Economics Letters 120 (2013) 57–60

Contents lists available at SciVerse ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Integration of choice probabilities in logit

Paolo Delle Site *

DICEA - Department of Civil, Architectural and Environmental Engineering, University of Rome "La Sapienza", Via Eudossiana, 18 – 00184, Rome, Italy
CTL - Research Centre for Transport and Logistics, University of Rome "La Sapienza", Via Eudossiana, 18 – 00184, Rome, Italy

HIGHLIGHTS

- Social utility is the latent objective function of a random utility model.
- Welfare measures are easy to compute and not affected by the before–after correlation.
- Social indirect utility is expressed by a sum of ordinary indefinite integrals.
- Without income effects social indirect utility reduces to the logsum formula.
- For translog specification social indirect utility has a simple closed form.

ARTICLE INFO

Article history:
Received 3 December 2012
Accepted 15 March 2013
Available online 26 March 2013

JEL classification:
C25
D11

Keywords:
Logit
Income effects
Social utility
Roy's identity

ABSTRACT

We find for logit with income effects a function that generates choice probabilities via Roy's identity. We show that it possesses all the properties to qualify as an indirect utility, it has a closed-form expression for the practically interesting translog specification of the systematic utilities, it reduces in the case without income effects to the expectation of the maximum utility. We discuss the use of the findings in welfare measurement and the extension to observed heterogeneity and mixed logit.

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. We consider the problem of demand integrability at the level of choice probabilities. Integration at this level makes it possible to treat demand for discrete alternatives within the classical framework of deterministic utility maximisation, and describe demand as if it were generated by a representative consumer with fractional consumption rates.

The utility function that results from this integration has an interpretation which depends on the interpretation of the random utility model. It is the latent objective function of a population of individuals characterised by identical systematic utilities and random tastes, or the latent objective function of a single individual with random tastes over repeated choices. McFadden (1981) refers to this function as social utility, implying that the first of the interpretations above is assumed. Hereafter, we keep McFadden's terminology with the warning that also the second of the interpretations above applies.

Social utility is not only a theoretical construct providing a mathematical programming formulation of random utility models. It is also of interest for welfare analysis. A compensating variation related to the social utility is easily computed by simply equating the social indirect utility in the state before the change of price and quality with the social utility in the state after the change and compensation. In addition to the computational advantage, welfare measures based on social utility have the merit, highlighted by Herriges and Kling (1999), of being independent of the before–after correlation of the random terms, an issue which needs to be taken into account when welfare measures are computed according to the mainstream approach which regards the compensating variation as a random quantity.

The mainstream approach considers as welfare measure the expectation of the compensating variation. To compute this, for reasons of tractability, the assumption is made that the random terms do not change between the state before and the state

* Tel.: +39 0644585923; fax: +39 0644585774.
E-mail address: paolo.dellesite@uniroma1.it.

0165-1765/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.econlet.2013.03.027

after, but other before–after correlation patterns can be considered to take into account changes in unobserved attributes or intrapersonal taste variation. Robustness of the expectation of the compensating variation to such assumptions is not guaranteed and is the subject of current research (Zhao et al., 2012). This is a drawback without the means to test the assumptions empirically.

A social indirect utility is available when the individual conditional indirect utilities do not exhibit income effects, i.e. when income does not influence choice. McFadden (1981) has shown that the expectation of the maximum utility satisfies Roy's identity and has all the other properties needed to qualify as an indirect utility. Later, Anderson et al. (1988) have found the related direct utility.

Income effects imply that a consumer who is facing a set of alternatives and alternative attributes may make different choices at different levels of her income. This circumstance is of interest when the expenditure on the alternatives is a significant share of income and the alternatives differ significantly in price. Income effects are used to be modelled by a translog specification of the systematic utilities (Herriges and Kling, 1999; Tra, 2013).

A social indirect utility for logit probabilities is not yet found when income effects are present. In this case, as McFadden (1999) notes, the expectation of the maximum utility does no longer satisfy Roy's identity. The present note provides a social indirect utility function when income effects are present. The approach in Hausman (1981), who integrates demand by solving the system of differential equations based on Roy's identity, is followed.

The organisation is as follows. In Section 2.1 the micro-economic foundation of logit is outlined. Section 2.2 deals with social utility. A proposition gives the form of the social indirect utility for the general specification of systematic utilities. A closed-form expression is then provided for the translog specification. The use of the results in welfare analysis is discussed. Section 2.3 concludes with extension to observed heterogeneity and mixed logit.

2. Integration of choice probabilities

2.1. Micro-economic foundation

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. We express u_i by the additively separable structure: $u_i = v_i + \varepsilon_i$, $i = 1, \ldots, J$, where v_i is the deterministic component, referred to as systematic utility, and ε_i is the random component. This structure for u_i defines the class of additive random utility models.

The systematic utility v_i of each alternative depends on income y, on the price p_i of the alternative, and on other qualitative attributes of the alternative distinct from price. Prices and income are deflated by the price of an outside commodity. Taking an additive structure, we have $v_i = v_i(y, p_i, \overline{v}_i) = f(y, p_i) + \overline{v}_i$, $i = 1, \ldots, J$, where $\overline{v}_i$ is a function of the qualitative attributes.

For consistency with the classical utility maximisation framework of micro-economics (based on McFadden (1981)), the maximum of utilities $u^* = \max_{i=1,\ldots,J} u_i$ needs to generate via Roy's identity the demand for the discrete alternative:

$$-\frac{\partial u^*}{\partial p_i}\bigg/\frac{\partial u^*}{\partial y} = \begin{cases} 1 & u_i \geq u_j \; i \neq j \\ 0 & \text{otherwise} \end{cases} \quad i = 1, \ldots, J. \tag{1}$$

We assume a functional form of systematic utilities in residual income $(y - p_i)$:

$$v_i = f(y - p_i) + \overline{v}_i \quad i = 1, \ldots, J \tag{2}$$

because this form satisfies Roy's identity (1). In fact, it satisfies each of the Roy's identities obtained by conditioning on the chosen alternative: $-\frac{\partial u_i}{\partial p_i}/\frac{\partial u_i}{\partial y} = 1, i = 1, \ldots, J$.

The function $f(y - p_i)$ is assumed to be strictly increasing in residual income $(y - p_i)$.

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

$$v_i = \alpha \cdot (y - p_i) + \overline{v}_i \quad i = 1, \ldots, J. \tag{3}$$

A functional form with income effects is the translog:

$$v_i = \alpha \cdot \ln(y - p_i) + \overline{v}_i \quad i = 1, \ldots, J. \tag{4}$$

The translog form has nice properties. The marginal utility of income depends on the alternative and is inversely related to the expenditure on the numéraire, being $\partial v_i/\partial y = \alpha/(y - p_i)$. The intuition is that a dollar has a higher value when the budget available for other purchases is lower. The marginal utility of income decreases with income, which is consistent with the intuition that a dollar provides a decreasing utility as income rises. In addition, if income increases, the difference in utilities of two alternatives with different prices, ceteris paribus, decreases, which means that the difference in market shares decreases too. The intuition is that with higher income the difference in price has a lower impact on demand.

The probability of choosing alternative i is $P_i = \Pr(u_i > u_j, j \neq i)$. Under the independently and identically distributed type I extreme value (Gumbel) assumption for the random component, the probability P_i that an individual chooses alternative i is the logit function: $P_i = e^{v_i/\theta}/\sum_{j=1}^{J} e^{v_j/\theta}$ where θ is the scale parameter. We assume that f and $\overline{v}_i$ are linear in the estimation parameters: this allows assuming $\theta = 1$ because the parameter θ cannot be distinguished from the overall scale of the other parameters.

2.2. Social indirect utility

A function V is termed a social indirect utility if it yields the choice probabilities via Roy's identity:

$$-\frac{\partial V}{\partial p_i}\bigg/\frac{\partial V}{\partial y} = P_i \quad i = 1, \ldots, J. \tag{5}$$

The function V measures the maximum, associated with a given price and income level, of the function that when maximised under the budget constraint yields the choice probabilities.

Consider first the case without income effects of Eqs. (3). McFadden (1981) has shown that a social indirect utility function V is given by the expectation $\overline{V}$ of the maximum utility:

$$\begin{aligned} V &= \overline{V} = E\left[\max_{i=1,\ldots,J} u_i\right] \\ &= \alpha \cdot y + E\left[\max_{i=1,\ldots,J} (-\alpha \cdot p_i + \overline{v}_i + \varepsilon_i)\right]. \end{aligned} \tag{6}$$

Satisfaction of Roy's identity (5) follows from the following property: the gradient of the expectation $\overline{V}$ of the maximum utility with respect to the systematic utilities equals the vector of choice probabilities: $\partial\overline{V}/\partial v_i = P_i$, $i = 1, \ldots, J$. For logit, Eq. (6) reduces to the logsum form:

$$\overline{V} = \alpha \cdot y + \ln \sum_{i=1}^{J} e^{-\alpha \cdot p_i + \overline{v}_i}. \tag{7}$$

In the general case of income effects with the functional form of Eqs. (2), the expectation $\overline{V}$ of the maximum utility does not satisfy Roy's identity (5). In fact, according to Roy's identity, the ratios $(\partial\overline{V}/\partial p_i)/P_i$, $i = 1, \ldots, J$, should be equal across alternatives, while they are not as we have by applying the chain rule of derivation: $(\partial \overline{V}/\partial p_i)/P_i = \partial v_i/\partial p_i$.

The objective here is, therefore, to find a function $V[v_i(y - p_i, \overline{v}_i); i = 1, \ldots, J]$ of prices and income, composed with the systematic utilities, which satisfies Eqs. (5).

We note, first, that we can write $\partial v_i/\partial p_i, i = 1, \ldots, J$, as a function of v_i: $\partial v_i/\partial p_i = \varphi_i(v_i)$. In fact, let $\chi_i = y - p_i$. We have by the chain rule of derivation: $\partial v_i/\partial p_i = \partial v_i/\partial \chi_i \cdot \partial \chi_i/\partial p_i = -\partial v_i/\partial \chi_i = \psi_i(\chi_i)$. Since $f(\chi_i)$ is increasing, it is invertible: $\chi_i = \chi_i(v_i)$. Hence $\partial v_i/\partial p_i = \varphi_i(v_i)$.

By an application of the chain rule of derivation for multivariable functions (Sydsæter et al., 1998, p. 26) to $\partial V/\partial y$ we obtain from Eqs. (5) for logit:

$$-\frac{\frac{\partial V}{\partial v_i} \cdot \frac{\partial v_i}{\partial p_i}}{-\sum_{j=1}^{J} \frac{\partial V}{\partial v_j} \cdot \frac{\partial v_j}{\partial p_j}} = \frac{e^{v_i}}{\sum_{j=1}^{J} e^{v_j}} \quad i = 1, \ldots, J. \tag{8}$$

The problem reduces to finding a function of systematic utilities $V(v_i, i = 1, \ldots, J)$ satisfying:

$$\frac{\partial V}{\partial v_i} = -\frac{e^{v_i}}{\varphi_i(v_i)} = \omega_i(v_i) \quad i = 1, \ldots, J. \tag{9}$$

Systems of partial differential equations of the form (9) are frequently found in applied mathematics, in particular physics. Solving (9) is equivalent to finding a function V with a given gradient $\boldsymbol{\omega}(v_i, i = 1, \ldots, J)$. If such function exists, the vector $\boldsymbol{\omega}$ is called a conservative vector field and the function V is called a scalar potential. A necessary condition for the existence of a potential is that the Jacobian of the vector field is symmetric. This condition is satisfied in the case here, because each component of the vector $\boldsymbol{\omega}$ depends on one component only of the systematic utility and, therefore, $\partial \omega_i/\partial v_j = 0, i \neq j$.

The following provides a social indirect utility function for logit as sum of ordinary indefinite integrals.

Proposition. *A social indirect utility function $V: \mathbb{R}^J_{++} \times \mathbb{R}_{++} \to \mathbb{R}$ yielding logit choice probabilities via Roy's identity (5) is:*

$$V = V[v_i(y - p_i, \overline{v}_i), i = 1, \ldots, J] = -\sum_{i=1}^{J} \int \frac{e^{v_i}}{\varphi_i(v_i)} \cdot dv_i. \tag{10}$$

Proof. The proof consists in showing that the function V in Eq. (10) qualifies as an indirect utility function because (Mas-Colell et al., 1995):

(a) satisfies Roy's identity;
(b) is continuous in $p_i, i = 1, \ldots, J$, and y;
(c) is homogeneous of degree 0 in $p_i, i = 1, \ldots, J$, and y;
(d) is non-increasing in $p_i, i = 1, \ldots, J$;
(e) is non-decreasing in y;
(f) is quasi-convex in $p_i, i = 1, \ldots, J$, and y.

The proof uses extensively the chain rule of derivation.

Part (a). Satisfaction of Roy's identity is verified by substitution of V of Eq. (10) into Eqs. (8) or (9).

Part (b). The first derivatives of V exist and are continuous. Hence V is continuous.

Part (c). Homogeneity of degree 0 of V follows immediately from the normalisation of prices $p_i, i = 1, \ldots, J$, and of income y.

Part (d). We aim to prove that $\frac{\partial V}{\partial p_i} \leq 0, i = 1, \ldots, J$. By the chain rule:

$$\frac{\partial V}{\partial p_i} = \frac{\partial V}{\partial v_i} \cdot \frac{\partial v_i}{\partial p_i} = -\frac{e^{v_i}}{\varphi_i(v_i)} \cdot \varphi_i(v_i) = -e^{v_i} < 0 \quad i = 1, \ldots, J.$$

Part (e). We aim to prove that $\frac{\partial V}{\partial y} \geq 0$. By the chain rule:

$$\frac{\partial V}{\partial y} = \sum_{i=1}^{J} \frac{\partial V}{\partial v_i} \cdot \frac{\partial v_i}{\partial y} = -\sum_{i=1}^{J} \frac{e^{v_i}}{\varphi_i(v_i)} \cdot \frac{\partial v_i}{\partial \chi_i} > 0$$

where the latter inequality follows from $\varphi_i(v_i) < 0$ and $\frac{\partial v_i}{\partial \chi_i} > 0$, which, in turn, result from $f(\chi_i)$ being increasing.

Part (f). By a theorem (Florenzano et al., 2001, p. 129), given any two feasible vectors $(p_i^*, i = 1, \ldots, J; y^*)$ and $(p_i^{**}, i = 1, \ldots, J; y^{**})$, V is quasi-convex if and only if the following function of the variable λ, with $\lambda \in [0, 1]$:

$$V[\lambda \cdot p_i^* + (1 - \lambda) \cdot p_i^{**}, i = 1, \ldots, J; \lambda \cdot y^* + (1 - \lambda) \cdot y^{**}]$$

is quasi-convex. We prove the quasi-convexity of $V(\lambda)$ by showing that this function is monotonic, because any monotonic function is also quasi-convex. First, we write v_i as function of λ in compact form:

$$v_i = f(\lambda) + \overline{v}_i = f[\chi_i(\lambda)] + \overline{v}_i = f(\lambda \cdot A_i + B_i) + \overline{v}_i \quad i = 1, \ldots, J$$

where:

$$A_i = y^* - y^{**} - (p_i^* - p_i^{**})$$
$$B_i = y^{**} - p_i^{**}.$$

By the chain rule:

$$\frac{dV}{d\lambda} = \sum_{i=1}^{J} \frac{\partial V}{\partial v_i} \cdot \frac{\partial v_i}{\partial \lambda} = -\sum_{i=1}^{J} \frac{e^{v_i}}{\varphi_i(v_i)} \cdot \frac{\partial v_i}{\partial \chi_i} \cdot \frac{\partial \chi_i}{\partial \lambda} = -\sum_{i=1}^{J} \frac{e^{v_i}}{\varphi_i(v_i)} \cdot \frac{\partial v_i}{\partial \chi_i} \cdot A_i.$$

The resulting sign of $\frac{dV}{d\lambda}$ is independent of λ hence $V(\lambda)$ is monotonic. □

The following establishes the relationship between the social utility V in Eq. (10) and the expectation $\overline{V}$ of the maximum utility.

Corollary 1. *In the case of absence of income effects of Eqs. (3), the social indirect utility V of Eq. (10) is a monotonic transformation of the expectation $\overline{V}$ of the maximum utility.*

Proof. When the systematic utility has the form of Eqs. (3), the integrals in Eq. (10) can be solved analytically and we obtain for the social utility: $V = \frac{1}{\alpha} \sum_{i=1}^{J} e^{v_i} + k$, where k is the integration constant, because $\varphi_i(v_i) = -\alpha, i = 1, \ldots, J$. Then, we can transform V and obtain by a strictly increasing transformation the expectation $\overline{V}$ of the maximum utility given by Eq. (7). □

In the case of a translog specification of the systematic utilities, the integrals in Eq. (10) can be solved analytically and the social indirect utility V takes a simple closed form.

Corollary 2. *In the case of the translog specification of Eqs. (4), the social indirect utility V of Eq. (10) takes the following closed-form expression:*

$$V = \sum_{i=1}^{J} \frac{1}{1 + \alpha} \cdot \frac{e^{\frac{1+\alpha}{\alpha} \cdot v_i}}{e^{\overline{v}_i/\alpha}} + k \tag{11}$$

where k is the integration constant.

Proof. Eq. (11) follows from solving the integrals in Eq. (10), having taken into account that with a translog specification: $\varphi_i(v_i) = -\frac{\alpha \cdot e^{\overline{v}_i/\alpha}}{e^{v_i/\alpha}}, i = 1, \ldots, J$. □

The social indirect utility function V in Eq. (11) is not a monotonic transformation of the expectation $\overline{V}$ of the maximum utility. This means that in the case with income effects the ordering of states provided by the social indirect utility V can be different from the ordering provided by the expectation of the maximum utility.

Welfare judgements based on social utility retain, generally, their positive, or descriptive, value, i.e. the desirability of states is judged according to the latent objective function of the population of individuals with identical systematic utilities and random tastes, or of the single individual with random tastes over repeated choices. In the interpretation of random utility models that considers a population of individuals with random tastes, in the case without income effects the value is strengthened. The expectation of the maximum utility is a utilitarian welfare function. Therefore, welfare judgements have also normative value because they are related to the individual preferences.

2.3. Extension to observed heterogeneity and mixed logit

Consider a population of individuals characterised by random tastes and observed heterogeneity, as an example a population with heterogeneous income. Consider the distribution density of income $h(y)$ and the social indirect utility function of Eq. (10) $V(y)$ conditional on the income y. The unconditional probabilities are the expectations over the distribution of income: $P_i = \int_y P_i(y) \cdot h(y) \cdot dy, i = 1, \ldots, J$, where the conditional probabilities $P_i(y)$, $i = 1, \ldots, J$, clearly satisfy Roy's identity: $-\frac{\partial V(y)}{\partial p_i} / \frac{\partial V(y)}{\partial y} = P_i(y)$, $i = 1, \ldots, J$.

We could define, similarly to probabilities, an unconditional social indirect utility function but this would not yield the unconditional probabilities via Roy's identity. Notice here the difference with the case without income effects where probabilities are independent of income and, therefore, a social indirect utility function is both a conditional and unconditional function yielding probabilities via Roy's identity.

The interest in mixed logit originates from two motivations (Hensher and Greene, 2003): to accommodate unobserved heterogeneity in the sensitivity of individuals to observed attributes (random-coefficients structure), and to accommodate flexible patterns of correlation of the random terms across alternatives (error-components structure). McFadden and Train (2000) have shown that the probabilities of any random utility model can be approximated with a desired accuracy by a mixed logit.

Suppose a logit model with $v_i = v_i(y - p_i, \overline{v}_i, \boldsymbol{\alpha}), i = 1, \ldots, J$, where $\boldsymbol{\alpha}$ is a latent mixing vector with distribution density $K(\boldsymbol{\alpha})$. Consider the social indirect utility function $V(\boldsymbol{\alpha})$ of Eq. (10) conditional on the vector $\boldsymbol{\alpha}$.

The mixtures of probabilities are given by: $P_i = \int_{\boldsymbol{\alpha}} P_i(\boldsymbol{\alpha}) \cdot K(\boldsymbol{\alpha}) \cdot d\boldsymbol{\alpha}, i = 1, \ldots, J$, where the conditional probabilities $P_i(\boldsymbol{\alpha}), i = 1, \ldots, J$, satisfy Roy's identity: $-\frac{\partial V(\boldsymbol{\alpha})}{\partial p_i} / \frac{\partial V(\boldsymbol{\alpha})}{\partial y} = P_i(\boldsymbol{\alpha}), i = 1, \ldots, J$.

We could define, similarly to probabilities, a mixture of social indirect utilities but this mixture would not yield the mixtures of probabilities via Roy's identity.

For welfare measurement in practice, in both cases of observed heterogeneity and mixed logit we can compute a conditional welfare measure based on the conditional social indirect utility, and obtain then the unconditional welfare measure as expectation over the distribution of the income or of the latent mixing vector.

References

Anderson, S.P., De Palma, A., Thisse, J.-F., 1988. A representative consumer theory of the logit model. International Economic Review 29 (3), 461–466.
Florenzano, M., Le Van, C., Gourdel, P., 2001. Finite Dimensional Convexity and Optimisation. Springer, Berlin.
Hausman, J.A., 1981. Exact consumer's surplus and deadweight loss. The American Economic Review 71 (4), 662–676.
Hensher, D.A., Greene, W.H., 2003. The mixed logit model: the state of practice. Transportation 30 (2), 133–176.
Herriges, J.A., Kling, C.L., 1999. Nonlinear income effects in random utility models. The Review of Economics and Statistics 81 (1), 62–72.
Mas-Colell, A., Whinston, M.D., Green, J.R., 1995. Microeconomic Theory. Oxford University Press, New York.
McFadden, D., 1981. Econometric models of probabilistic choice. In: Manski, C., McFadden, D. (Eds.), Structural Analysis of Discrete Data with Econometric Applications. MIT Press, Cambridge, MA, pp. 198–272.
McFadden, D., 1999. Computing willingness-to-pay in random utility models. In: Melvin, R., Moore, J.C., Riezman, R. (Eds.), Trade Theory and Econometrics—Essays in Honor of John Chipman. Routledge, London, pp. 253–274.
McFadden, D., Train, K., 2000. Mixed MNL models for discrete response. Journal of Applied Econometrics 15 (5), 447–470.
Sydsæter, K., Strøm, A., Berck, P., 1998. Economists' Mathematical Manual. Springer, Berlin.
Tra, C.I., 2013. Nonlinear income effects in random utility models: revisiting the accuracy of the representative consumer approximation. Applied Economics 45 (1), 55–63.
Zhao, Y., Kockelman, K., Karlström, A., 2012. Welfare calculations in discrete choice settings: an exploratory analysis of error term correlation with finite populations. Transport Policy 19 (1), 76–84.