

# Chapter 7

## Calculating Marginal and Non-marginal Welfare Measures



**Abstract** This chapter focuses on the calculation of marginal and non-marginal welfare measures. It outlines how the calculation of welfare measures is related to the specified model and the assumptions underlying that model. It further describes how the calculation of these measures is affected by the inclusion of preference heterogeneity, including the incorporation of interaction terms to capture observed preference heterogeneity or random parameters to capture unobserved preference heterogeneity. Finally, it discusses how these measures can be aggregated and compared.

### 7.1 Calculating Marginal Welfare Measures

WTP in the context of DCEs is defined as the amount of income a person is willing to give up for a certain improvement of an attribute or a combination of attributes, so that the overall change in utility is zero. Similarly, WTA is the minimum amount of extra income required to compensate for a certain deterioration of an attribute. WTP and WTA are based on microeconomic theory and correspond to the Hicksian welfare measures (see Sect. 1.1). Freeman et al. (2014, p. 68) describe these concepts in detail and Bateman et al. (2006) present the application of different welfare measures. This section will specifically focus on marginal WTP (mWTP), whereas Sect. 7.2 will discuss welfare implications of larger changes (e.g., in multiple attributes) and related issues of aggregation.

The concept of mWTP is defined as the marginal rate of substitution between the attribute and the price attribute in the indirect utility function and hence relates to the notion of indifference (Dekker 2014).

$$mWTP = -V'(a)/V'(c) \quad (7.1)$$

where  $a$  and  $c$  are the attribute of interest and the cost attribute, respectively and  $V'$  is the first partial derivative of the indirect utility function. For readability, indices for individuals, alternatives and choice situations are omitted. If the attributes enter utility linearly, the mWTP boils down to  $mWTP = -\frac{\beta_a}{\beta_c}$ , where  $\beta_a$  and  $\beta_c$

are the corresponding parameters of the attribute of interest and the cost attribute, respectively.

When using models that incorporate interaction terms to capture observed preference heterogeneity, i.e. allowing marginal utilities of attributes to vary across people, some caution is required. Assume that we interact attribute  $a$  with a continuous, case-specific variable, e.g., age. As utility depends on age, so does mWTP. The value of mWTP is thus a function of age and can be conveniently written as

$$mWTP(age) = -V'(a)/V'(c) = -\frac{\beta_a + \beta_{a-age} \cdot age}{\beta_c},$$

where  $\beta_{a-age}$  is the corresponding parameter of the interaction term. It is possible to substitute any possible value of age to calculate mWTP for that specific age. The ratio  $\beta_a/\beta_c$  will not provide a useful value as it would be the mWTP for a person of age zero. If the mean value of age can be taken as a representative value, it could be advisable to mean-centre this variable before the interaction is formed. Then the mean age would be zero, and the ratio  $\beta_a/\beta_c$  would represent the mWTP for the mean value of age. Nevertheless, in some cases other values of the case-specific variables (e.g. median value), can be more representative.

Further difficulties in calculating mWTP may arise when we specify the cost parameter as random in a RP-MXL model. This is because the ratio of two random variables follows a different, often unknown distribution. For some distributions, the first and second moments of the resulting ratio distribution are not defined, which makes it impossible to report means and standard deviations (Daly et al. 2012). For other distributions, moments may be defined but cannot be calculated analytically. If we are interested in knowing the shape of such a distribution, we can use simulation. The basic idea is to randomly draw from the distributions of the relevant random parameters and calculate mWTP for each draw (Krinsky and Robb 1986, 1991). See Daly et al. (2012, 2020) for the limitations of this approach.

If the random parameters are correlated, these correlations need to be taken into account when generating the random draws to compute their ratios. This is accommodated by drawing from a multivariate distribution. Such draws are feasible for two normally distributed random variables, yet it becomes more difficult if the any of the coefficients have a distributional form that is not a transformation of a normal distribution (Yang 2008). Hensher et al. (2015) recommend only drawing from the multivariate normal distribution.

Simulating the distributions has another advantage: it provides a good indication of whether the assumptions on the random parameters are meaningful. For example, a log-normally distributed cost coefficient may provide a good model fit. However, its large standard deviation may produce very unrealistic results. In a simulation, it can quickly become obvious that many mWTP values are not plausible (see Sect. 8.3 on cross-validation). A recent example of a study using this type of simulation is Knoefel et al. (2018). Simulation is, however, not the key solution because many problems with the resulting distribution can be masked.

The log-normal distribution is frequently used in RP-MXLs to specify the price coefficient. The distribution has the advantage that its values cannot become negative ( $\exp(x) > 0, \forall x$ ). Similar to the normal distribution (and many other parametric distributions), the log-normal distribution is characterised by two parameters, the location parameter  $\mu$  and the scale parameter  $\sigma$ . These parameters determine the shape of the distribution. In the normal distribution,  $\mu$  and  $\sigma$  represent the mean (and the median as the distribution is symmetric) and the standard deviation of the distribution. In the log-normal distribution, the relevant statistics (median, mean and standard deviation) have to be calculated using the formulas presented in Sect. 5.4.

Note that the log-normal distribution is not symmetric, and the mean depends on  $\sigma$ . If  $\sigma$  is large (i.e. the distribution has a fat tail), the mean will quickly become (too) large as well. In such cases, it is useful to report the median as a central tendency measure. In our experience, the median provides, in many cases, a more useful value (Daly et al. 2012). Recent examples for median WTP values from log-normal distributions are Sagebiel et al. (2017) and Rommel and Sagebiel (2017). Note that most statistical software packages output  $\mu$  and  $\sigma$ , and the researcher has to calculate mean, median and standard deviation using the above-mentioned formulas. For policy and welfare analysis however, it is helpful to report the distribution of mWTP, especially, when the distribution does not follow a clearly defined shape.

Finally, RP-MXL models allow the calculation of so-called individual-specific (conditional) utility parameters and mWTP values (Train 2009, Chap. 11; Sarrias 2020). This approach is useful when the researcher aims to predict future choices for specific individuals or when individuals are used for subsequent analysis. However, these conditional values are only meaningful with a sufficiently large number of choices per individual, depending on the complexity of the choice tasks. Also, blocked designs may cause imprecise conditional estimates as individuals faced different sequences of choices. As a consequence, following this approach may only be recommended when it is necessary for a follow-up analysis.

The value of mWTP in a LCM can be calculated for each class separately, using the standard approach  $mWTP_l = -V'_l(a)/V'_l(c)$  where  $l$  denotes the class  $l = (1, \dots, L)$ . The value usually reported is the weighted mean of the within-class mWTP values weighted by the class share.

$$mWTP_w = \sum_{l=1}^L classShare_l \cdot mWTP_l$$

The literature that uses this formula is vast as it is applied in almost all case studies based on LCM, but let us mention for example Scarpa and Thiene (2005).

mWTP is a key concept in welfare economics and rooted in microeconomic theory. Some microeconomic background is required to fully understand the concept of mWTP. While simple model specifications allow a straightforward and easy calculation and interpretation of mWTP, researchers need to be careful once specifications become more complicated. Most importantly, non-centred interaction terms can easily lead to a misinterpretation. Similarly, when using RP-MXL models with

randomly distributed cost coefficients, calculation and interpretation become more complicated.

Nearly all empirical applications in environmental economics rely on mWTP, even in contexts where welfare effects are not the main subject. In the latter cases, mWTP serves as a way of obtaining a meaningful interpretation of results and some kind of importance ranking. As mWTP in RP-MXL models is rather difficult to obtain, many studies investigated this topic. A good summary is provided in Hensher et al. (2015). A detailed, and somewhat advanced discussion of mWTP for random cost parameters is Daly et al. (2012). Conditional mWTP values are described in detail in Train (2009, Chap. 11). In general, practitioners should look into distributions of mWTP rather than specific statistics (mean, median, standard deviation), and only use those models where a meaningful interpretation of mWTP is feasible. Researchers should double check if the mWTP distributions are realistic. If a large percentage of the distribution falls outside the range of acceptable values (e.g. mWTP for renewable energy per kWh is more than 1€, or a yearly payment for water quality improvements of 10,000 €), something may be wrong, even if model fit statistics indicate differently.

## 7.2 Aggregating Welfare Effects

Often researchers are not interested in the welfare implications of marginal changes, but wish to derive the monetary value of a policy intervention (or product change). Policy interventions often involve changes in multiple attributes and of a reasonable but non-marginal size. Economic theory provides the tools to derive such values. First, we need to establish the do-nothing scenario and the corresponding utility level for each individual, denoted by  $V_0$ . Any policy intervention, assuming a quality improvement, will increase the utility to  $V_1$ . Effectively, we are interested in the value of the utility difference  $V_1 - V_0$  for each individual and the aggregation of these individual effects. This is exactly what the Hicksian welfare measures do (see Sect. 1.1).

A key issue with discrete choice modelling is, however, that a priori we do not know which goods individuals will select in the do-nothing scenario and whether they would switch as a result of the policy change. We typically work with what is known as the LogSum (Ben-Akiva and Lerman 1985), which denotes the expected maximum utility of a choice set here denoted for the do-nothing scenario—where  $j$  denotes an alternative in the choice set:

$$LS_0 = \ln \left( \sum_j \exp(V_{j0}) \right).$$

The change in the expected maximum utility as a result of the policy intervention is then denoted by:

$$LS_1 - LS_0 = \ln \left( \sum_j \exp(V_{j1}) \right) - \ln \left( \sum_j \exp(V_{j0}) \right).$$

Utility is, however, not informative for cost–benefit purposes and we require a translation into monetary terms using the marginal utility of income  $\lambda$ . Batley and Ibáñez Rivas (2013) highlight that when the indirect utility function is linear in income and price, we can use the negative of the cost coefficient  $\beta_c$  for this purpose, i.e.  $\lambda = -\beta_c$ , such that the monetary change in e.g. compensating surplus is denoted by:

$$\Delta CS = \frac{LS_1 - LS_0}{\lambda}.$$

When the discrete choice model is linear in both the non-cost *and* cost attributes, the LogSum, i.e. the change in compensating surplus, is identical to the sum of the constant marginal WTP estimates for the individual attributes. When non-linear non-cost attributes are introduced, aggregation over non-marginal quality improvements is far less trivial than simply deriving the change in the LogSum for a given scenario. Hence, the use of the LogSum is recommended in such instances.

The introduction of non-linear costs in the indirect utility functions causes significant theoretical and computational challenges. The issue here is that due to the inclusion of non-linear cost effects the marginal utility of income is no longer constant and thus invalidates the use of the LogSum in these cases. Batley and Dekker (2019) show that in a discrete choice setting non-linear cost effects are non-compatible with economic theory, despite the fact that an econometric model may suggest that non-linear costs are likely. Karlstrom and Morey (2003) and McFadden (1996) have developed methods to derive the resulting change in “compensating variation” using methods of integration and simulation, respectively. These methods are hardly implemented in the literature due to their challenging computational burden.

The LogSum thus allows the change in compensating surplus to be derived for a given individual. However, we are typically interested what the policy intervention implies for the population and not the sample. When mWTP estimates are used, such as the case for the Value of Statistical Life (Robinson and Hammitt 2016) one can simply use these as a multiplier as stated in national cost–benefit analysis guidelines such as the Green Book in the UK (HM Treasury 2018). Alternatively, one can derive the welfare implications for different representative socio-economic groups using the estimated indirect utility functions and the LogSum and aggregate over the population. Essentially, this will indicate whether the net WTP is positive or negative for society, and thus be a reflection of the Kaldor and Hicks compensation criteria. However, Nyborg (2014) highlights there are significant controversial value judgements made when simply aggregating WTP and WTA across people. She argues that effectively more weight is given to richer people in the social welfare analysis as a result.

To sum up, firstly, the output of discrete choice models is commonly a set of mWTP measures. These are not always informative when quality effects are non-linear in the indirect utility function. Secondly, the LogSum facilitates aggregation of welfare effects, particularly versus the do-nothing scenario. Thirdly, more complicated calculations are required when non-linearities are associated with income and price variables.

### 7.3 WTP Comparison

In some applications, it is interesting to compare welfare measures from different samples. For example, a researcher has collected two samples from different cities and wants to find out if WTP is larger in one city than in another city. Or the researcher has conducted a split sample to answer a specific methodical question and wants to find out whether there is a difference in the two samples. Direct testing with *t*-tests is not appropriate, as welfare measures such as mWTP are ratios of coefficients (non-linear combinations) and they are, therefore, usually non-normally distributed. One way to test differences by comparing simulated distributions. The idea is to simulate mWTP values and count in how many cases the mWTP value from one sample is larger than that from the other sample. This procedure has been proposed by Poe et al. (1994, 2005). A step-by-step guide can be found in Haab and McConnell (2002, p. 112).

The Poe test can be conducted for basically any model including RP-MXL models, but be aware that mean and median mWTP are sometimes calculated from the location and scale parameters (e.g. for log-normally distributed price coefficients, see Sect. 7.1), which requires a different formula for mWTP. One can use Poe test to compare other welfare measures such as compensating surplus of a specific policy scenario. The process is similar to that with mWTP, with the only difference that, for each draw, the compensating surplus formula is used instead of the mWTP formula.

If the Poe test is not feasible, or a formal test is not required, one can conduct a graphical analysis of confidence intervals. Plotting mean mWTP and their respective confidence intervals of two samples offer a good initial insight into the magnitude of the differences. When confidence intervals overlap, mWTP are not likely to be statistically different. Note that both in the Poe test and in the overlapping confidence interval approach, the null hypothesis of equality of mWTP is less likely to be rejected the larger the variations and confidence intervals are.

In summary, comparing independent samples with respect to welfare measures can be done with the Poe et al. (2005) test or with the overlapping confidence interval method. Several empirical applications rely on the Poe test to establish if there are differences in WTP between samples. See, for example, Liebe et al. (2015) or Glenk et al. (2019) for methodological applications, and Brouwer et al. (2010, 2016) and Knoefel et al. (2018) for empirical applications.

## References

- Bateman IJ, Day BH, Georgiou S, Lake I (2006) The aggregation of environmental benefit values: welfare measures, distance decay and total WTP. *Ecol Econ* 60:450–460. <https://doi.org/10.1016/j.ecolecon.2006.04.003>
- Batley R, Dekker T (2019) The intuition behind income effects of price changes in discrete choice models, and a simple method for measuring the compensating variation. *Environ Resource Econ*. <https://doi.org/10.1007/s10640-019-00321-2>
- Batley R, Ibáñez Rivas JN (2013) Applied welfare economics with discrete choice models: implications of theory for empirical specification. In: Hess S, Daly A (eds) *Choice modelling*. Edward Elgar Publishing, pp 144–171
- Ben-Akiva M, Lerman S (1985) *Discrete choice analysis. theory and applications to travel demand*. The MIT Press, Cambridge, MA
- Brouwer R, Bliem M, Getzner M et al (2016) Valuation and transferability of the non-market benefits of river restoration in the Danube river basin using a choice experiment. *Ecol Eng* 87:20–29. <https://doi.org/10.1016/j.ecoleng.2015.11.018>
- Brouwer R, Martin-Ortega J, Berbel J (2010) Spatial preference heterogeneity: a choice experiment. *Land Econ* 86:552–568. <https://doi.org/10.3368/le.86.3.552>
- Daly A, Hess S, Train K (2012) Assuring finite moments for willingness to pay in random coefficient models. *Transportation* 39:19–31. <https://doi.org/10.1007/s11116-011-9331-3>
- Daly A, Hess S, Ortúzar J (2020) Estimating willingness-to-pay from discrete choice models: setting the record straight. Working paper. Choice Modelling Centre (CMC). University of Leeds. <https://www.stephanehess.me.uk/publications.html>
- Dekker T (2014) Indifference based value of time measures for Random Regret Minimisation models. *J Choice Model* 12:10–20. <https://doi.org/10.1016/j.jocm.2014.09.001>
- Freeman AMI, Herriges JA, Kling CL (2014) *The measurement of environmental and resource values : theory and methods*. Routledge
- Glenk K, Meyerhoff J, Akaichi F, Martin-Ortega J (2019) Revisiting cost vector effects in discrete choice experiments. *Resource Energy Econ* 57:135–155. <https://doi.org/10.1016/j.reseneeco.2019.05.001>
- Haab TC, McConnell KE (2002) *Valuing environmental and natural resources. The econometrics of non-market valuation*. Edward Elgar Publishing Limited, Cheltenham, UK
- Hensher DA, Rose JM, Greene WH (2015) *Applied choice analysis, 2nd edn*. Cambridge University Press, Cambridge
- HM Treasury (2018) *The green book: central government guidance on appraisal and evaluation*. OGL Press, London, UK
- Karlstrom A, Morey ER (2003) Calculating the exact compensating variation in logit and nested-logit models with income effects: theory, intuition, implementation, and application. Social Science Research Network, Rochester, NY
- Knoefel J, Sagebiel J, Yildiz Ö et al (2018) A consumer perspective on corporate governance in the energy transition: evidence from a Discrete Choice Experiment in Germany. *Energy Econ* 75:440–448. <https://doi.org/10.1016/j.eneco.2018.08.025>
- Krinsky I, Robb A (1986) On approximating the statistical properties of elasticities. *Rev Econ Stat* 68:715–719. <https://doi.org/10.2307/1924536>
- Krinsky I, Robb AL (1991) Three methods for calculating the statistical properties of elasticities: a comparison. *Empirical Econ* 16:199–209. <https://doi.org/10.1007/BF01193491>
- Liebe U, Glenk K, Oehlmann M, Meyerhoff J (2015) Does the use of mobile devices (tablets and smartphones) affect survey quality and choice behaviour in web surveys? *J Choice Model* 14:17–31. <https://doi.org/10.1016/j.jocm.2015.02.002>
- McFadden D (1996) *Computing willingness-to-pay in random utility models*. University of California at Berkeley, Econometrics Laboratory Software Archive

- Nyborg K (2014) Project evaluation with democratic decision-making: what does cost–benefit analysis really measure? *Ecol Econ* 106:124–131. <https://doi.org/10.1016/j.ecolecon.2014.07.009>
- Poe GL, Giraud KL, Loomis JB (2005) Computational methods for measuring the difference of empirical distributions. *Am J Agr Econ* 87:353–365. <https://doi.org/10.1111/j.1467-8276.2005.00727.x>
- Poe GL, Severance-Lossin EK, Welsh MP (1994) Measuring the difference (X–Y) of simulated distributions: a convolutions approach. *Am J Agr Econ* 76:904–915. <https://doi.org/10.2307/1243750>
- Robinson LA, Hammitt JK (2016) Valuing reductions in fatal illness risks: implications of recent research. *Health Econ* 25:1039–1052. <https://doi.org/10.1002/hec.3214>
- Rommel K, Sagebiel J (2017) Preferences for micro-cogeneration in Germany: policy implications for grid expansion from a discrete choice experiment. *Appl Energy* 206:612–622. <https://doi.org/10.1016/j.apenergy.2017.08.216>
- Sagebiel J, Glenk K, Meyerhoff J (2017) Spatially explicit demand for afforestation. *Forest Policy Econ* 78:190–199. <https://doi.org/10.1016/j.forpol.2017.01.021>
- Sarrias M (2020) Individual-specific posterior distributions from Mixed Logit models: properties, limitations and diagnostic checks. *Journal of Choice Modelling* 100224. <https://doi.org/10.1016/j.jocm.2020.100224>
- Scarpa R, Thiene M (2005) Destination choice models for rock climbing in the Northeastern Alps: a latent-class approach based on intensity of preferences. *Land Econ* 81:426–444
- Train K (2009) *Discrete choice methods with simulation*, 2nd edn. Cambridge University Press, New York
- Yang M (2008) Normal log-normal mixture, leptokurtosis and skewness. *Appl Econ Lett* 15:737–742. <https://doi.org/10.1080/13504850600749073>

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), 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 license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license 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.

