ . With respect to the effect of cognitive burden including more attributes decreases response rates. Increasing the number of attributes from between two and four attributes to five attributes, six attributes or seven or more attributes decreases response rates. Including an opt-out increases response rates. We find a mixed effect of the number of

choice sets: compared to 8 choice sets, response rates are higher for DCEs with fewer (3-7) and more (>8) choice sets. Contrary to expectations, studies with risk attributes have higher response rates. The inclusion of a cost attribute, time preferences or using a constant comparator design does not significantly affect survey response rates.

The perceived benefit of response increases response rates. Surveys of patients or health care professionals rather than the general public have higher response rates. For other study variables and controls, reminders increase responses rates and response rates are lower in the UK compared to the rest of the world.

## **Discussion**

This is the first study to use MRA to model DCE response rates. Our results are consistent with a social exchange theory of survey response. We find that the cognitive burden of the survey reduces response rates and the perceived benefits increase response rates. Our findings suggest that researchers who want to minimise the cognitive burden should focus on the number of attributes included in the DCE choice sets. We find that including a cost attribute does not affect response rates. We find counter intuitive results for both the number of choice sets and inclusion of risk attributes. We suggest that these variables may be confounded with survey quality differences across studies. Researchers who include risk attributes or a high number of choice sets may compensate by reducing cognitive burden in other (unmeasured) aspects of the study design. It is difficult to control for study quality in a systematic way. We extracted all study data that are regularly reported. But few studies include detailed information about the entire survey and do not provide the survey as an appendix, routinely. We join

other authors in calling for authors to report routinely more study design details (Harrison et al, 2014; Lancsar and Louviere, 2008). Furthermore, studies with 8 choice sets are likely to be early DCE applications with generally less well-designed surveys.

There are three caveats in the MRA results. 1. The study quality will vary (Nelson and Kennedy, 2009), and we cannot control for this except through the ability of our regressors to capture it. 2. Quality assessment checklists for MRA of RCTs, such as Evers et al (2005), are not applicable to DCEs, and existing DCE checklists are not quality measures. 3. Publication bias is present in all MRA (Nelson and Kennedy, 2009). In our case selection bias is present if studies with low response rates are less likely to be published. We focus on response rates for postal studies because response rates were often missing for other modes. Recent DCE checklists encourage the reporting of response rates (Lancsar and Louviere, 2008). When more data are available, future work should consider if our results hold across modes.

## References

- Asenso-Boadi, F. Peters, T. and Coast, J. (2008). Exploring differences in empirical time preference rates for health: An application of meta-regression. *Health Economics*, 17, 235-248.
- Baker, R. Blumberg, S. Brick, J. Couper, M. Courtright M. Dennis, J. Dillman, D. Frankel, M. Garland, P. Groves, R. Kennedy, C. Krosnick, J. and Lavrakas, P. (2010). Research Synthesis: AAPOR report on online panels. *Public Opinion Quarterly*, 74, 711-781.
- Barrio, M. and Loureiro, M. (2010). A meta-analysis of contingent valuation forest studies. *Ecological Economics*, 69, 1023-1030.
- Baum, C. (2008). Stata tip 63: Modelling proportions. *The Stata Journal*, 8(2), 299-303.
- Bech, M. Kjaer, T. and Lauridsen, J. (2011). Does the number of choice sets matter? Results from a web survey applying a discrete choice experiment. *Health Economics*, 20, 273-286.
- Bryan, S. Buxton, M. Sheldon, R. and Grant, A. (2001). Magnetic resonance imaging for the investigation of knee injuries: an investigation of preferences. *Health Economics*, 7, 595-603.
- de Bekker-Grob, E. Ryan, M. and Gerard, K. (2012). Discrete choice experiments in health economics: a review of the literature. *Health Economics*, 21, 145-172.

- DeShazo, J.R. and Fermo G. (2002). Designing choice sets for stated preference methods: The effects of complexity on choice consistency. *Journal of Environmental Economics and Management*, 44, 123-43.
- Dillman, D. Smyth, JD. Christian, LM. (2009). *Internet, mail and mixed-mode surveys: The tailored design method*. New Jersey, Wiley.
- Evers, S. Goossens, M. de Vet, H. van Tulder, M. and Ament, A. (2005). Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. *Int J Technol Assess Health Care*. 21: 240–245.
- Glass, G. (1976). Primary, secondary and meta-analysis of research. *Education Research*, 10, 3-8.
- Groves, R.M. (1989) *Survey Errors and Survey Costs*. Wiley: Philidephia
- Hanley, N. Ryan, M. & Wright, R. (2003). Estimating the monetary value of health care. *Health Economics*, 12: 3-16.
- Harrison M, Rigby D, Vass C, Flynn T, Louviere J, Payne K. (2014). Risk as an attribute in discrete choice experiments: A systematic review of the literature. *Patient*, 7, 151-170.
- Huque, MF. (1988). Experiences with meta-analysis in NDA submissions. *Proceedings of the Biopharmaceutical Section of the American Statistical Association*, 2:28-33.
- Johnson, FR. Fries, E. and Banzaf, HS. (1997). Valuing morbidity: An interation of the willingness-to-pay and health-status index. *Journal of Health Economics*, 16, 641-665.
- Lancsar, E. Louviere, J. Donaldson, C. Currie, G. and Burgess, L. (2013). Best Worst Discrete Choice Experiments in Health: Methods and an Application. *Social Science and Medicine*, 76, 74-82.
- Lancsar, E. Louviere, J. (2008). Conducting discrete choice experiments to inform healthcare decision making: A user's guide. *Pharmacoeconomics*, 26, 661-677.
- Lindjem, H. Navrud, S. (2008). How reliable are meta-analyses for international benefit transfers? *Ecological Economics*, 66, 425-435.
- List, J. and Gallet, C. (2001). What Experimental Protocol Influence Disparities between Actual and Hypothetical Stated Values? *Environmental and Resource Economics*, 20, 241-254
- Little, J. and Berrens, R. (2004). Explaining disparities between actual and hypothetical stated values: further investigation using meta-analysis. *Economics Bulletin*, 3, 1-13.
- Lloyd, A. (2001). The extent of patients' understanding of the risk of treatments. *Quality in Health Care*, 10 suppl 1: i14-i18.
- Mazzotta, M. and Opaluch, J. (1995). Decision making when choices are complex: A test of Heiner's hypothesis. *Land Economics*, 71, 500--515.
- Murphy, J. Allen, P. Stevens, T. and Weatherhead D. (2005). A Meta-analysis of Hypothetical Bias in Stated Preference Valuation. *Environmental and Resource Economics*, 30. 313-325.
- Nelson, J. and Kennedy, P. (2011). The use (and abuse) of meta-analysis in environmental and natural resource economics: an assessment. *Environmental and Resource Economics*, 42, 345-377.

- Özdemir, S. Johnson, FR. and Hauber, AB. (2009). Hypothetical Bias, Cheap Talk, and Stated Willingness to Pay for Health Care. *Journal of Health Economics*. 28, 894-901.
- Papke, LE. and Wooldridge, JM. (1996). Econometric methods for fractional response variables with an application to 401(K) plan participation rates. *Journal of Applied Econometrics*. 11, 619–632.
- Pedersen, L. Kjaer, T. Kragstrup, J. and Gyrd-Hansen, D. (2011). Does the inclusion of a cost attribute in forced and unforced choices matter? Results from a web survey applying the discrete choice experiment. *Journal of Choice Modelling*. 4, 88-109.
- Presser, S. Rothgeb, J. Couper, M. Lessler, J. Martin, E. Martin, J. and Singer, L. (2004). *Methods for testing and evaluating survey questionnaires*. Wiley.
- Ryan, M. and Gerard, K. (2003). Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Applied Health Economics Health Policy*. 2, 55–64.
- Ryan, M. and Wordsworth, S. (2000). Sensitivity of willingness to pay estimates to the level of attributes in discrete choice experiments. *Scottish Journal of Political Economy* 47, 504-524.
- Swait, J. and Adamowicz, W. (2001). The influence of task complexity on consumer choice: A latent class model of decision strategy switching. *Journal of Consumer Research*. 28, 135-48.

**Table 1: Characteristics of included studies (postal surveys only including observations with missing data, n=64)**

Variables		Summary statistics		
		%	Mean (Standard deviation)	% missing
Response rate			50.44 (19.46)	12.90
<i>Cognitive burden</i>				
Number of attributes			5.30 (1.99)	1.08
	2-4 attributes*	31.18		
	5 attributes	22.58		
	6 attributes	27.96		
	7+ attributes	17.20		
Number of alternatives			2.56 (1.74)	2.15
	2 alternatives*	63.44		
	3+ alternatives	34.41		
Opt-out option		31.18		2.15
Number of choice sets			12.66 (17.32)	3.23
	No. of choice sets: 3-7	23.66		
	No. of choice sets: 8*	17.20		
	No. of choice sets: 9-15	31.18		
	No. of choice sets: 16	15.05		
	No. of choice sets: 17+	9.68		
Risk attribute		38.71		1.08
Price proxy		43.01		1.08
Time preferences		3.23		0
Constant comparator		20.43		2.15
<i>Perceived benefit of response</i>				
Population	General public*	20.43		0
	Patients and carers	51.61		
	Health care professionals	21.51		
	Other	6.45		
Disease	Cancer*	21.51		0
	Other	78.49		
Perspective	Own*	79.57		1.08
	Other	19.35		
<i>Other study characteristics</i>				
Reminder	0*	17.20		17.20
	1+	65.59		
Country	UK*	49.46		1.08
	Other	49.46		
Year of publication	2001-2008*	56.99		0
	2009-2011 (2012 epub)	43.01		

\* indicates the reference group for the analysis.

**Table 2: Meta analysis results using linear regression model – OLS (n=64)**

Variables		Coefficient $\beta$	Significance	Robust SE
<i>Cognitive burden</i>				
Number of attributes	2-4 attributes			
	5 attributes	-0.13	*	0.07
	6 attributes	-0.20	***	0.09
	7+ attributes	-0.26	***	0.09
Number of alternatives	2 alternatives			
	3+ alternatives	-0.21		0.15
Opt-out option		0.18	**	0.11
Number of choice sets	No. of choice sets: 3-7	0.22	***	0.09
	No. of choice sets: 8			
	No. of choice sets: 9-15	0.12	*	0.06
	No. of choice sets: 16	0.11	*	0.06
	No. of choice sets: 17+	0.12	***	0.05
Risk attribute		0.10	**	0.05
Price proxy		-0.08		0.06
Time preferences		-0.02		0.08
Constant comparator		0.01		0.06
<i>Perceived benefit of response</i>				
Population	General public			
	Patients and carers	0.25	***	0.05
	Health care professionals	0.17	*	0.10
	Other	0.06		0.10
Disease	Cancer	-0.08		0.10
	Other			
Perspective	Own	-0.01		0.06
	Other			
<i>Other study characteristics</i>				
Reminder (1=yes; 0=no)		0.14	*	0.09
Country	UK	-0.12	**	0.06
	Other			
Year of publication	2001-2008			
	2009-2011	-0.11		0.10
Constant		0.51	***	0.12
R-squared		0.61		

\*\*\* at 1% level, \*\* at 5% level, \* at 10% level

## **Online appendix**

### **Systematic literature review**

A systematic literature review identified the studies included in the meta regression analysis (MRA). The literature review combined an existing literature review for the period 2001 to 2008 with an updated review to the end of 2011 (de Bekker-Grob et al, 2012).

The search terms used and databases searched were the same as those used by Ryan and Gerard (2003) and de Bekker-Grob et al (2012): MEDLINE; EMBASE; HEALTHSTAR; Social Science Citation Index; PsychLIT; EconLIT; and Health Management Information Consortium. For consistency, we searched these databases using the same search terms as Ryan and Gerard (2003) and de Bekker-Grob et al (2012), these were devised to encompass all the terms used to refer to discrete choice experiments (DCEs): discrete choice experiment(s), discrete choice model/models/modelling, discrete choice method(s), discrete choice study, stated preference(s), part-worth utilities, functional measurement(s), paired comparison(s), pairwise choice(s), conjoint analysis/analyses, conjoint study/studies, conjoint measurement(s), conjoint approach, conjoint choice experiment. In addition, to ensure a comprehensive set of papers, we checked reference lists of obtained papers and cross checked our review with another being carried out at the same time by Drs Karen Gerard and Tim Bolt. During this process we noticed several papers in the journal 'Health Economics' were missing from the search. We also conducted a search of papers published in 'Health Economics' using the keyword search feature on the website. Studies were included if they were choice-based, not ranking or rating

exercises, published as a full-text in English and were applied to eliciting preferences in a health or healthcare context.

The search identified 371 papers of which 306 papers met the inclusion criteria and 144 were eligible to be included for analysis. Based on eligibility criteria, studies were excluded that either described methodological work and did not report the results of a DCE survey, or were already included in our dataset as a different publication from the same study. For example two different papers may report the results of the same survey but emphasise different results depending on the intended audience for the paper. In this case, we included the study only once to identify unique studies and to avoid inclusion of duplicated studies.

A study may contribute more than one observation to our dataset. For example, if the same questionnaire was administered to two different populations or slightly different questionnaires were administered to subgroups of the same population. In each case, we include these as separate observations but with the same study identifier. For all identified studies, we extracted, when possible, the response rate and the variables of interest as detailed in the main paper. When data were missing from the published paper, we contacted the corresponding author of the paper to request the missing data. We created a standardised request letter and modified this depending on the variable or variables that we were requesting missing data on. All authors of this study requested missing data and studies were allocated across authors based on personal contacts in the first instance, and then at random. We contacted 102 corresponding authors, of which the email was undeliverable for 16 and an internet search failed to identify a new email address. We received responses from 55 corresponding authors, providing a response rate of 63%.

## **Data and robustness tests**

MRA can be sensitive to outliers and lack of variability in the data. When creating the dummy variables based on each variable, we grouped data into categories based on a-priori hypotheses about how these variables affect response rates. We checked the number of observations in each category and merged comparable categories where data were sparse. We test for multicollinearity between variables by estimating pairwise correlations and use a correlation of 0.7 or greater as an indicator that two variables are collinear.

In addition to the linear regression reported in the main body of the paper, we also specified the model as a fractional response model using a logit link function and the binomial distribution (Papke and Wooldridge, 1996; Baum, 2008). We estimated this model using a generalised linear model (GLM) estimator using a cluster robust estimator of variance clustered by a study identifier.

We reweight observations based on the variance of the measure of interest to account for uncertainty around the measure in the original study. In studies with a small sample size, one response more or less would have a larger effect on the reported response rate than it would in a study with a larger sample size. We calculate the variance of the reported response rates using the formula,  $np(1-p)$ , where  $n$  is the sample size and  $p$  is the response rate. We reweight both the linear response model and fractional response model by 1 divided by the square root of the response rate variance.

The fractional response model reproduces most of the linear response model results. The predictions obtained from the fractional response and linear response models are highly correlated ( $p=0.999$ ). Therefore, the linear response model estimated using OLS

is our preferred model. The results of the fractional response model are available from authors on request. Introducing weights based on the response rate variance or sample size increases the goodness of fit for both linear models. The weighted models differs slightly from the unweighted in that the dummy variable capturing the number of alternatives is statistically significant.