An experimental analysis of the implicit choice set generation using the Constrained Multinomial Logit model

Michel Bierlaire * Ricardo Hurtubia * Gunnar Flötteröd *

4th Kuhmo-Nectar Conference and Summer School

"Transport and Urban Economics"

June 29 - July 3, 2009

Abstract

In this paper, we compare two methods to model the formation of choice sets in the context of discrete choice models. The first method is the probabilistic approach proposed by Manski (1977), who models the choice probability as the joint probability of selecting a choice set and an alternative from this set. This approach is theoretically sound and unbiased, but it is hard to implement due to the complexity that arises from the combinatorial number of possible choice sets. The second method, known as the Constrained Multinomial Logit (CMNL), uses explicit alternative elimination and is easier to implement but can only be understood as an approximation of Manski's approach. We analyze in which situations this approximation is appropriate by estimating models with both approaches over synthetic data and comparing the results.

Keywords: Choice Set Generation, Constrained Multinomial Logit

*Transport and Mobility Laboratory, École Polytechnique Fédérale de Lausanne, {michel.bierlaire, ricardo.hurtubia, gunnar.floetteroed}@epfl.ch

1 Introduction

In standard choice models, it is assumed that the alternatives considered by the decision maker can be exogenously specified by the analyst. The choice set is thus characterized by deterministic rules based on the decision-maker and the choice context. For example, single-room apartments are not considered by families with children in a house choice context, train is not considered as a possible transportation mode if it involves a long walk to reach the train station, etc. There are, however, many situations where the deterministic choice set generation procedure is not satisfactory, or even possible. It may be due to fuzzy rules (how long is a "long walk"?), or unavailability of data (the number of children in the household is unknown to the analyst).

Modeling explicitly the choice set generation process involves a combinatorial complexity, which make the models intractable except for some specific instances (Manski, 1977). Therefore, some heuristics have been proposed in the literature that derive tractable models by approximating the choice set generation process, such as the Implicit Availability/Perception (IAP) model by (Cascetta & Papola, 2001) or the Constrained Multinomial Logit (CMNL, see (Martinez et al., 2009)). The objective of this paper is to empirically analyze the quality of the CMNL model when estimated from synthetic data for which the true choice set formation logic is known.

The paper is organized as follows. In Section 2 and 3, we review the probabilistic choice set approach and the Constrained Multinomial Logit. In Section 4, we compare these two approaches, first through a very simple example and, second, by estimating both models over synthetic data. Section 5 concludes the paper and identifies possible further work.

2 Probabilistic choice set generation

The most commonly used discrete choice model, the Multinomial Logit (MNL), assumes that the decision maker considers a subset of feasible alternatives (the choice set) when facing a choice. The probability of individual n choosing alternative i is

$$P_n(i) = \frac{e^{V_{in}}}{\sum_{j \in C_n} e^{V_{jn}}} \tag{1}$$

where V_{in} is the deterministic utility of alternative i for individual n and C_n is this individual's choice set. The choice set is a subset of the universal choice set C and may vary across individuals. A very usual and convenient way to account for this is the use of deterministic rules to define the availability of an alternative. This is expressed through indicators that describe the availability of an alternative to a decision maker: $A_{in} = 1$ if alternative i is available to individual n, 0 otherwise. Using these indicators, we can write any choice model as

$$\begin{aligned} P_n(i|C_n) &= \Pr\left(U_{in} \geq U_{jn}, \forall j \in C_n\right) \\ &= \Pr\left(U_{in} + \ln A_{in} \geq U_{jn} + \ln A_{jn}, \forall j \in C\right) \end{aligned} \tag{2}$$

where U_{in} is the random utility of alternative i for decision maker n. For an unavailable alternative, this adds $\ln 0 = -\infty$ to its utility, whereas the addition of $\ln 1 = 0$ has no effect on the utility of an available alternative.

In the case of a Multinomial Logit, this generates choice probabilities of the form

$$P_n(i) = \frac{e^{V_{in} + \ln A_{in}}}{\sum_{j \in C} e^{V_{jn} + \ln A_{jn}}}. \tag{3}$$

This probability is equivalent to (1) apart from the fact that (i) the utilities are modified to account for the availability of alternatives and (ii) the sum in the denominator runs over all alternatives in the universal choice set. Usually, the availability indicators are defined from information about the decision maker. A classical example of this is choosing the car mode only if a car is owned or not taking the train if the train station is located too far away. However, in many cases, the availability of an alternative may depend on unobserved attributes or complex interactions between the decision maker and his environment. In these cases, from the point of view of the analyst, the choice set is a latent construct since nothing is observed about it except the chosen alternative. This requires the use of a choice set generation model.

The most general way to account for choice set generation is to define the probability of choosing an alternative i as conditional on the probability of observing the different possible choice sets, as proposed by (Manski, 1977):

$$P_n(i) = \sum_{C_m \subseteq C} P_n(i|C_m) \cdot P_n(C_m) \tag{4}$$

where $P_n(i|C_m)$ is the conditional probability for individual n of choosing alternative i given the choice set C_m and $P_n(C_m)$ is the probability of individual n considering choice set C_m. Since the true choice set (C_n in equation 1) is unknown, we must account for every possible subset C_m that can be built from the universal choice set. This form for the choice probability can be understood as a a two stage approach, where the alternative selection and the choice set generation are two different processes. This approach, which we from now on call the "Probabilistic Choice Set" (PCS) model, is appealing because it is theoretically sound and allows using different models at each stage, but it is hardly applicable to large scale choice problems due to the computational complexity that arises from the combinatorial number of possible choice sets: If the number of alternatives in the universal choice set is J, the number of possible choice sets is $(2^J - 1)$.

Several authors, including (Swait & Ben-Akiva, 1987), (Ben-Akiva & Boccara, 1995), and (Swait, 2001), have addressed the choice set generation process starting from Manski's approach. We analyze them in the following, with an emphasis on the modeling of the choice set probability.

Both (Swait & Ben-Akiva, 1987) and (Ben-Akiva & Boccara, 1995) propose the use of explicit random constraints to determine the choice set generation probability. This methodology defines the probability of considering a choice set as a function of the availability of the different alternatives in the universal choice set:

$$P_n(C_m) = \frac{\prod_{i \in C_m} \phi_{in} \cdot \prod_{j \notin C_m} (1 - \phi_{jn})}{1 - \prod_{k \in C} (1 - \phi_{kn})} \tag{5}$$

where $\phi_{in}$is the probability of alternative i being available to user n. The previous expression assumes independency of the availability probabilities, which are modeled as binary logits that depend on some of the alternative's attributes. As mentioned before, despite the sound theoretical base of this model, there are limitations for its application due to the combinatorial number of possible choice sets.

(Swait, 2001) proposes to model the choice set generation as an implicit part of the choice process in a multivariate extreme value (MEV) framework, requiring no exogenous information. Here, choice sets are not separate constructs but another expression of preferences. The probability of considering a choice set is defined as the probability of that choice set giving the maximum expected utility to an individual n:

$$P_n(C_m) = \frac{e^{\mu I_{n,C_m}}}{\sum_{C_k \subseteq C} e^{\mu I_{n,C_k}}} \tag{6}$$

where μ is the scale parameter for the higher level decision (choice set selection) and I_{n,C_m} is the inclusive value (the "logsum" or expected maximum utility) of choice set C_m for decision maker n:

$$I_{n,C_m} = \frac{1}{\mu_m} \ln \sum_{j \in C_m} \mu_m e^{V_{nj}}. \tag{7}$$

Swait's probabilistic choice set generation approach does not require additional assumptions by the analyst on which attributes affect the alternative's availability, but it is expensive to apply since it also requires the enumeration of all possible sub-sets that can be constructed from the universal choice set. Beyond this, it requires the estimation of a scale parameter μ_m for each one of these subsets.

3 Constrained Multinomial logit

Given the limitations of the probabilistic choice set approach presented in the previous section, some authors have proposed simplified approximations for modeling choice set generation as an implicit part of the choice process. (Cascetta & Pappola, 2001) introduce the Implicit Availability/Perception Logit model as a way to incorporate awareness of paths into route choice modeling without requiring an explicit choice set generation step. The IAP Logit has a similar form as (3) but, instead of using discrete availability indicators, the utility function is shifted using the probability of an alternative being available. A similar approach that penalizes the utilities of "dominated" alternatives is proposed in (Cascetta et al., 2007).

(Martinez et al., 2009) expand the IAP idea and propose the Constrained Multinomial Logit. They assume that the utility is separated in a compensatory and a non-compensatory part, where the compensatory part accounts for the trade-off between the alternative's attributes, while the non-compensatory part indicates the availability of the alternative:

$$\tilde{V}_{in} = V_{in} + \ln \phi_n(i) \tag{8}$$

where V_{in} is the "classical" utility function of alternative i for decision maker n and $\ln \phi_n(i)$ represents the non-compensatory part of the utility. Using this utility function, the expression for the logit choice probability becomes (3) with the availability indicators A_{in} replaced by the cut-off functions $\phi_n(i)$. These cut-off functions can be interpreted as counterparts to the availability probabilities used in the random constraints approach (5).

The functional form for $\phi_n(i)$ is assumed to be a binary logit, considering that the availability of an alternative is related with constrains/thresholds for its attributes:

$$\phi_n(i) = \frac{1}{1 + \exp(\omega(Y_{in} - a))} \tag{9}$$

where Y_{in} is a variable related to the availability of alternative i for decision maker n, the a parameter is the value at which the constraint is most likely to bind, and ω is the scale parameter of the binary logit. Both a and ω are to be estimated. The intuition is that when the attribute Y_{in} exceeds a, the availability $\phi_n(i)$ of alternative i tends to zero, while this availability tends to one when the value of the attribute is below a:

$$\phi_n(i) = \frac{1}{1 + \exp(\omega(Y_{in} - a))} = \begin{cases} 1 \text{ if } Y_{in} - a \to -\infty \\ 0 \text{ if } Y_{in} - a \to +\infty. \end{cases} \tag{10}$$

The previous expression represents an upper value cut-off, where a represents the maximum value that the attribute Y_{in} can have in order to consider alternative i. To model a lower value cut-off we only need to invert the sign of the scale parameter ω.

The term $\ln \phi_n(i)$ can be interpreted as a penalty in the utility function when the constraint is violated since it has no effect when $\phi_n(i) = 1$ ($\ln(1) = 0$) and decreases the utility to minus infinity when $\phi_n(i) = 0$ ($\ln(0) = -\infty$). Another way to interpret this function is as the probability of an alternative belonging to the true choice set that decision maker n considers when making his choice.

The cut-off function can be generalized to account for more than one constraint using the following expression:

$$\tilde{\phi}_n(i) = \prod_k \phi_{nk}(i) \tag{11}$$

where ϕ_{nk} represents a constraint function related with the kth attribute of the alternative.

The CMNL approach has advantages over the PCS model from a practical point of view since it does not require enumerating the choice sets, which makes it easier to specify and estimate. However, the CMNL is a heuristic that is based on assumptions about the functional form of the utility function. The CMNL can thus be understood as an approximation to the PCS model. The next section evaluate the quality of this approximation.

4 Comparison of CMNL with PCS

This section compares the CMNL model with the PCS model. For this, we first present a simple example where we analytically analyze the difference between the choice probabilities obtained using both models. Second, we estimated CMNL and PCS models over synthetic data and compare the results. For notational simplicity, we subsequently omit the index n for the decision-maker.

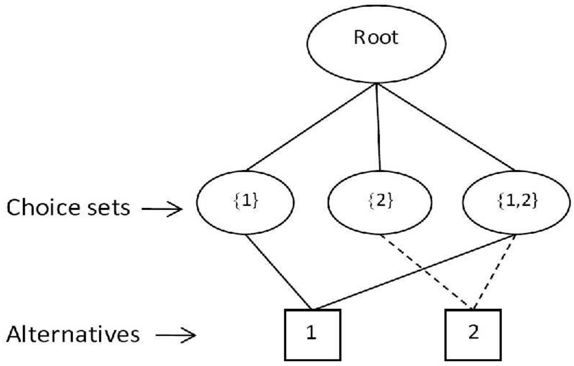


Figure 1: Example of PCS

4.1 Simple example

Consider a multinomial logit model with only 2 alternatives, where alternative 1 is always available ($\phi(1) = 1$) and alternative 2 has a certain probability $\phi(2)$ of being considered as available by the decision maker. Figure 1 shows the structure of the decision tree if we consider every possible combination of alternatives as a choice set. This can be seen as a situation where the decision maker is to some extent captive to alternative 1, see also the captivity logit model proposed in (Gaudry & Dagenais, 1979).

Under the CMNL assumption, the probability of choosing alternative 1 is easily specified as

$$P(1) = \frac{e^{V_1}}{e^{V_1} + e^{V_1 + \ln \phi(2)}}. \tag{12}$$

Under the PCS assumption (equation 4), the probability of choosing alternative 1 becomes

$$P\,(1) = P(\{1\}) \cdot \frac{e^{V_1}}{e^{V_1}} + P(\{1,2\}) \cdot \frac{e^{V_1}}{e^{V_1} + e^{V_2}} \tag{13}$$

where $P(\{1\})$ is the probability of considering the choice set composed only of alternative 1 and $P(\{1,2\})$ is the probability of considering the choice set including both alternatives. According to (5), the choice set probabilities are

$$P(\{1\}) = \frac{\phi(1) \cdot (1 - \phi(2))}{1 - (1 - \phi(1)) \cdot (1 - \phi(2))} = 1 - \phi(2) \tag{14}$$

and

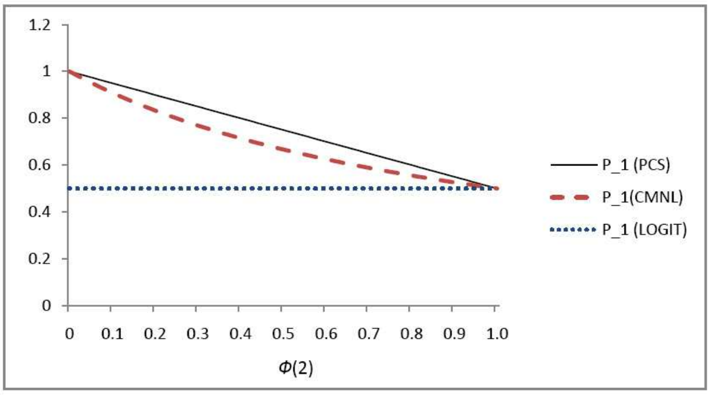


Figure 2: Choice probability of alternative 1 ($V_1 = V_2$)

$$P(\{1,2\}) \quad = \quad \frac{\phi(1)\cdot\phi(2)}{1-(1-\phi(1))\cdot(1-\phi(2))} = \phi(2). \tag{15}$$

The probability of considering choice set $\{2\}$ is zero because alternative 1 is always be available.

In the deterministic limit ($\phi(2) = 0$ or $\phi(2) = 1$), both approaches collapse into (3) and yield the same results. However, it is interesting to see what happens when $\phi(2)$ takes values between zero and one. The resulting choice probabilities are shown in Figure 2, assuming the same utility level $V_1 = V_2$ for both alternatives. Results for a plain multinomial logit model assuming only the full choice set are included for comparison. This figure shows that the CMNL is a good approximation of the PCS only when $\phi(2)$ becomes either zero or one, but it underestimates the probability of alternative 1 elsewhere. If the utility for alternative 1 is bigger than the utility for alternative 2, the approximation improves, which can be seen in Figure 3. However, if the utility of alternative 1 is smaller than the utility of alternative 2, the CMNL becomes a poor approximation of the PCS for intermediate $\phi(2)$ values, which is demonstrated in Figure 4.

These results can be interpreted as an unwanted compensatory effect in the CMNL: The constraint is enforced by modifying the utility of the constrained alternative.

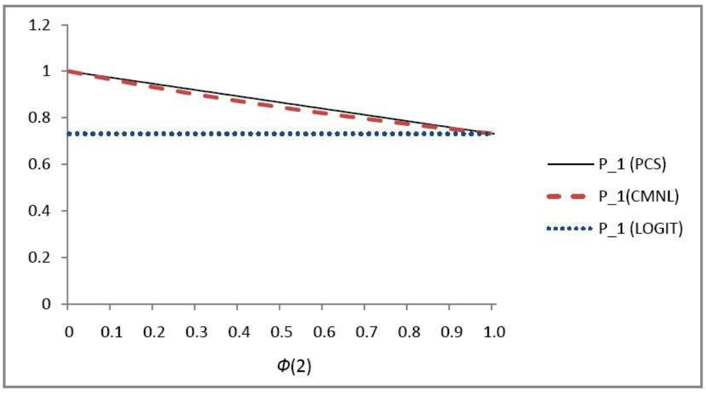


Figure 3: Probability of alternative 1 ($V_1 > V_2$)

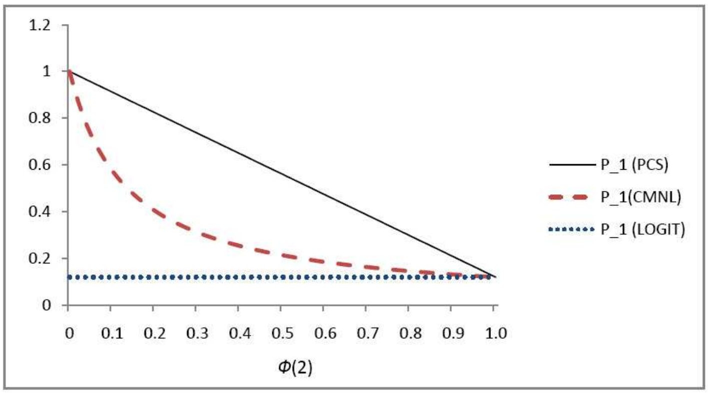


Figure 4: Choice probability of alternative 1 ($V_1 < V_2$)

However, as the utility of this alternative gets higher, it eventually does compensate the constraint function to some extent. This explains that for $V_1 > V_2$, where alternative 2 is relatively unattractive anyway, the CMNL provides a good approximation to the PCS, whereas for $V_1 < V_2$, where the attractiveness of alternative 2 works against its constraint, the deviation from the PCS becomes large.

This indicates at least two situations in which the CMNL is a good approximation of the PCS: (i) the constraint functions are steep enough to approach a deterministic limit, and (ii) a positive correlation between the availability of an alternative and its utility level exists. The performance of the CMNL when applied to a data set with rather soft constraints is investigated in the next section.

4.2 Synthetic data

This section describes a series of controlled experiments where some of the data is synthetically generated. We start from a real stated preference data set that was collected for the analysis of a hypothetical high speed train in Switzerland (Bierlaire et al., 2001). The alternatives are:

1. Driving a car (CAR)
2. Regular Train (TRAIN)
3. Swissmetro, the future high speed train (SM)

From this data set, which consists of 5607 observations, we use only the attributes of the alternatives to simulate synthetic choices based on pre-specified utility functions and constraints.

Regarding the constraints, it is assumed that the TRAIN and the SM alternative are always available, whereas the availability of the CAR alternative depends on the travel time according to

$$\phi(\text{CAR}) = \frac{1}{1 + \exp(\omega(TT_{\text{CAR}}/60 - a))}, \tag{16}$$

which states that the probability of considering CAR as an available alternative decreases with the travel time TT_{CAR}, in minutes, and that this probability is 0.5 when the availability threshold a, in hours, is reached.

This implies that, depending on the availability of the CAR alternative, there are two possible choice sets: the full choice set and the choice set containing only the TRAIN and the SM alternative. The random constraints approach (Ben-Akiva & Boccara, 1995) defines the probability of each choice set as follows:

$$\begin{aligned} P(\text{TRAIN, SM}) &= \frac{\phi(\text{TRAIN})\phi(\text{SM})(1-\phi(\text{CAR}))}{1-(1-\phi(\text{CAR}))(1-\phi(\text{TRAIN}))(1-\phi(\text{SM}))} \\ &= 1-\phi(\text{CAR}) \end{aligned} \tag{17}$$

and, accordingly,

$$P(\text{CAR, TRAIN, SM}) = \phi(\text{CAR}). \tag{18}$$

The synthetic choices are simulated by (i) sampling a choice set for each decision maker and (ii) sampling an alternative for each decision maker from his choice set using a multinomial logit model with the following specification for the utility functions:

$$\begin{aligned} V_n(\text{CAR}) &= \text{ASC}_{\text{CAR}}+\beta_{\text{cost}} \cdot \text{COST}_{\text{CAR}} + \beta_{\text{tt}} \cdot \text{TT}_{\text{CAR}} \\ V_n(\text{TRAIN}) &= \beta_{\text{cost}} \cdot \text{COST}_{\text{TRAIN}} + \beta_{\text{tt}} \cdot \text{TT}_{\text{TRAIN}}+\beta_{\text{he}} \cdot \text{HE}_{\text{TRAIN}} \\ V_n(\text{SM}) &= \text{ASC}_{\text{SM}}+\beta_{\text{cost}} \cdot \text{COST}_{\text{SM}} + \beta_{\text{tt}} \cdot \text{TT}_{\text{SM}} + \beta_{\text{he}} \cdot \text{HE}_{\text{SM}} \end{aligned} \tag{19}$$

A description of the parameters and their values used for the choice sampling procedure are given in Table 1. 100 choice data sets were generated for each value of ω. These values generate constraints with different levels of dispersion, from very wide constraints ($\omega = 1$) to very steep constraints ($\omega = 10$). Figure 5 shows the shape of some of these constraint functions. Estimation results for both the PCS and the CMNL model are given in Tables 2 and 3. The presented values are averaged over all 100 experiments per ω value.

Table 1: Parameter descriptions and values

Parameter	Description	Value	Unit
ASC_{CAR}	Alternative specific constant for car	0.3	–
ASC_{SM}	Alternative specific constant for Swissmetro	0.4	–
β_{cost}	Cost parameter	-0.001	CHF
β_{tt}	In vehicle travel time parameter	-0.001	Minutes
β_{he}	Headway (betw. successive trains/metros)	-0.005	Minutes
a	Availability threshold	3	Hours
ω	Availability dispersion	1,2,3,5,10	–

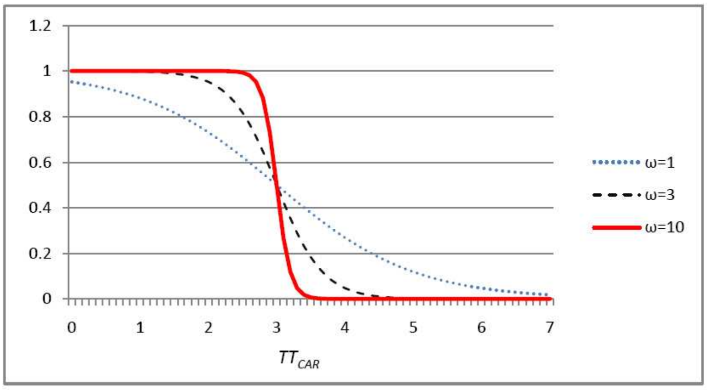


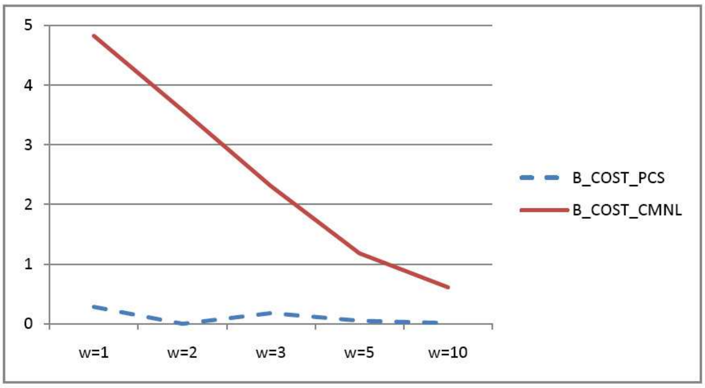
Figure 5: Shape of the constraint for different values of ω

Table 2: Estimation results for PCS model

real ω value		1		2		3		5		10	
parameter	real value	estimate	t-test	estimate	t-test	estimate	t-test	estimate	t-test	estimate	t-test
ASC_{CAR}	0.3	0.304	0.027	0.288	0.113	0.300	0.010	0.301	0.012	0.314	0.184
ASC_{SM}	0.4	0.396	0.044	0.399	0.010	0.405	0.053	0.401	0.017	0.410	0.151
β_{cost}	-0.01	-0.010	0.283	-0.010	0.001	-0.010	0.179	-0.010	0.052	-0.010	0.012
β_{he}	-0.005	-0.005	0.241	-0.005	0.010	-0.005	0.048	-0.005	0.082	-0.005	0.078
β_{time}	-0.01	-0.01	0.074	-0.010	0.050	-0.010	0.049	-0.010	0.003	-0.010	0.001
a	3	2.963	0.019	3.008	0.118	3.000	0.100	2.998	0.081	3.002	0.101
ω	see top	1.003	0.028	2.014	0.079	3.066	0.210	5.095	0.170	10.523	0.353

Table 3: Estimation results for CMNL model

real ω value		1		2		3		5		10	
parameter	real value	estimate	t-test	estimate	t-test	estimate	t-test	estimate	t-test	estimate	t-test
ASC_{CAR}	0.3	0.503	0.950	0.421	1.153	0.406	1.365	0.380	0.988	0.326	0.313
ASC_{SM}	0.4	0.565	2.013 *	0.550	2.375 *	0.536	1.804	0.506	1.485	0.463	0.872
β_{cost}	-0.01	-0.008	4.825 *	-0.008	3.580 *	-0.009	2.309 *	-0.009	1.182	-0.010	0.613
β_{he}	-0.005	-0.005	0.202	-0.005	0.151	-0.005	0.071	-0.005	0.120	-0.005	0.090
β_{time}	-0.01	-0.007	3.929 *	-0.008	3.645 *	-0.008	2.813 *	-0.009	2.316 *	-0.009	1.523
a	3	2.186	1.753	2.656	3.073 *	2.773	3.762 *	-2.869	3.305 *	2.948	1.864
ω	see top	1.043	0.239	2.094	0.403	3.118	0.431	5.238	0.424	12.146	3.149 *



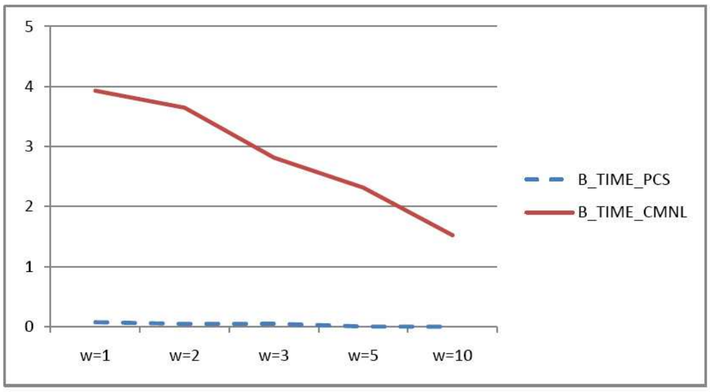


Figure 6: T-statistics for the cost and time parameter over ω

As expected, the results for the PCS model are unbiased and all estimates are significant with 95% probability when tested against the real values. The estimates for the CMNL are also close to the real values, but many of them (marked with *) are not significant when tested against the real values. The quality of the CMNL estimates improves with decreasing dispersion (increasing ω). This is consistent with the findings of Section 4.1.

Figure 6 shows the t-statistics for the cost and travel time parameter over different ω values for the PCS model and the CMNL model. The quality of the estimates is constant across different values of ω for the PCS model. The quality of the CMNL estimates increases with ω, and their t-statistics reach acceptable values when the constraint functions become very steep. The frequent occurrence of insignificant

CMNL parameters for smaller ω values results from the model's inadequacy in these conditions: If the functional form of the CMNL is inconsistent with that of the PCS model based on which the synthetic data is generated, then the CMNL parameters become weak in explaining the data, which translates in estimation results of low significance.

5 Conclusions and further work

The CMNL is a proper approximation of the PCS only when the constraints (or the availabilities of alternatives) tend to be deterministic. This reduces the spectrum of possible applications of the CMNL but, at the same time, confirms that it is convenient to use, especially when solving problems with numerous alternatives where the behavior is expected to tend to a deterministic alternative elimination process.

In further work, specific ways to determine when it is recommendable to use the CMNL will be investigated, together with exploring a possible correction or modification to the CMNL specification in order to make it a reasonable approximation to the PCS approach even in scenarios with large dispersion in the alternative availabilities. Appendix A provides some additional insight into this problem. It demonstrates that the PCS also collapses into the CMNL if all possible choice sets have the same logsum value. An interesting question is to what extent this type of result can be used to evaluate *ex post* if the CMNL is an appropriate model specification for a given data set.

A Equivalence of PCS and CMNL for constant logsums

Equating the choice probabilities for PCS and CMNL and omitting the subscript n, we have, for all i in C,

$$\sum_{C_m \subseteq C} \frac{\mathbf{1}(i \in C_m) e^{V_i}}{\sum_{j \in C_m} e^{V_j}} P(C_m) = \frac{\phi(i) e^{V_i}}{\sum_{j \in C} \phi(j) e^{V_j}} \tag{20}$$

and, equivalently,

$$\sum_{C_m \subseteq C} \frac{\mathbf{1}(i \in C_m)}{\sum_{j \in C_m} e^{V_j}} P(C_m) = \frac{\phi(i)}{\sum_{j \in C} \phi(j) e^{V_j}} \tag{21}$$

where $\mathbf{1}(\cdot)$ is the indicator function. We assume that

$$\sum_{j \in C_m} e^{V_j} = I \tag{22}$$

for all C_m with $P(C_m) > 0$. Then,

$$\frac{\sum_{C_m \subseteq C} \mathbf{1}(i \in C_m) P(C_m)}{I} = \frac{\phi(i)}{\sum_{j \in C} \phi(j) e^{V_j}}. \tag{23}$$

To begin with, we equate only the numerators

$$\sum_{C_m \subseteq C} \mathbf{1}(i \in C_m) P(C_m) = \phi(i), \tag{24}$$

which specifies $\phi(i)$ as the marginal probability that alternative i is contained in any considered choice set. This is consistent with the definition given in Section 3.

Equating the denominators, substituting (24), and rearranging terms then leads to

$$
\begin{aligned}
I &= \sum_{j \in C} \phi(j) e^{V_j} \\
&= \sum_{j \in C} \sum_{C_m \subseteq C} \mathbf{1}(j \in C_m) P(C_m) e^{V_j} \\
&= \sum_{C_m \subseteq C} P(C_m) \sum_{j \in C_m} e^{V_j}, \qquad (25)
\end{aligned}
$$

where the right-hand-side is nothing but the expectation of the constant I value defined in (22). Consequently, specification (24) equates the PCS and the CMNL for constant logsum terms. This also implies that the CMNL model is a reasonable approximation of the PCS model in situations where all choice sets have approximately the same logsums. Operationally, this poses the question of how to identify such situations.

References

M. E. Ben-Akiva & B. Boccara (1995). 'Discrete choice models with latent choice sets'. *International Journal of Research in Marketing* **12**:9–24.

M. Bierlaire, et al. (2001). 'Acceptance of modal innovation: the case of the Swiss-Metro'. In *Proceedings of the 1st Swiss Transportation Research Conference*, Ascona, Switzerland. www.strc.ch.

E. Cascetta, et al. (2007). 'The use of dominance variables in choice set generation'. In *Proceedings of the 11th World Conference on Transport Research*, University of California at Berkeley.

E. Cascetta & A. Papola (2001). 'Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand'. *Transportation Research Part C* **9**:249–263.

M. Gaudry & M. Dagenais (1979). 'The dogit model'. *Transportation Research Part B* **13**:105–111.

C. Manski (1977). 'The structure of random utility models'. *Theory and Decision* **8**:229–254.

F. Martinez, et al. (2009). 'The Constrained Multinomial Logit Model: A Semi-Compensatory Choice Model'. *Transportation Research Part B* **43**:365–377.

J. Swait (2001). 'A non-compensatory choice model incorporating attribute cut-offs'. *Transportation Research Part B* **35**:903–928.

J. Swait & M. Ben-Akiva (1987). 'Incorporating random constraints in discrete models of choice set generation'. *Transportation Research Part B* **21**(2).