Deriving welfare measures in discrete choice experiments: a comment to Lancsar and Savage (2)

J. M. C. Santos Silva*
ISEG/Universidade Técnica de Lisboa, Portugal

Introduction

Lancsar and Savage (Deriving welfare measure from discrete choice experiments: inconsistency between current methods and random utility and welfare theory. Health Econ 2004; this issue) point out that the methods currently used to derive measures of welfare in the health economics literature are often inappropriate, and recommend the use of the expected compensating variation formula derived by Small and Rosen [1], especially when data are collected through discrete choice experiments. This is an important issue and the contribution of Lancsar and Savage (2004) may be a significant step to correct the misuse of welfare measures. Unfortunately, the work of Lancsar and Savage (2004) is marred by some errors and imprecisions which may be misleading. Indeed, the authors go too far in their criticism of the welfare measures currently used in the health economics literature and fail to point out some limitations of the Small and Rosen [1] method, which they advocate. Overall, I believe that Lancsar and Savage (2004) provide an unbalanced and poorly substantiated view of the problem.

This note reconsiders the computation of measures of welfare in discrete choice problems and explores the relation between the marginal willingness to pay and the concepts of compensating variation (for an individual) and expected compensating variation (for a population). Moreover, some econometric issues related to the estimation of these welfare measures are also discussed. This aspect was ignored by Lancsar and Savage (2004), but has a critical role in the choice of the appropriate welfare measure to use in a particular application. In contradistinction to what is claimed by Lancsar and Savage (2004), it is argued that the marginal willingness to pay, the compensating variation and the expected compensating variation, can be important instruments to measure welfare changes.

Like in Lancsar and Savage (2004), only the case of the multinomial logit model is considered, but the results can be easily extended to other settings.

The framework

Consider an individual choosing one of $J \geq 2$ alternatives according to the utility they provide. Let $U_{ij}$ denote the utility provided by alternative $j$ to individual $i$ and assume that

$$U_{ij} = X_j \beta + u_{ij},$$

where $X_j$ is a set of observable characteristics of the alternative and $u_{ij}$ is a component of the utility not observed by the researcher. For sake of simplicity, it is assumed that the utility of the alternatives does not depend on observable characteristics of the individuals. However, the results presented here can be easily extended to more general cases.

Assuming that $u_{ij}$ is an extreme value type 1 random variable, the probability that the alternative $j$ is chosen by individual $i$ is given by

$$\pi_j = \frac{\exp(X_j \beta)}{\sum_{k=1}^{J} \exp(X_k \beta)},$$

which can also be interpreted as the probability that an individual drawn at random from the population will choose alternative $j$. That is, $\pi_j$...
also represents the share of the population that chooses $j$.

To estimate the unknown parameters it is necessary to obtain data on the observable characteristics of the alternatives and on the proportion of individuals choosing each alternative. When all the alternatives under study are available, obtaining the appropriate data is generally trivial. When such revealed preference data are inappropriate or simply not available, stated preference data can be gathered by means of a discrete choice experiment in which the subjects are asked to choose one of the hypothetical alternatives presented to them. These answers can be used to estimate the proportion of individuals that would choose each alternative, in the same way a survey of the population is used when all alternatives are available.

Both stated and revealed preference data have advantages and drawbacks but, whatever method is used to collect the data, the interpretation of the underlying probabilistic model is the same (see [2], for a concise but clarifying exposition of the advantages and disadvantages of the two types of data). Therefore, although Lancsar and Savage (2004) put particular emphasis on cases in which data are obtained using a discrete choice experiment, the theoretical welfare measures studied here are applicable in any discrete choice problem, and not only when using stated preference data. As it is discussed below, the particular type of the data that is available has implications both for the choice of estimation methods and for the interpretation of the results. However, these are implementation issues which have little to do with the validity of the theoretical welfare measures.

**Welfare measures**

In order to clarify the relation between different measures of welfare, it is useful to write

$$V_j = X_j \beta = X_{1j} \beta_1 + p_j \beta_p + \tilde{X} \tilde{\beta}$$

where $X_1$ is the characteristic whose effect on the welfare we want to measure, $p_j$ is the price of alternative $j$, $\tilde{X}$ denotes the remaining characteristics of alternative $j$ and $\beta_1$, $\beta_p$ and $\tilde{\beta}$ are parameters.

Consider an individual that chooses alternative $j$ and suppose that the value of $X_{1j}$ changes by an amount $\Delta X_{1j}$. If, after this change, the individual continues to choose alternative $j$, the change in its utility is given by $\Delta X_{1j} \beta_1$. In this case the utility will remain unchanged if $\Delta X_{1j} \beta_1 + \Delta p_j \beta_p = 0$. Therefore, the offsetting price variation in this case is given by

$$\Delta p_j = -\Delta X_{1j} \frac{\beta_1}{\beta_p}$$

Equivalently, defined in terms of income, the compensating variation (CV) is given by

$$CV = \Delta X_{1j} \frac{\beta_1}{\beta_p}$$

where the CV is defined as the increase in income that is needed to bring the individual back to its initial utility level, after a price or quality change.

It is interesting to notice that Lancsar and Savage (2003) define the CV in a nonstandard way. Indeed, the authors state that ‘CV is a measure of how much money needs to be given to or taken from a consumer after a price or quality change to leave them at their initial level of utility’ (Lancsar and Savage, 2004). Because this definition does not specify if the money needs to be given or taken, it only identifies the magnitude of the welfare change, but not its direction. That is, in Lancsar and Savage (2004) only the absolute value of the CV can be interpreted, but not its sign. Of course, this ambiguity is problematic when the direction of the change is uncertain.

The compensating variation defined in (1) is only useful to evaluate welfare changes for individuals that do not choose a different alternative after $X_{1j}$ changes. For example, it can be used to calculate the change in price needed to compensate the consumers for the deterioration in the quality of a service provided by some company. Because the drop in quality is offset by a reduction of the price, all utilities will be unchanged and therefore there is no reason for individuals to choose a different alternative.

Another situation in which choice probabilities can be treated as constant is when infinitesimal changes in $X_{1j}$ are considered. In this case $\partial X_{1j} \beta_1 + \partial p_j \beta_p = 0$, which leads to the marginal willingness to pay (MWTP) for $X_1$ (or marginal rate of substitution between income and $X_1$):

$$MWTP = \frac{\partial p_j}{\partial X_{1j}} = \frac{\partial V_j / \partial X_{1j}}{\partial V_j / \partial p_j} = \frac{\beta_1}{\beta_p}$$

That is, the MWTP measures the change in income that individuals are willing to accept in exchange for marginal changes in $X_1$, whatever the
alternative they choose. Notice that, assuming that $\beta_j$ and $\beta_p$ do not vary across alternatives or individuals, the MWTP does not depend on the alternative chosen by the individual and consequently it is the same for the entire population.

By its own nature, the MWTP is not appropriate to evaluate the total welfare change resulting from discrete changes in the characteristics of the alternatives. In an heterogeneous population, the expected compensating variation (ECV) required by a change in $X_{ij}$ depends on the proportion of individuals that chooses each alternative before and after the change in this characteristic. Small and Rosen [1] show that, in the case of the logit, the expected compensating variation in a population is given by

$$ECV = -\frac{1}{\lambda} \left[ \ln \sum_{k=1}^{J} \exp[V_k + \delta_{kj} A_{X_{ij}} \beta_1] 
- \ln \sum_{k=1}^{J} \exp[V_k] \right]$$

(2)

where $\lambda$ is the marginal utility of income and $\delta_{kj}$ denotes the Kronecker delta which is equal to 1 when $k = j$, being 0 otherwise. Since $\lambda$ is generally not available, it is usually replaced by the negative of the coefficient on the price [2]. That is, $\lambda = -\beta_p$.

This expression for the ECV assumes that the marginal utility of income is constant. Morey et al. [3] provide an example of how this assumption can be relaxed in the context of a logit model for the choice of malaria treatments. A more general approach to this problem has been recently developed [4–6], but this is not pursued here.

The form of (2) does not make its meaning immediately clear. In their discussion of the ECV expression, Lancsar and Savage (2004) emphasize that the ECV depends on the difference between two inclusive values or log-sums, which have interesting interpretations in other settings. However, in this context, these inclusive values are just the result of the form of the logit probabilities and have no economic interpretation [2]. For example, if the model is a probit rather than a logit the ECV does not depend on inclusive values [1]. Perhaps the clearer interpretation of the ECV expression results from seeing it as the difference between the value of the consumer surplus before and after the change in the characteristics of the alternatives [2, 7].

To gain some insight into the meaning of (2), it is interesting to consider the case in which $X_{ij}$ varies by a small amount. Expanding the ECV in a Taylor series around $A_{X_{ij}} = 0$, leads to

$$ECV \approx -\frac{1}{\lambda} \left[ \frac{A_{X_{ij}} \beta_1 \exp[V_j]}{\sum_{k=1}^{J} \exp[V_k]} \right] = \frac{A_{X_{ij}} \beta_1}{\beta_p} \pi_j$$

That is, for small values of $A_{X_{ij}}$, the expected compensating variation is approximately given by the compensating variation corresponding to the individuals that choose $j$ (see Equation (1)), times the share of the population that chooses this alternative. Of course, for large changes in the utility of at least one alternative, this approximation may not be adequate as it ignores the consequent changes in the shares of each alternative. Albeit only valid for small variations, this approximation helps to understand the meaning of (2) and its relation with the CV and the MWTP.

The discussion above suggests that whenever the variation in the characteristics of the utilities leaves the choice probabilities unchanged, the MWTP has a leading role in the evaluation of welfare change. Indeed, suppose that $X_1$ changes for all alternatives by the same amount. In this case

$$ECV = -\frac{1}{\lambda} \left[ \ln \sum_{k=1}^{J} \exp[V_k + \beta_1 A_{X_{1i}}] 
- \ln \sum_{k=1}^{J} \exp[V_k] \right]$$

and a simple algebraic manipulation leads to

$$ECV = -\frac{1}{\lambda} \left[ \beta_1 A_{X_{1i}} \right] = A_{X_{1i}} \frac{\beta_1}{\beta_p}$$

Therefore, in this special case, the ECV is identical to the CV and can be directly obtained as the product of the MWTP by $A_{X_{1i}}$. Lancsar and Savage (2004) claim that the ECV only coincides with (1) when $\pi_j = 1$ and $J = 2$ (the importance of the case in which $\pi_j = 1$ is emphasized by Ryan (Deriving welfare estimates from discrete choice experiments: a response to Lanscar and Savage. Health Econ 2004)). However, the equality also holds when the change in the expected utility is equal for all alternatives, whatever the value of $J$.

More generally, the equality holds whenever variations of the attributes leave the choice probabilities unchanged.
Estimation issues

The estimates of the MWTP and of the ECV have different degrees of robustness to the possible misspecification of the discrete choice model used in the estimation. The MWTP is estimated as the ratio of two coefficients in the utility function, a quantity that is often consistently estimated even in presence of some forms of misspecification [8–10]. On the other hand, the ECV requires the consistent estimation of the choice probabilities, which generally depends on the correct specification of the parametric model. Therefore, the ECV is less robust to possible model misspecification than the MWTP. Given the possible sensitivity of the results to specification problems, it is important to carefully test the models used to evaluate welfare changes, particularly if the ECV is to be used. Chesher and Santos Silva [11] discuss several specification tests for the correct specification of logit models and propose a particular specification that may accommodate some forms of misspecification.

It is also worth noting that the sensitivity of the ECV to the value of the choice probabilities makes its use problematic when estimation is performed exclusively using data from discrete choice experiments. Indeed, it is widely believed that models estimated from stated preferences data do not accurately describe the choice probabilities, unless some sort of calibration is used to correct the alternative specific intercepts [2,12]. When only stated preferences data are available, for instance because some of the alternatives considered are not currently available, it may not be possible to perform this calibration. Therefore, although Lancsar and Savage (2004) argue that this is the case in which the use of the ECV is more strongly recommended, when only stated preferences data are available, estimates of the ECV should be viewed with great caution.

As an illustration, some simple exercises can be performed using the numerical results presented in Lancsar and Savage (2004). In these exercises the results in Lancsar and Savage (2004) are taken at face value. Their validity is studied in the next section. The estimate for the intercept in the model presented by Lancsar and Savage (2004) implies that if both the current and hypothetical alternatives had exactly the same observed characteristics, the hypothetical alternative would be chosen by only 2

The empirical illustration

Although the numerical results presented in Lancsar and Savage (2004) are just an illustration of the use of different welfare measures, and clearly not the main focus of their paper, it is important to clarify a few issues about this particular application which can be misleading for the less experienced practitioner.

As explained in [13], the empirical illustration presented in Lancsar and Savage (2004) uses data on a sample of 27 asthmatics who were faced with 12 scenarios in which they chose between their current medication and a hypothetical alternative. Actually, the questionnaire had a third option of no medication but, because it was very seldom selected, the authors decided to drop this alternative and treat the data as coming from a binary choice question. Since each individual provides a set of 12 observations, Lancsar and Savage (2004) account for the possible presence of individual specific effects estimating a conditional fixed effects logit model.

The use of the conditional fixed effects logit model is potentially appropriate to describe this type of data where there are multiple observations for each individual. However, its use raises important issues when computing welfare measures. Indeed, although Lancsar and Savage (2004)
report an estimate for the ECV, it is actually impossible to compute the ECV from the results of the conditional fixed effects logit! The MWTP, however, can be computed as usual.

As described in many standard textbooks [14, 15], the fixed effects conditional logit model only identifies a subset of the parameters of the utility function. Indeed, this estimator does not permit the estimation of the individual fixed effects, or the identification of the parameters associated with the regressors that do not vary across alternatives. Therefore, the results provided by this estimator are not enough to estimate the probability that an individual will choose a given alternative because this probability depends on the fixed effects. These probabilities determine the share of the population that chooses each alternative and without them it is not possible to estimate the ECV.

The probabilities that are used in the construction of the likelihood function of a fixed effects binary logit model are the probabilities that a particular set of choices is made by an individual, conditional on the total number of times one of the alternatives is chosen by him. These probabilities, which are implicitly used by Lancsar and Savage (2004) to compute the ECV, bear little relation to the probabilities that are needed to compute the ECV. On the other hand, the MWTP is just the ratio of the parameters of two covariates that vary across alternatives, and thus it can be computed without any problems.

Besides this major issue, there are some other problems with this empirical illustration. Upon examining the data set used, it is possible to conclude that for one of the individuals there is no data for its current medication, and still the information on the hypothetical alternative is used in the estimation. Since the choice probabilities depend only on the difference between the attributes of the competing alternatives, it is not clear how this information can be used in the estimation of a binary model. The authors informed me (personal communication by e-mail on 5 September 2003) that reestimating the model dropping the data for this individual leads to a log-likelihood of $-194.31$. Given that the log-likelihood value reported in Lancsar and Savage (2004) is $-197.97$, it is clear that the results presented in the paper somehow use the information on this individual. Unfortunately, no reasonable explanation for how this is possible could be found.

Further examination of the data shows that five individuals whose data is used in the empirical illustration always choose the current medication. However, it is well known that individuals that always choose the same alternative do not contribute to the log-likelihood of a fixed effects logit model [14, 15]. Since these $5 \times 12 = 60$ observations could not possibly be used in the estimation of the model described by Lancsar and Savage (2004), it is difficult to understand how the results reported in the paper were actually obtained.

In view of these problems, the value for the ECV reported by Lancsar and Savage (2004) is meaningless. It is ironic that the authors use in their empirical illustration an estimator that does not actually permit the computation of the welfare measure they advocate so strongly. However, this unfortunate situation may have its positive side if it serves as a warning to the less experienced practitioners about the dangers of using powerful econometric software without a reasonable understanding, not only of the software itself, but also of the econometric methods used.

**Conclusion**

The results in the ‘welfare measures’ section highlight that the concepts of compensating variation, expected compensating variation and marginal willingness to pay, are all closely related. Moreover, the results presented here show that, when appropriately used, all these concepts are interesting and important instruments to measure welfare changes. Therefore, I cannot agree with Lancsar and Savage (2004) when they claim that the marginal willingness to pay is not consistent with welfare theory and with the random utility framework in which the discrete choice models considered here are based.

Naturally, the use of the compensating variation and marginal willingness to pay is not always appropriate and the work of Lancsar and Savage (2004) has the merit of emphasizing the usefulness of measuring welfare changes using the expected compensating variation, which is especially interesting when the researcher is relatively confident of the correct specification of the model. However, the use of this measure of welfare change may be problematic when only stated preference data are available, precisely the case for which Lancsar and Savage (2004) recommend its use.
Acknowledgements

I am most grateful to Pedro Pita Barros, the Editor Jim Burgess and Mandy Ryan for many helpful discussions and suggestions. I am also very grateful to Emily Lancsar for providing the data set and the code used in the estimation of the empirical illustration presented in Lancsar and Savage (2004). I have also benefited from helpful and extensive discussions with Emily Lancsar on the details of this particular application. The usual disclaimer applies. Partial financial support from Fundação para a Ciência e Tecnologia, program POCTI, partially funded by FEDER, is gratefully acknowledged.

References