CREI Working Paper no. 4/2012

THE IMPACT OF THE BEFORE-AFTER ERROR TERM CORRELATION ON WELFARE MEASUREMENT IN LOGIT

by

Paolo Delle Site
“La Sapienza” University

Marco Valerio Salucci
“La Sapienza” University

available online at http://host.uniroma3.it/centri/crei/pubblicazioni.html

ISSN 1971-6907

Outputs from CREI research in progress, as well contributions from external scholars and draft reports based on CREI seminars and workshops, are published under this series. Unless otherwise indicated, the views expressed are attributable only to the author(s), not to CREI nor to any institutions of affiliation.
The impact of the before-after error term correlation on welfare measurement in logit

Paolo Delle Site and Marco Valerio Salucci
DICEA – Department of Civil, Architectural and Environmental Engineering
University of Rome “La Sapienza”
Via Eudossiana, 18 – 00184
Rome, Italy

Abstract

We consider random utility models with independent and identical type I extreme value distribution of the error terms. To compute the expectation of the compensating variation it is necessary to consider the correlation of the error terms between the state before the price and quality change and the state after. We investigate the impact of the before-after correlation of the error terms on the expectation of the compensating variation. We consider each error term to be correlated between the before state and the after state independently and identically across alternatives. We prove the theoretical property that in the case without income effect the logsum formula holds for any assumption on the before-after correlation. We use numerical evidence to show that in the case with income effect the variability of the expectation of the compensating variation with the assumption on the before-after correlation increases with the size of the income effect.

Keywords: logit, compensating variation, before-after correlation, income effect

JEL codes: C25, D60
1. Introduction

Logit is the simplest and most popular of the broad class of discrete choice models which can be derived by assuming that each individual maximises a random utility function. The mainstream approach to welfare measurement, today, is based on stochastic compensating variation. The compensating variation measures the welfare associated with a policy which changes the price and quality of the choice alternatives. Given that logit belongs to the class of random utility models, it is most natural to adopt as welfare measure the expectation of the compensating variation derived from random utility. Computation of this welfare measure depends on how income enters the specification of the (conditional indirect) utility functions.

McFadden (1981) was the first to lay the micro-economic foundation of random utility models. He shows that if utilities depend on residual income, i.e. on the expenditure on the numeraire, Roy’s identity is satisfied. However, this condition may be violated by econometric specifications used in applied work. This circumstance has led a few authors (Viton, 1985; Jara-Díaz and Videla, 1989) to investigate the consistency between the micro-economic foundation and the econometric specification of random utility models, and to argue that, to restore consistency, income variables appearing in the specification may need to be re-interpreted, as an example as proxy for tastes rather than as true expenditure entering Roy’s identity.

The majority of the applications of random utility models have considered that there is no income effect, i.e. income does not influence choice. A non-income sensitive specification can be obtained, as shown by McFadden in his 1981 essay, by assuming that the marginal utility of income is constant across alternatives. This assumption greatly simplifies the computation of the expectation of the compensating variation. This is because the maximum utility is a linear function of income, and, hence, a closed form expression for both the compensating variation and its expected value can be computed. McFadden (1999) proved that the expectation of the compensating variation equals the difference, normalized by the marginal utility of income, of the expectations before and after the change of the maximum utility and takes a logsum form for logit.

There are, however, applications where the hypothesis of absence of income effect is, at least in principle, questionable. One example is travel demand. In developing countries expenditure in transport can be a significant fraction of many socio-economic groups’ budget, as shown by Jara-Díaz and Ortúzar (1989). In addition, income effect can be postulated in applications with high travel prices, which may occur for airline tickets and for charges resulting from stiff road pricing policies. Income effect implies that a consumer who is facing a set of alternatives and alternative attributes may make different choices at different levels of her income.
The first approach proposed to compute the expectation of the compensating variation in the case with income effect is based on the simulation of the distribution of the error terms (McFadden, 1999; Herriges and Kling, 1999; Cherchi et al., 2004). Later, a few authors developed analytic approaches. Formulas of the expectation of the compensating variation for additive models have been provided, on the basis of the distribution of the compensating variation, by Dagsvik and Karlström (2005), who used Hicksian, i.e. compensated, probabilities, and by De Palma and Kilani (2011), who used Marshallian, i.e. observed, probabilities. The formulas available for the case with income effect are not in closed form and require numerical integration even for the logit model.

To compute the expectation of the compensating variation it is necessary to make an assumption on the correlation of the error terms between the state before the price and quality change and the state after the change. For reasons of mathematical convenience, all the methods developed to provide the expectation of the compensating variation analytically have assumed that the error terms remain the same before and after, i.e. an assumption of perfect before-after correlation is made. The assumption that the error terms remain the same before and after is usually retained also when the measure is computed by simulation.

Relaxing the perfect correlation assumption can be justified on several grounds. There are two interpretations of random utility models. One accounts for inter-individual variability of tastes as it is assumed that the error terms are individual specific, i.e. utility is deterministic for the decision maker, stochastic only for the modeller. The other accounts for intra-individual variability of tastes as it is assumed that the individual draws from a distribution each time a choice is made, i.e. utility is stochastic also for the decision maker. In the first interpretation, changes in the error terms before and after would arise from unobserved attributes of the alternatives or of the individual that are changed by the policy under consideration. In the second interpretation, we need to assume that the error terms do change before and after because the choice is stochastic.

Zhao et al. (2012) appear to be the first to publish on the impact of the before-after correlation. They use simulation to provide the expectation of the compensating variation in logit in the case without income effect. Before and after error terms are generated based on rank order correlation, independently and with identical correlation across alternatives. They find that the logsum formula is robust to the assumption on the before-after correlation because the differences of the logsum with the measures obtained from simulation with varying before-after correlation result to be negligible. The impact of the before-after correlation in the case with income effect has not been investigated to date.

The paper provides a theoretical and numerical insight on the impact of the before-after
correlation on the expectation of the compensating variation for logit. Formulas, in both cases without and with income effect, are available but hold under a perfect before-after correlation assumption. Formulas are computationally less demanding than the simulation approach. Therefore, it is of practical relevance to investigate their robustness with respect to the before-after correlation assumption.

First, the paper shows that in the case without income effect the independence of the logsum formula from the before-after correlation is a theoretical property of logit. Second, it provides with the support of a numerical example evidence on the impact of the before-after correlation in the case with income effect. The deviations of the measures obtained by simulation from the measure obtained analytically in the perfect correlation case are evaluated. Following the proof that the absence of income effect results in no impact of the before-after correlation, it is of interest to explore how this impact changes with the size of the income effect. The paper addresses this question. The organisation is as follows. Section 2 outlines the theoretical framework for welfare measurement in logit when the before-after correlation of the error terms is considered. Section 3 reports on the theoretical result in the case without income effect. Section 4 on the numerical results in the case with income effect. Section 5 concludes the paper.

2. Welfare measurement with the logit model

An individual endowed with income $y$ faces a set of $J$ mutually exclusive alternatives. Given the utility-maximising behaviour subject to a constraint on income spent, when alternative $i$ is chosen the individual will be characterised by a conditional indirect utility function $u_i$. The utility $u_i$ is expressed by the additively separable structure: $u_i = v_i + \varepsilon_i$, $i=1,..,J$, where $v_i$ is the deterministic component, referred to as systematic utility, and $\varepsilon_i$ is the error component. This structure for $u_i$ defines the class of additive random utility models.

The probability of choosing alternative $i$ is $P_i = \Pr(u_i > u_j, j \neq i)$. Under the assumption that the error terms are independently and identically distributed (i.i.d.) according to a type I extreme value (Gumbel) distribution with location parameter equal to 0 and scale parameter $\theta$, the probability $P_i$ that an individual chooses alternative $i$ is the logit function:

$$P_i = \frac{e^{v_i/\theta}}{\sum_{j=1}^{J} e^{v_j/\theta}}.$$

We assume that the systematic utility of each alternative $v_i$ depends on residual income, i.e. the part of income $(y - p_i)$ not spent on the alternative, where $p_i$ is the price of the alternative. Prices and income are deflated by the price of an outside commodity. We assume
the additive structure:

\[ v_i = v_i(y - p_i, \bar{v}_i) = f_i(y - p_i) + \bar{v}_i \quad i = 1, \ldots, J \]  

(1)

where \( f_i \) is an increasing function of the argument and \( \bar{v}_i \) is a function of qualitative attributes of the alternative distinct from price. In addition, we assume that \( f_i \) and \( \bar{v}_i \) are linear in the estimation parameters: this allows assuming \( \theta = 1 \) because the parameter \( \theta \) cannot be distinguished from the overall scale of the other parameters.

The functional form in residual income of eqn (1) is justified because it ensures consistency with the micro-economic framework at the individual level (McFadden, 1981). In fact, the maximum of utilities \( v^* = \max_i u_i \) generates via Roy’s identity the demand for the discrete alternative:

\[
- \frac{\partial v^*}{\partial p_i} / \frac{\partial v^*}{\partial y} = \begin{cases} 1 & \text{if } u_i > u_j \quad i \neq j \\ 0 & \text{otherwise} \end{cases}
\]  

(2)

As stressed by Jara-Díaz and Videla (1990), income in eqns (2) should refer to the same period as the market demand for the discrete good, i.e. that of a unit consumption. If the econometric specification considers income in a longer time frame (typically a month or a year), Roy’s identity would still hold as soon as price is referred to money expenditure on the alternative in the same time frame.

The case without income effect is obtained when the systematic utility is linear in residual income with a coefficient that does not depend on alternative, as in this case income drops out of the econometric specification:

\[ v_i = \alpha \cdot (y - p_i) + \bar{v}_i \quad i = 1, \ldots, J \]  

(3)

Consider a policy inducing a change in prices from \( p_i' \) to \( p_i'' \), \( i = 1, \ldots, J \), and in quality from \( \bar{v}_i' \) to \( \bar{v}_i'' \), \( i = 1, \ldots, J \). The stochastic compensating variation \( cv \), conditional on the vector of before error terms \( \epsilon_i^T = [\epsilon_i', i = 1, \ldots, J] \) and the vector of after error terms \( \epsilon_i''^T = [\epsilon_i'', i = 1, \ldots, J] \), satisfies:

\[
\max_{i=1, \ldots, J}[v_i(y - p_i', \bar{v}_i') + \epsilon_i'] = \max_{i=1, \ldots, J}[v_i(y - cv - p_i'', \bar{v}_i'') + \epsilon_i'']
\]  

(4)

We consider each error term to be correlated between the state before the change and the state
after the change independently and identically across alternatives. We denote with $\rho$ the Pearson correlation coefficient. The theory of bivariate Gumbel distributions developed by Tiago de Oliveira (1980), and recalled in Garrow et al. (2010), can be used to generate a vector of after error terms $\mathbf{\varepsilon}''$ which has the following relationship with the vector of before error terms $\mathbf{\varepsilon}'$. Suppose $\varepsilon_i'$ follows a standard Gumbel, i.e. a Gumbel with location parameter equal to 0 and scale parameter equal to 1. Then, each component $\varepsilon_i''$ also follows a standard Gumbel and the Pearson correlation coefficient between $\varepsilon_i'$ and $\varepsilon_i''$ is the given value $\rho$. To satisfy this relationship, each after error term $\varepsilon_i''$ needs to be generated as follows:

$$
\varepsilon_i'' = \max[\varepsilon_i' + \ln \Phi, \eta + \ln(1 - \Phi)] \quad i=1,..,J
$$

(5)

where $\eta$ is a standard Gumbel distributed variable generated independently from $\varepsilon_i'$, and $\Phi$ satisfies:

$$
\rho(\Phi) = -\frac{6}{\pi^2} \int_0^\Phi \frac{\ln t}{1-t} \, dt
$$

(6)

The values of $\Phi$ for given $\rho$ are obtained from solving eqn (6). In particular, for $\rho=0$ we have $\Phi=0$, and for $\rho=1$ we have $\Phi=1$. As it is immediately verified, when $\rho=0$, i.e. zero correlation, we have that $\varepsilon_i'$ and $\varepsilon_i''$ are independent. When $\rho=1$, i.e. perfect correlation, we have $\varepsilon_i' = \varepsilon_i''$.

### 3. The case without income effect: a theoretical result

The following proposition provides the expectation of the compensating variation $E[cv]$ in a logit without income effect.

**Proposition.** The expectation of the compensating variation in a logit with the systematic utilities of the form (3) is given by the following expression:

$$
E[cv] = \frac{1}{\alpha} \left[ \ln \sum_{i=1}^J \exp(-\alpha \cdot p_i'' + \bar{v}_i'') - \ln \sum_{i=1}^J \exp(-\alpha \cdot p_i' + \bar{v}_i') \right]
$$

(7)

which holds for any before-after correlation $\rho$ of the random terms.

**Proof.** We follow the lines of the proof provided by McFadden (1999) for the case where the
random terms do not change before and after. In the event $B_{ik}$ that the individual chooses alternative $i$ before and alternative $k$ after the price and quality change and income compensation, we have the conditional compensating variation:

$$cv|_{B_{ik}} = \frac{1}{\alpha} \left[ -\alpha \cdot p_{ik}'' + \overline{\nu}_k'' - (\alpha \cdot p_{ik}' + \overline{\nu}_i'') + \varepsilon_{ik}'' - \varepsilon_i' \right]$$

(8)

By the law of total expectation:

$$E[cv] = \frac{1}{\alpha} \cdot \sum_{i=1}^{j} \sum_{k=1}^{j} P_{ik} \cdot \left\{ -\alpha \cdot p_{ik}'' + \overline{\nu}_k'' - (\alpha \cdot p_{ik}' + \overline{\nu}_i'') + E[\varepsilon_{ik}''|B_{ik}] \right\}$$

(9)

where $P_{ik}$ denotes the probability of the event $B_{ik}$. Since the expectation of the difference of two random variables equals the difference of the expectations, we can re-arrange eqn (9):

$$E[cv] = \frac{1}{\alpha} \cdot \sum_{i=1}^{j} \sum_{k=1}^{j} P_{ik} \cdot \left\{ -\alpha \cdot p_{ik}'' + \overline{\nu}_k'' - (\alpha \cdot p_{ik}' + \overline{\nu}_i'') - E[\varepsilon_{ik}''|B_{ik}] \right\} + \frac{1}{\alpha} \cdot \sum_{i=1}^{j} \sum_{k=1}^{j} P_{ik} \cdot \left\{ -\alpha \cdot p_{ik}' + \overline{\nu}_i' - E[\varepsilon_{ik}'|B_{ik}] \right\}$$

(10)

Denote now by $B_k''$ the event that $k$ is chosen after and by $P_k''$ the probability of this event. Denote by $B_i'$ the event that $i$ is chosen before and by $P_i'$ the probability of this event. Using again the law of total expectation:

$$E[cv] = \frac{1}{\alpha} \cdot \sum_{k=1}^{j} P_k'' \cdot \left\{ -\alpha \cdot p_{k}'' + \overline{\nu}_k'' - (\alpha \cdot p_{k}' + \overline{\nu}_i'') - E[\varepsilon_{k}''|B_k''] \right\} + \frac{1}{\alpha} \cdot \sum_{i=1}^{j} P_i' \cdot \left\{ -\alpha \cdot p_{i}' + \overline{\nu}_i' - E[\varepsilon_{i}'|B_i'] \right\}$$

(11)

Therefore, the expectation of the compensating variation $E[cv]$ equals the difference, normalised by the marginal utility of income, of the expectations before and after of the maximum utility:

$$E[cv] = \frac{1}{\alpha} \cdot \left\{ E\left[ \max_i ( -\alpha \cdot p_{i}'' + \overline{\nu}_i'' + \varepsilon_{i}'') \right] - E\left[ \max_i ( -\alpha \cdot p_{i}' + \overline{\nu}_i' + \varepsilon_{i}') \right] \right\}$$

(12)

The expectation of the maximum utility in the before state equals the logsum expression plus the Euler’s constant $\gamma$: $E\left[ \max_i ( -\alpha \cdot p_{i}'' + \overline{\nu}_i'' + \varepsilon_{i}'') \right] = \ln \sum_{i=1}^{j} \exp(-\alpha \cdot p_{i}' + \overline{\nu}_i') + \gamma$.

To compute the expectation of the maximum utility in the after state we need to take into account that the vector of the after error terms $\varepsilon''$ is not independent of the vector of the
before error terms $\varepsilon'$. The marginal density of the after error terms $\varepsilon''$, which is needed to compute the expectation of the maximum utility, depends on the correlation between the two vectors $\varepsilon'$ and $\varepsilon''$. The marginal density of the vector $\varepsilon''$ results from the product of the marginal densities of the components $\varepsilon_i''$ because they are i.i.d. As the density of each after component $\varepsilon_i''$ equals that of the before component $\varepsilon_i'$, also the expectation of the maximum utility in the after state will equal the logsum expression plus the Euler’s constant:

$$E\left[\max_i(-\alpha \cdot p_i'' + \nu_i'' + \varepsilon_i'')\right] = \ln \sum_i \exp(-\alpha \cdot p_i'' + \nu_i'') + \gamma . \Box$$

4. The case with income effect: numerical results

According to the proposition of the previous section, in logit without income effect the before-after correlation of the error terms does not matter in welfare measurement. The question arises whether this is also the case with income effect. In this section, we present the numerical results from the simulations carried out to assess the sensitivity of the expectation of the compensating variation $E[\text{cv}]$ to the before-after correlation of the error terms in logit with income effect.

We use a synthetic dataset which has been generated to simulate the choice of the transport mode of morning commuters to a city centre. The dataset is based on a kernel of revealed data which refer to current conditions in the city of Rome. A logit is considered with the systematic utility of the alternatives having the following linear-in-income specification:

$$v_m = \alpha_{mb} \cdot (y - p_{mb}) + \beta_1 \cdot t_{1,m} + \beta_2 \cdot t_{2,m} + ASC_m$$

$$v_b = \alpha_{mb} \cdot (y - p_{mb}) + \beta_1 \cdot t_{1,b} + \beta_2 \cdot t_{2,b} + ASC_b$$

$$v_c = \alpha_c \cdot (y - p_c) + \beta_2 \cdot t_{2,c}$$  \hspace{1cm} (13)

where the indexes $m$, $b$ and $c$ denote metro, bus and car, $t_i$ the access time (including the waiting time at station/stop), $t_2$ the in-vehicle time, $ASC$ the alternative-specific constant. The estimate of the coefficients are in Table 1. The values of the attributes are shown in Table 2.

The income effect is controlled by the difference between the marginal utility of income for the two public transport modes $\alpha_{mb}$ and the marginal utility of income for car $\alpha_c$. In the model estimated, preference for car results to be relatively higher for increasing income. Empirical evidence of income effect at the level of the transport mode choice is still lacking. It is not the aim of the paper to fill this gap. The specification adopted with constant alternative-specific marginal utilities of income makes extremely easy to assess the sensitivity
of the welfare measure to the size of the income effect.

The impacts of the policy considered on the attributes of the alternatives consist in a significant increase of the price of the car mode, +80 EUR/month. This policy, which reflects current orientations, especially in European countries, tending towards increased prices for transport modes that are less “sustainable”, makes income effect of peculiar relevance. In practice, the increase may result from parking pricing, which is a widely adopted policy, or from congestion pricing, a scheme found in a few European cities.

### Table 1. Logit model

<table>
<thead>
<tr>
<th>Coefficient Symbol</th>
<th>Name (unit)</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_{mb} )</td>
<td>Residual income public transport (metro &amp; bus)</td>
<td>0.00221</td>
</tr>
<tr>
<td>( \alpha_c )</td>
<td>Residual income car</td>
<td>0.00290</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>Access time</td>
<td>-0.19987</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>In-vehicle time</td>
<td>-0.09850</td>
</tr>
<tr>
<td>( ASC_{\text{m}} )</td>
<td>Alternative specific constant metro</td>
<td>2.90246</td>
</tr>
<tr>
<td>( ASC_{b} )</td>
<td>Alternative specific constant bus</td>
<td>1.85917</td>
</tr>
</tbody>
</table>

Number of (simulated) observations: 2916  
Rho-squared: 0.05425

### Table 2. Values of the attributes

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>Name (unit)</td>
</tr>
<tr>
<td>( y )</td>
<td>Income (EUR/month)</td>
</tr>
<tr>
<td>( p_{mb} )</td>
<td>Monetary cost public transport (EUR/month)</td>
</tr>
<tr>
<td>( p_c )</td>
<td>Monetary cost car (EUR/month)</td>
</tr>
<tr>
<td>( t_{1,m} )</td>
<td>Access time metro (minutes)</td>
</tr>
<tr>
<td>( t_{1,b} )</td>
<td>Access time bus (minutes)</td>
</tr>
<tr>
<td>( t_{2,m} )</td>
<td>In-vehicle time metro (minutes)</td>
</tr>
<tr>
<td>( t_{2,b} )</td>
<td>In-vehicle time bus (minutes)</td>
</tr>
<tr>
<td>( t_{2,c} )</td>
<td>In-vehicle time car (minutes)</td>
</tr>
</tbody>
</table>

We use simulation to compute \( E[cv] \) for different values of the before-after correlation \( \rho \) of the error terms. Pairs, before and after, of each error term are created using the methodology.
reported on in section 2. The range from zero \((\rho = 0)\) to perfect correlation \((\rho = 1)\) is explored.

The simulation uses \(T\) draws of the error terms generated from a pseudo-random sequence, and generates \(R\) distinct sequences. If \(t\) denoted the draw in the sequence, \(r\) the replication, the unbiased estimator \(cv_{TR}\) of \(E[cv]\) is given by:

\[
 cv_{TR} = \frac{1}{T \cdot R} \cdot \sum_{r=1}^{R} \sum_{t=1}^{T} cv_{tr}
\]  

(14)

The 95% confidence intervals are \(cv_{TR} \pm 1.96 \cdot \sqrt{\frac{S^2}{R}}\) where \(S^2\) is the sample variance:

\[
 S^2 = \frac{1}{R-1} \cdot \sum_{r=1}^{R} \left[ \frac{1}{T} \cdot \sum_{t=1}^{T} cv_{tr} - cv_{TR} \right]^2.
\]

We use \(T=10,000\) and \(R=1,000\).

For each draw, the compensating variation \(cv_{tr}\) that solves eqn (4) is found as follows. Eqn (4) is constituted by a constant term \(u_0\) in the left-hand side. The right-hand side is a function of the compensating variation. One needs first to solve the equation: \(u_0 = u_i(cv_{tr})\) for each alternative. As the systematic utility of each alternative \(v_i\) is assumed to be monotonic in residual income, each eqn will have one solution only. Then, because of the decreasing monotonicity of the systematic utilities with respect to the compensating variation, the solution of eqn (4) will be, as it is easily seen, the maximum of the solutions of these alternative-specific eqns.

The results for \(E[cv]\) obtained from simulation for varying \(\rho\) are compared, in terms of percentage difference, with the formula by Dagsvik and Karlström (2005) which provides \(E[cv]\) when \(\rho = 1:\)

\[
 E[cv] = y - \sum_{i=1}^{J} \left\{ \mu_{ui} \right\} P[g_1(m),...,g_J(m)] \cdot dm 
\]

(15)

where:

\(\mu_{ui}\) is the expenditure required after the change that would keep the individual at
the same level of utility as before the change if he chose alternative $i$ both before and after the change, i.e.

$$v_i(y - p_i', \bar{v}_i') = v_i(\mu_i - p_i'', \bar{v}_i'') \quad i = 1, \ldots, J$$

and

$$g_i(m) = \max [v_i(y - p_i', \bar{v}_i'), v_i(m - p_i'', \bar{v}_i'')] \quad i = 1, \ldots, J$$

Table 3 shows the results. Four values of the marginal utility of income for car $\alpha_c$ are considered: 0.00221, 0.00230, 0.00260 and 0.00290. The latter is the estimation value, the other three are considered to assess the sensitivity to the size of the income effect. The value $\alpha_c = 0.00221$ represents the case without income effect because the marginal utilities of income for the two public transport modes $\alpha_{mb}$ and for car $\alpha_c$ are equal.

The variability of the expectation of the compensating variation with the correlation $\rho$, measured by the percentage difference on $E[cv]$ for $\rho = 1$, amplifies as the income effect gets larger, i.e. when the Table is read from top to bottom. The highest percentage difference on $\rho = 1$ is obtained with the largest income effect ($\alpha_c = 0.00290$) when $\rho = 0$. The value is in the range of 32%. This proves that the variability of the stochastic measure with the correlation can be numerically significant. The values of the percentage differences that are found when there is no income effect ($\alpha_c = 0.00221$), which theoretically should be zero, are justified in the light of the finite number of draws, i.e. finite size of the population, used in the simulations.

Another interesting issue is the variation of the stochastic measure around the expectation. The Table shows that the 95% confidence intervals get narrower as we move towards perfect correlation ($\rho = 1$), i.e. moving along the Table from left to right, a result also found by Zhao et al. (2012) in their investigation restricted to the case without income effect.

5. Conclusions

Without income effect the logsum is a robust formula which provides the expectation of the compensating variation for any value of the before-after error term correlation. This is a theoretical property and is validated numerically. In contrast, in the case with income effect the assumption made on the dynamics of the error terms does matter. Numerical evidence suggests that the expectation of the compensating variation can change significantly if a different assumption on the before-after correlation is made and that the change increases with
the size of the income effect. This is a critical issue without the means to test the assumption empirically. The paper is a small step in the area of dynamic random utility models. A natural extension for future enquiry relates to random utility models with different patterns of correlation across alternatives.

References


Table 3. Sensitivity to before-after error term correlation

<table>
<thead>
<tr>
<th>Before-after correlation $\rho$</th>
<th>0</th>
<th>0.2</th>
<th>0.4</th>
<th>0.6</th>
<th>0.8</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_c = 0.00221$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average cv from simulation $^*$</td>
<td>-10.04</td>
<td>-9.96</td>
<td>-9.95</td>
<td>-9.94</td>
<td>-9.91</td>
<td>-9.93</td>
</tr>
<tr>
<td>Formula $E[cv] ^*$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-9.92</td>
</tr>
<tr>
<td>% difference on formula $E[cv]$</td>
<td>1.28%</td>
<td>0.43%</td>
<td>0.38%</td>
<td>0.25%</td>
<td>-0.02%</td>
<td>0.13%</td>
</tr>
<tr>
<td>$\alpha_c = 0.00230$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average cv from simulation $^*$</td>
<td>-10.96</td>
<td>-10.88</td>
<td>-10.86</td>
<td>-10.80</td>
<td>-10.71</td>
<td>-10.63</td>
</tr>
<tr>
<td>Formula $E[cv] ^*$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-10.62</td>
</tr>
<tr>
<td>% difference on formula $E[cv]$</td>
<td>3.22%</td>
<td>2.49%</td>
<td>2.25%</td>
<td>1.71%</td>
<td>0.85%</td>
<td>0.15%</td>
</tr>
<tr>
<td>$\alpha_c = 0.00260$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formula $E[cv] ^*$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-13.46</td>
</tr>
<tr>
<td>% difference on formula $E[cv]$</td>
<td>-6.07%</td>
<td>-0.56%</td>
<td>2.59%</td>
<td>3.68%</td>
<td>2.69%</td>
<td>0.09%</td>
</tr>
<tr>
<td>$\alpha_c = 0.00290$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average cv from simulation $^*$</td>
<td>-11.51</td>
<td>-14.49</td>
<td>-16.33</td>
<td>-17.31</td>
<td>-17.45</td>
<td>-16.90</td>
</tr>
<tr>
<td>Formula $E[cv] ^*$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-16.88</td>
</tr>
<tr>
<td>% difference on formula $E[cv]$</td>
<td>-31.82%</td>
<td>-14.20%</td>
<td>-3.28%</td>
<td>2.53%</td>
<td>3.37%</td>
<td>0.07%</td>
</tr>
</tbody>
</table>

$^*$ Values of $cv$ and $E[cv]$ in EUR/month