A Note on Price Effects in Conditional Logit Models

Department of Economics and Statistics "Cognetti de Martiis"
Campus Luigi Einaudi, Lungo Dora Siena 100/A, 10153 Torino (Italy)
www.est.unito.it

Dipartimento
Economia e Statistica
Cognetti de Martiis

Working Paper Series

45/14

A NOTE ON PRICE EFFECTS IN CONDITIONAL LOGIT MODELS

UGO COLOMBINO

UNIVERSITÀ
DEGLI STUDI
DI TORINO

A Note on Price Effects in Conditional Logit Models

Ugo Colombino

Department of Economics and Statistics Cognetti De Martiis

1. Introduction

We consider a setting in which an individual chooses among M alternatives. The utility attached to alternative j is

$$U_j = U_j(z_j, m - \pi_j) \tag{1.1}$$

where $z_j = (z_{j1}, ..., z_{jN})'$ is a vector of N attributes, *m* is the individual's exogenous wealth, π_j is the price of alternative *j*. Typically, π depends on *z* :

$$\pi_j = \pi(z_{j1}, ..., z_{jN}) \tag{1.2}$$

Example 1.

The alternatives are standard consumption bundles, each component of *z* representing the quantity of a consumption good. Accordingly, π_j is the price of bundle j: in the simplest (textbook) case, $\pi_j = \sum_{i=1}^{N} w_i z_{ji}$, where w_i is the (constant) unit price of good i .

Example 2.

The alternatives are cars of different types. The attributes measure characteristics such as maximum speed, number of seats, interior space, presence of ABS etc. π_j is the price of type j. Also, $\partial\pi_j / \partial z_{ji}$ is the marginal price of attribute i (in type j). Analogous examples are generated by replacing cars with goods or services that can be defined as vectors of attributes, such as houses, computers, plant locations, fishing or hiking sites, telephone calls patterns etc. (e.g. Train, 1980; McFadden, 1997; Train et al., 1987; Colombino, 1998; Trajtenberg, 1989).

Example 3.

The alternatives are jobs, job *j* being characterised by h_j hours required. In the simplest case, the utility attached to job *j* would be $U_j(z_j, m - wz_j)$, where $z_j \equiv -h_j$ and *w* is a fixed wage rate. More generally, z_j might be a vector, as in Example 1 and 2, and $-h_j$ one of its components (e.g. Van Soest, 1995; Aaberge et al. 1999; Colombino, 2013)

In what follows we limit ourselves to the special case where the marginal prices of goods, attributes, or characteristics are constant, i.e.

$$\pi_j = \sum_{i=1}^{N} w_i z_{ji} \tag{1.3}$$

In Example 1, if the goods are produced by a perfectly competitive industry, w_i is the minimum average cost of producing good *i*. In Example 2, if the cars are produced by a perfectly competitive industry, the marginal price of attribute *i* is the minimum unit production cost of attribute *i*. In Example 3, *w* is the wage rate.

We assume that the analyst specifies:

$$U_j = V(z_j, m - \pi_j) + \varepsilon_j \tag{1.4}$$

where $V(z_j, m - \pi_j)$ is a parametric function and ε_j is i.i.d. Type I Extreme Value random variable. It is well known that under the above assumptions, the probability that alternative j is chosen has the following expression (e.g. Ben-Akiva, M., and Lerman, S.R., 1985):

$$P_j = \frac{\exp(V_j)}{\sum_{k=1}^{M} \exp(V_k)} \tag{1.5}$$

We can define the expectation of the chosen value z_i^* as

$$E(z_{*i}) = \sum_{j=1}^{M} z_{ji} P_j \tag{1.6}$$

We are interested in evaluating the effect of w_k upon $E(z_{*i})$, i.e. the effect of the price of attribute *k* upon the expected value of the chosen quantity of attribute *i*. This is the usual focus of interest in standard consumer theory, as in Example 1. In the cases illustrated by example 2, the literature has focussed upon a different question: what is the effect of π_k upon P_j, e.g. the effect of the price of car type *k* upon the probability that car type *j* is chosen? The following expressions are well known:

$$\frac{\partial P_j}{\partial \pi_k} = \frac{\partial V_k}{\partial \pi_k} P_j P_k \quad (1.7)$$
$$\frac{\partial P_j}{\partial \pi_j} = \frac{\partial V_j}{\partial \pi_j} P_j \left(1 - P_j\right)$$

However, also in this setting, we might be interested in a different question, namely the effect of the price of maximum speed upon the expected value of the chosen maximum speed. Analogously, in the case illustrated by Example 3, we might be interested in the effect of the wage rate upon the expected hours of work. We are interested in uncovering the implications of (1.5) upon this type of price effect.

2. Price Effects

Using (1.6) we have:

$$\frac{\partial E(z_i^*)}{\partial w_k} = \sum_{j=1}^{M} z_{jk} \frac{\partial P_j}{\partial w_k} \qquad (2.1)$$

We write P_j as

$$P_j = \frac{1}{\sum_{i=1}^{M} \exp(V_i - V_j)} \qquad (2.2)$$

Then we find:

$$\frac{\partial P_j}{\partial w_k} = -\frac{\sum_i \exp(V_i - V_j)\left(\frac{\partial V_i}{\partial w_k} - \frac{\partial V_j}{\partial w_k}\right)}{\left(\sum_i \exp(V_i - V_j)\right)^2} =$$

$$= -P_j^2 \sum_i \frac{\exp(V_i)}{\exp(V_j)}\left(\frac{\partial V_i}{\partial w_k} - \frac{\partial V_j}{\partial w_k}\right) = -P_j^2 \sum_i \frac{\frac{\exp(V_i)}{\sum_x \exp(V_x)}}{\frac{\exp(V_j)}{\sum_x \exp(V_x)}}\left(\frac{\partial V_i}{\partial w_k} - \frac{\partial V_j}{\partial w_k}\right) = \tag{2.3}$$

$$= -P_j \sum_i P_i \left(\frac{\partial V_i}{\partial w_k} - \frac{\partial V_j}{\partial w_k}\right) = -P_j \left(\sum_i P_i \frac{\partial V_i}{\partial w_k} - \frac{\partial V_j}{\partial w_k}\right) = -P_j (E(\mu_{*k}) - \mu_{jk})$$

where we have defined

$$\mu_{jk} \equiv \frac{\partial V_j}{\partial w_k} \tag{2.4}$$

and

$$E(\mu_{*k}) \equiv \sum_i P_i \frac{\partial V_i}{\partial w_k} \tag{2.5}$$

Now we substitute (2.3) into (2.1) to obtain:

$$\begin{aligned} \frac{\partial E(z_i^*)}{\partial w_k} &= -\sum_{j=1}^{M} z_{ji} P_j (E(\mu_{*k}) - \mu_{jk}) = \\ &= -E(z_{*i})E(\mu_{*k}) + \sum_{j=1}^{M} P_j z_{ji} \mu_{jk} \text{ (using (1.6))} \\ &= \operatorname{cov}(z_{*i}, \mu_{*k}). \end{aligned} \tag{2.6}$$

Using (2.4), we also have:

$$\frac{\partial E(z_i^*)}{\partial w_k} = \operatorname{cov}(z_{*i}, -\lambda_* z_{*k}) \tag{2.7}$$

where $\lambda_j \equiv \frac{\partial V_j}{\partial (m - \pi_j)}$ = marginal utility of income evaluated at alternative j.

3. A special case: the quasi-linear utility function

It is interesting to consider the case with $\lambda_j = \lambda$ (constant), i.e. a utility function linear in the income term. The linear-in-income specification is very common in empirical analysis adopting the MNL framework. Moreover, even when the utility function is not linear in the income term, if utility is additively separable in z and $(m-\pi)$ and if m is large with respect to π[1], then the marginal utility of income will have little variation across alternatives. Rewriting (2.7) with a constant λ, we get:

$$\frac{\partial E(z_i^*)}{\partial w_k} = -\lambda \operatorname{cov}(z_{*i}, z_{*k}) \tag{3.1}$$

and

$$\frac{\partial E(z_i^*)}{\partial w_i} = -\lambda \operatorname{var}(z_{*i}) \tag{3.2}$$

Focussing on expression (3.2), let us write the variance as

$$\operatorname{var}(z_{*i}) = \sum\nolimits_j P_j z_{ji}^2 - \left(\frac{1}{M} \sum\nolimits_j P_j z_{ji} \right)^2 \tag{3.3}$$

By adding and subtracting $\frac{1}{M} \sum_j z_{ji}^2 - \left(\frac{1}{M} \sum_j z_{ji} \right)^2$ we obtain

$$\begin{aligned} \operatorname{var}(z_{*i}) = {} & \left[\frac{1}{M} \sum\nolimits_j z_{ji}^2 - \left(\frac{1}{M} \sum\nolimits_j z_{ji} \right)^2 \right] + \\ & + \sum\nolimits_j z_{ji}^2 \left(P_j - \frac{1}{M} \right) + \\ & + \left[\left(\sum\nolimits_j \frac{1}{M} z_{ji} \right)^2 - \left(\sum\nolimits_j P_j z_{ji} \right)^2 \right] \end{aligned} \tag{3.4}$$

The first term in square brackets is the "arithmetic variance" of the values of attribute i across the alternatives, or equivalently the variance computed according to a uniform distribution of choice probabilities. The second term is a measure of "non-uniformity" of the choice probabilities, where the addends contribute with a positive or negative sign depending on P_j being larger or smaller than *1/M*. The last term in square brackets is a

---

[1] This will easily be case in example 2 (but not in examples 1 or 3) of section 1.

measure of asymmetry of the distribution of choice probabilities: it is zero if the distribution is symmetric, positive if the distribution is asymmetric to the right, negative if the distribution is asymmetric to the left.

Note that if $P_j = {}^{1}\!/\!_{M}, \forall j,$ the second and third terms disappear. Thus in a "poorly informative" model (i.e. a model with P_j close to ${}^{1}\!/\!_{M}$, $\forall j$) the own price effect of an attribute will be dominated by the arithmetic variance of the values of that attribute across the alternatives. Even in informative models, that variance will have some weight on the own price effect. This seems to have some interesting implications on the specification of the choice set, which are unexplored so far. A common procedure consists in representing continuous choice sets with a (often small) set of discrete values. However, how many values, and which values, are selected will in general affect the arithmetic variance of attributes across the alternatives and therefore in turn affect the own price effects of the attributes. This suggests that when it is adopted the strategy of approximating a continuous (or even a discrete but very large) choice set with a relatively small set of discrete alternatives, some care should be used in building the discrete set so as not to artificially restrict or inflate the variance of the attribute values across the alternatives.

4. Discrete vs Continuous Choice Sets

Expression (2.7) – or in the case of quasi-linear utility – expressions (3.1) and (3.2), carry over to continuous choice sets: simply replace sums with integrals. This remains true however only if the choice density function is non-degenerate. If the variance of the random component in expression (1.5) goes to 0, also the covariances or variances appearing in (2.7), (3.1) and (3.2) go to 0, i.e. the price effects fade out. This makes sense with a discrete choice set. If the model predicts a particular choice with probability 1, then an infinitesimal change in a price will not change the optimal (discrete) choice (if the alternatives are sufficiently far away). However, if the choice set is continuous, as the variance of the random component goes to 0 we simply approach the deterministic case were the optimal choice is the solution to:

$$\max_z U(z, m - \pi(z))$$

and the optimal choice of *z* will be a deterministic function of *w* and *m*. For example, suppose $U = \sum_i \alpha_i \ln(z_i) + \lambda(m - \sum_i w_i z_i)$. Then the optimal (interior) solution is $z_i^* = \alpha_i / \lambda w_i$, and the own price effect is $\partial z_i^* / \partial w_i = -\alpha_i / \lambda w_i^2$.

Therefore, the discrete and the continuous choice set cases seem to diverge when we approach a deterministic model. This again sounds as a *caveat* for the common procedure of approximating continuous choice sets with a discrete set of (fixed) points. If the approximating choice set contains too few alternatives, we risk to force toward 0 the price effects of attributes.[2] It is also worthwhile noting that the problem emerges to the extent that the model is "too good" (overfitting), i.e. the variance of the stochastic component is "too small". In this perspective, the strategy of maximizing the fitting performance of the systematic part $V(z_j, m - \pi_j)$ - e.g. using very general and flexible forms with lots of parameters – might not always be the most appropriate one.

---

[2] Aaberge, Colombino and Wennemo (2009) present a simulation analysis of alternative procedures to generate the choice sets.

References

Aaberge, R., Colombino, U. and S. Strøm, 1999, “Labor Supply in Italy: An Empirical Analysis of Joint Household Decisions, with Taxes and Quantity Constraints”, *Journal of Applied Econometrics*, 14(4), 403-422.

Aaberge, R., Colombino U. and T. Wennemo, 2009, "Evaluating Alternative Representations of the Choice Sets in Models of Labour Supply", *Journal of Economic Surveys*, 23(3), 586-612.

Ben-Akiva, M., and S. Lerman, 1985, *Discrete choice analysis*, MIT Press: Cambridge.

Colombino, U., 1998, “Evaluating the effects of new telephone tariffs on residential users’ demand and welfare. A model for Italy*”, Information Economics and Policy*, 10(3), 283-303.

Colombino, U., 2013, "A new equilibrium simulation procedure with discrete choice models", *International Journal of Microsimulation*, 6(3), 25-49.

McFadden, D., 1978, “Modelling the Choice of Residential Location*”*, in Karlqvist et al. (eds.), *Spatial Interaction Theory And Planning Models*, 75-96, North Holland: Amsterdam.

Train, K., 1980, "A Structured Logit Model of Auto Ownership and Mode Choice", *Review of Economic Studies*, 47(2), 357-370.

Train, K., McFadden, D. and M. Ben-Akiva, 1987, “The demand for local telephone service”, *Rand Journal of Economics*, 18(1), 109-123.

Trajtenberg, M., 1989, "The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners", *Journal of Political Economy*, 97(2), 444-479.

Van Soest, A., 1995, “Structural Models of Family Labor Supply: A Discrete Choice Approach”, *Journal of Human Resources* 30(1), 63-88.