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

Michel Bierlaire, Ricardo Hurtubia and Gunnar Flötterröd

Transport and Mobility Laboratory

Introduction

Choice model:

$$P_n(i|\mathcal{C}_n)$$

- Common practice: $\mathcal{C}_n$ characterized by deterministic rules
- Modeling the choice set generation (Manski, 1977):

$$P_n(i) = \sum_{\mathcal{C}_m \subseteq \mathcal{C}} P_n(i|\mathcal{C}_m) P_n(\mathcal{C}_m)$$

- Combinatorial complexity
- Operational instances:
  - Random constraints (Swait and Ben-Akiva, 1987, Ben-Akiva and Boccara, 1995)
  - MEV framework (Swait, 2001)

Introduction

- Heuristics:
  - Implicit Availability/Perception model (Cascetta and Papola, 2001)
  - Constrained Multinomial Logit model (Martinez et al., 2009)

Objective: analyze the quality of the CMNL as a choice set generation process.

Deterministic Choice Set Generation

- Availability conditions
- Exogenous variables

$$A_{in} = \begin{cases} 1 & \text{if alternative } i \text{ is considered by individual } n, \\ 0 & \text{otherwise.} \end{cases}$$

- Choice model

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

- Note: a choice model with deterministic choice set generation can always be written in terms of the universal choice set

Deterministic Choice Set Generation

Logit model:

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

What if variables A_{in} are not exogenously given?

Probabilistic Choice Set Generation

Approaches:

- Correct model: Manski (1977) most of the time impractical
- Sampling of alternatives:
  - Assume $\mathcal{C}_n = \mathcal{C}, \forall n$
  - Sample a subset for estimation
  - see Frejinger, Bierlaire and Ben-Akiva (forthcoming) for route choice
- Replace A_{in} by a probability distribution
  - Availability/Perception (Cascetta and Papola, 2001)
  - Cutoffs (Martinez et al., 2009)

Cutoffs

Optimization problem of rational consumer n:

$$\max_{\delta_{ni}} \sum_{i \in \mathcal{C}} \delta_{ni} U_{in}(X_i)$$

subject to

$$\sum_{i \in \mathcal{C}} \delta_{ni} = 1, \quad \delta_{ni} \in \{0, 1\}, \forall i \in \mathcal{C}$$

But attributes are meaningful only within some bounds

$$\ell_{nk} \leq X_{ik} \leq u_{nk} \; \forall i \in \mathcal{C}, \forall k$$

An alternative i with one of its attributes is out of bounds is not considered

Cutoffs

Examples:

- Item too expensive
- Traveling by train involves a too long walking distance to the station
- etc.

If these rules are deterministic, the variables A_{in} can be derived
If not, what can be done?

Cutoffs

Idea: relax the constraint in a probabilistic way
Example: constraint $\ell \leq X$

$$\begin{array}{rcccc} V_{\text{not considered}} & = & \ell & + & \varepsilon_1 \\ V_{\text{considered}} & = & X & + & \varepsilon_2 \end{array}$$

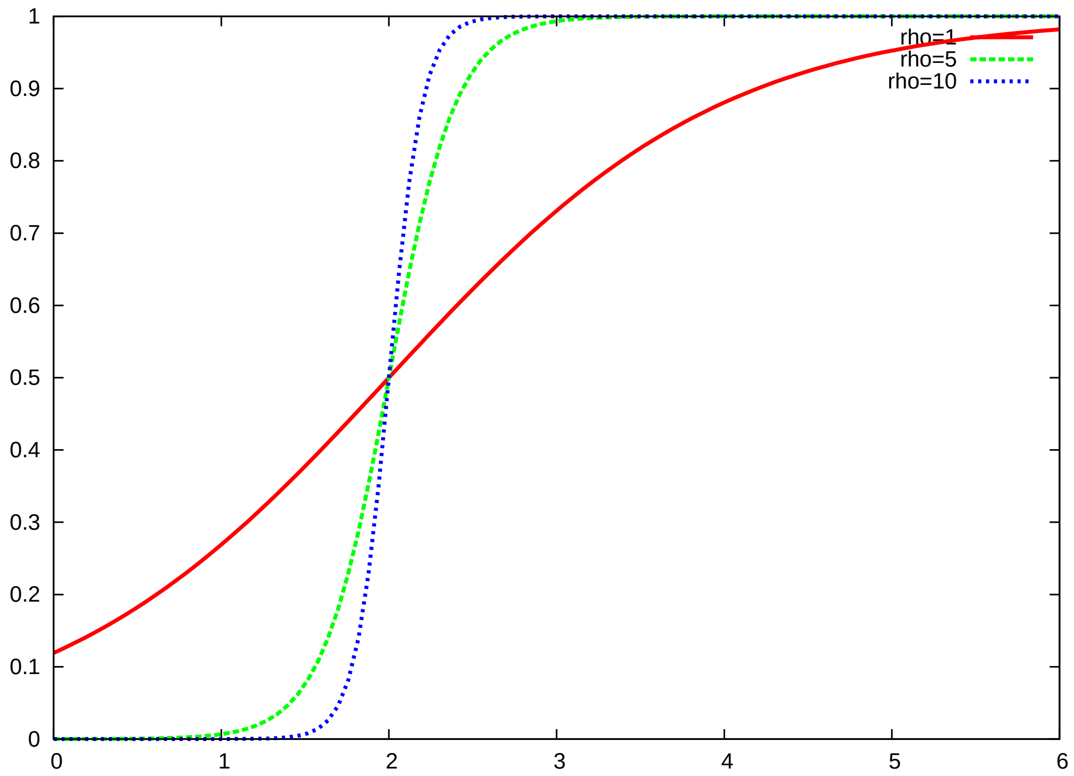
$$P(\text{considered}) = \frac{e^{\rho X}}{e^{\rho X} + e^{\rho \ell}} = \frac{1}{1 + e^{\rho(\ell - X)}}$$

Example: constraint $X \leq u$

$$P(\text{considered}) = \frac{e^{-\rho X}}{e^{-\rho X} + e^{-\rho u}} = \frac{1}{1 + e^{\rho(X - u)}}$$

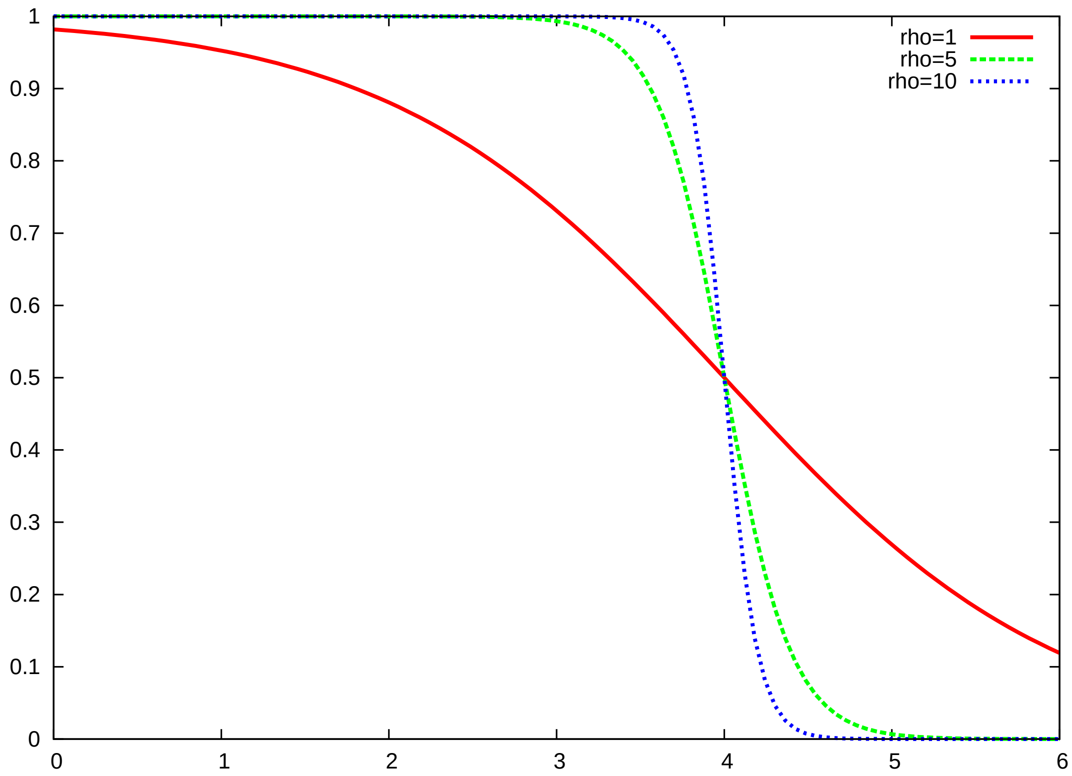
Cutoffs

Example: $2 \leq X$



Cutoffs

Example: $X \leq 4$



Cutoffs

Constraint $\ell \leq X \leq u$

$$P(\text{considered}) = \frac{1}{1+e^{\rho(\ell - X)}} \frac{1}{1+e^{\rho(X-u)}}$$

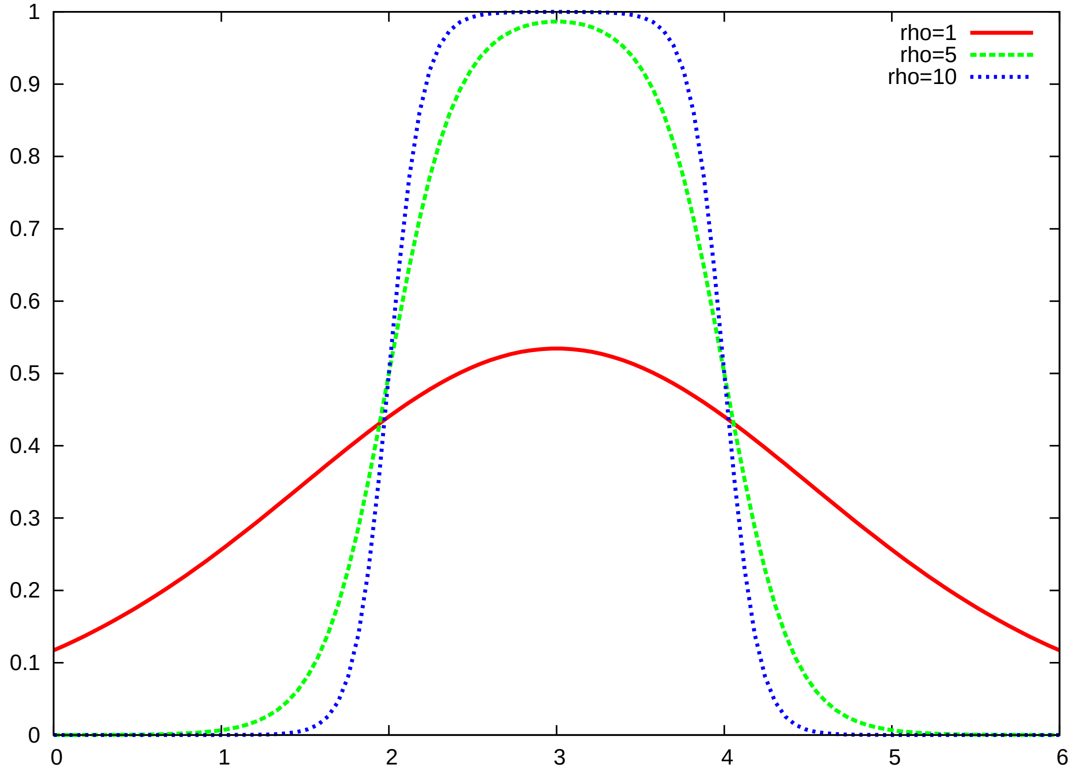
We denote this quantity by $\phi_n(X)$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Cutoffs

Example: $2 \leq X \leq 4$



Cutoffs

The utility function now becomes

$$V_i = \sum_k \beta_k X_{ik} + \sum_{k^*} \frac{1}{\rho} \ln \phi_n(X_{ik^*})$$

where k^* ranges only on constrained attributes. Note that

$$\begin{aligned} \ln \phi(X) &= -\ln(1 + e^{\rho(\ell - X)}) - \ln(1 + e^{\rho(X-u)}) \\ &= -\ln(1 + e^{\rho\ell} e^{-\rho X}) - \ln(1 + e^{\rho X} e^{-\rho u}) \end{aligned}$$

Can be estimated, although it is difficult

Comparison of CMNL and Manski

Simple example:

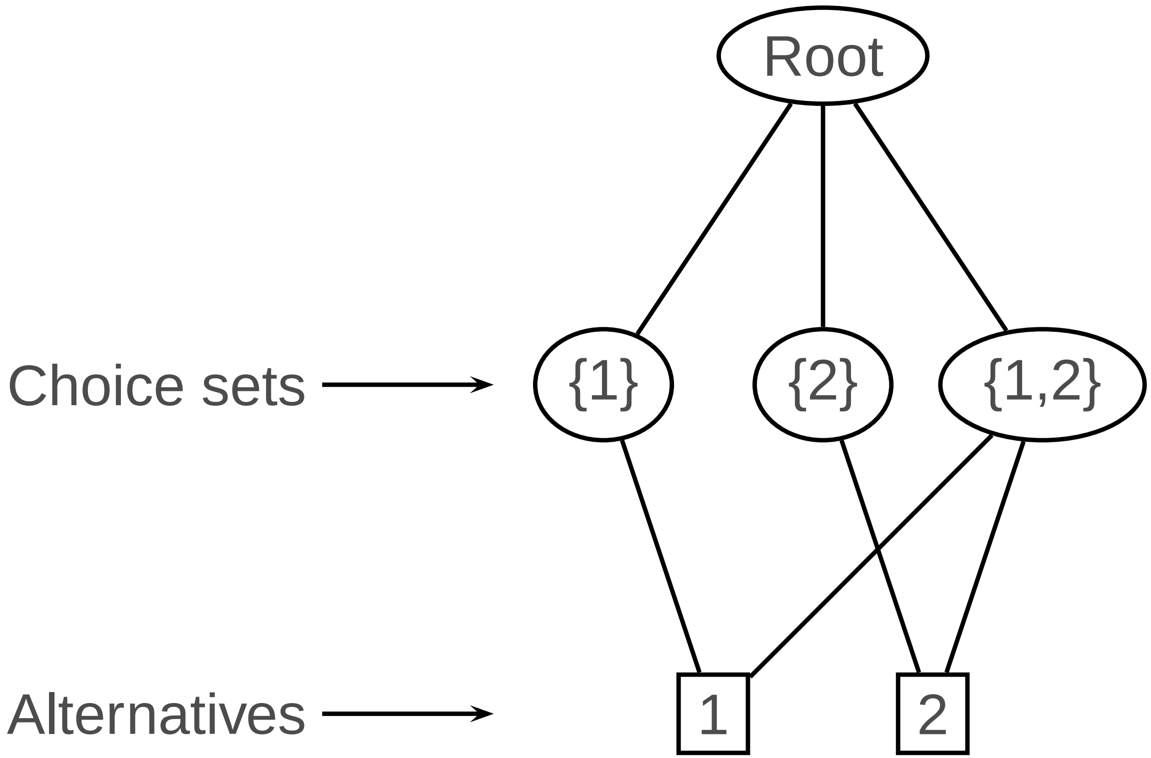
- Binary logit: $\mathcal{C} = \{1, 2\}$
- Alternative 1 is always available
- Alternative 2 is considered with probability ϕ_2

We have

- $P(\mathcal{C}_n = \{1\}) = 1 - \phi_2$
- $P(\mathcal{C}_n = \{2\}) = 0$
- $P(\mathcal{C}_n = \{1, 2\}) = \phi_2$

Comparison of CMNL and Manski

Manski's model



Comparison of CMNL and Manski

Manski's model

$$P(1) = P(\mathcal{C}_n = \{1\}) \frac{e^{V_1}}{e^{V_1}} + P(\mathcal{C}_n = \{2\}) 0 + P(\mathcal{C}_n = \{1,2\}) \frac{e^{V_1}}{e^{V_1} + e^{V_2}}$$

$$= (1 - \phi_2) + \phi_2 \frac{e^{V_1}}{e^{V_1} + e^{V_2}}$$

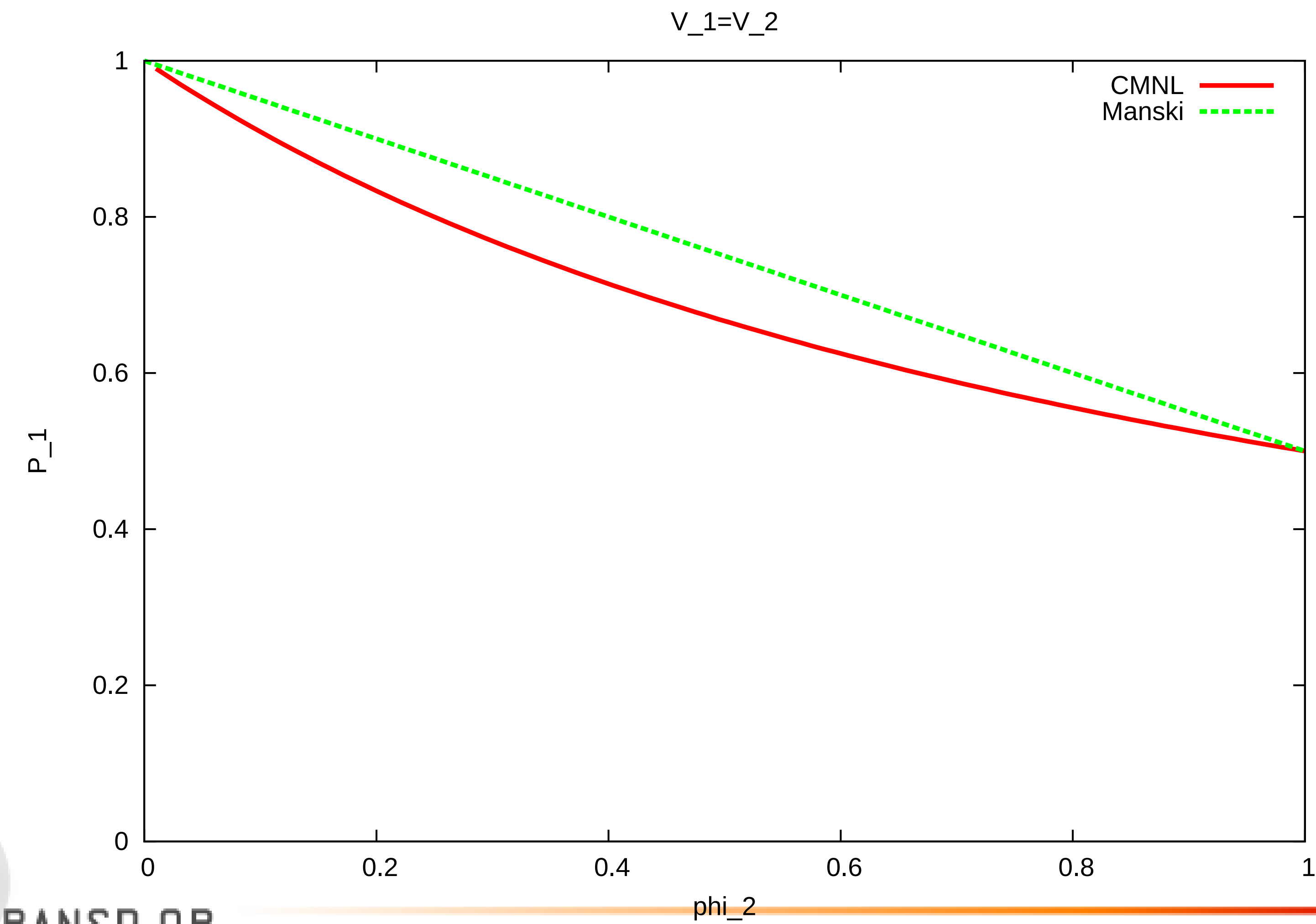
CMNL model

$$P(1) = \frac{e^{V_1}}{e^{V_1} + e^{V_2 + \ln \phi_2}}.$$

Note: for given V's, Manski is linear in ϕ_2, not CMNL

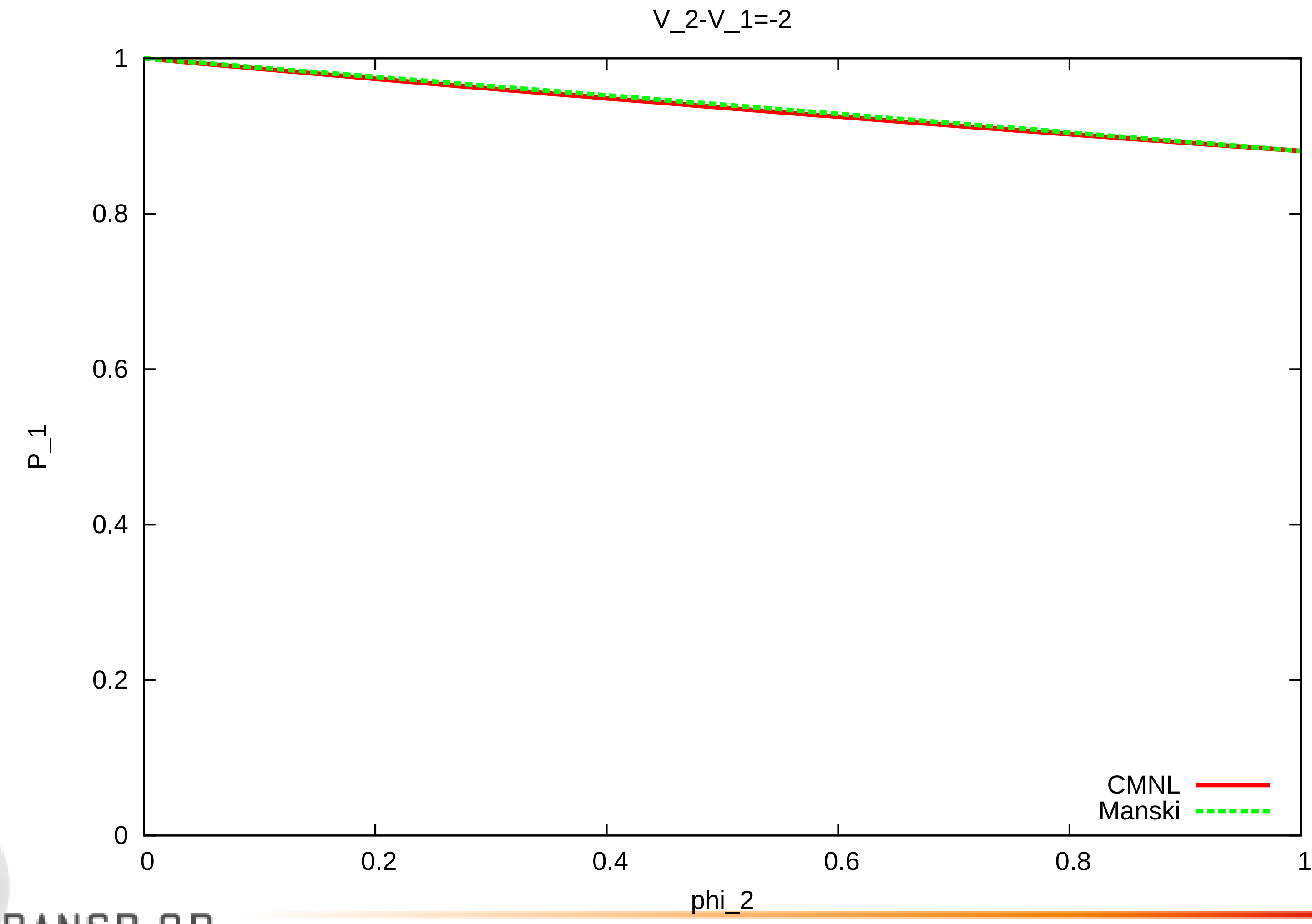
Comparison of CMNL and Manski

Equal utility



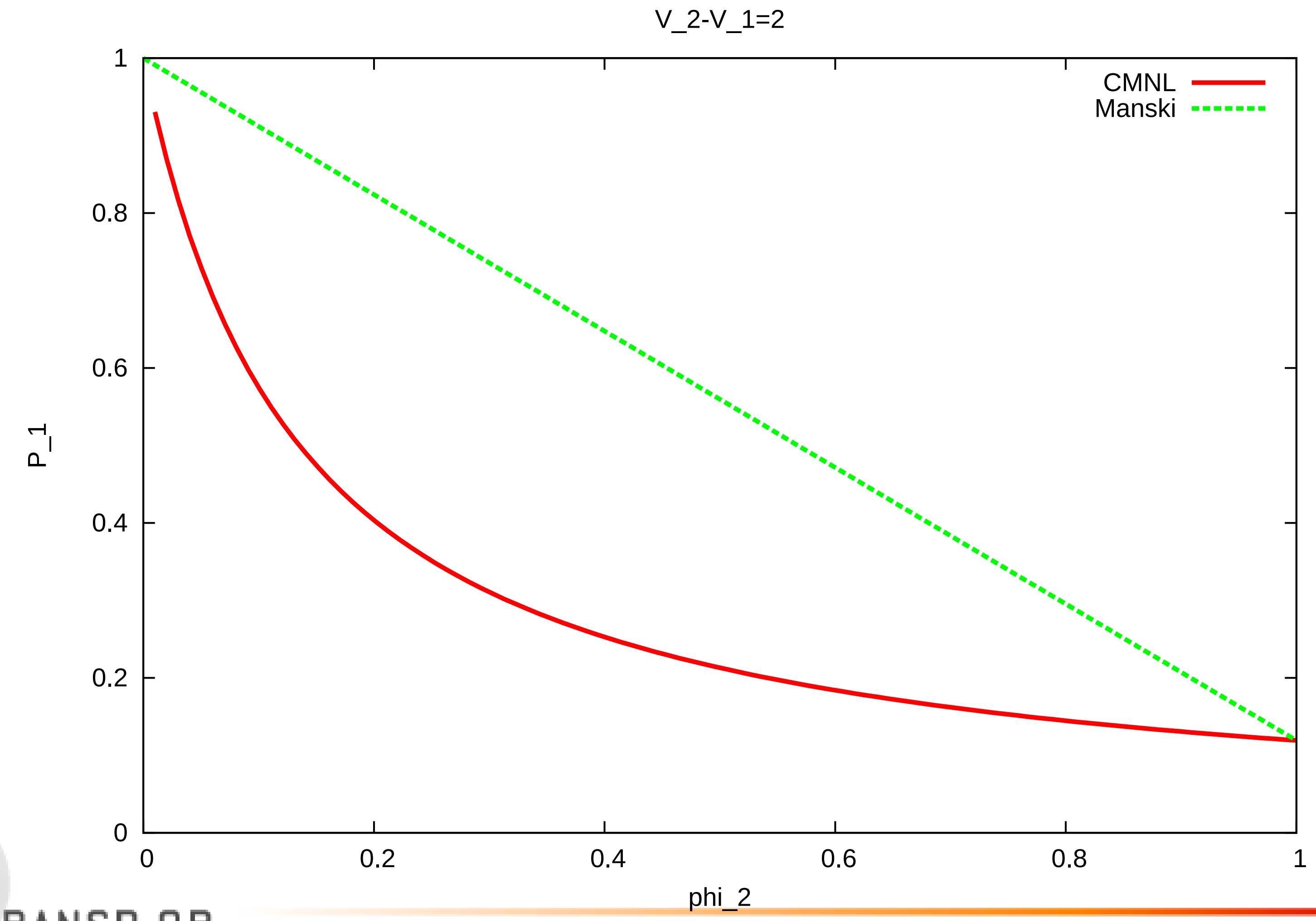
Comparison of CMNL and Manski

Alt. 1 is dominant



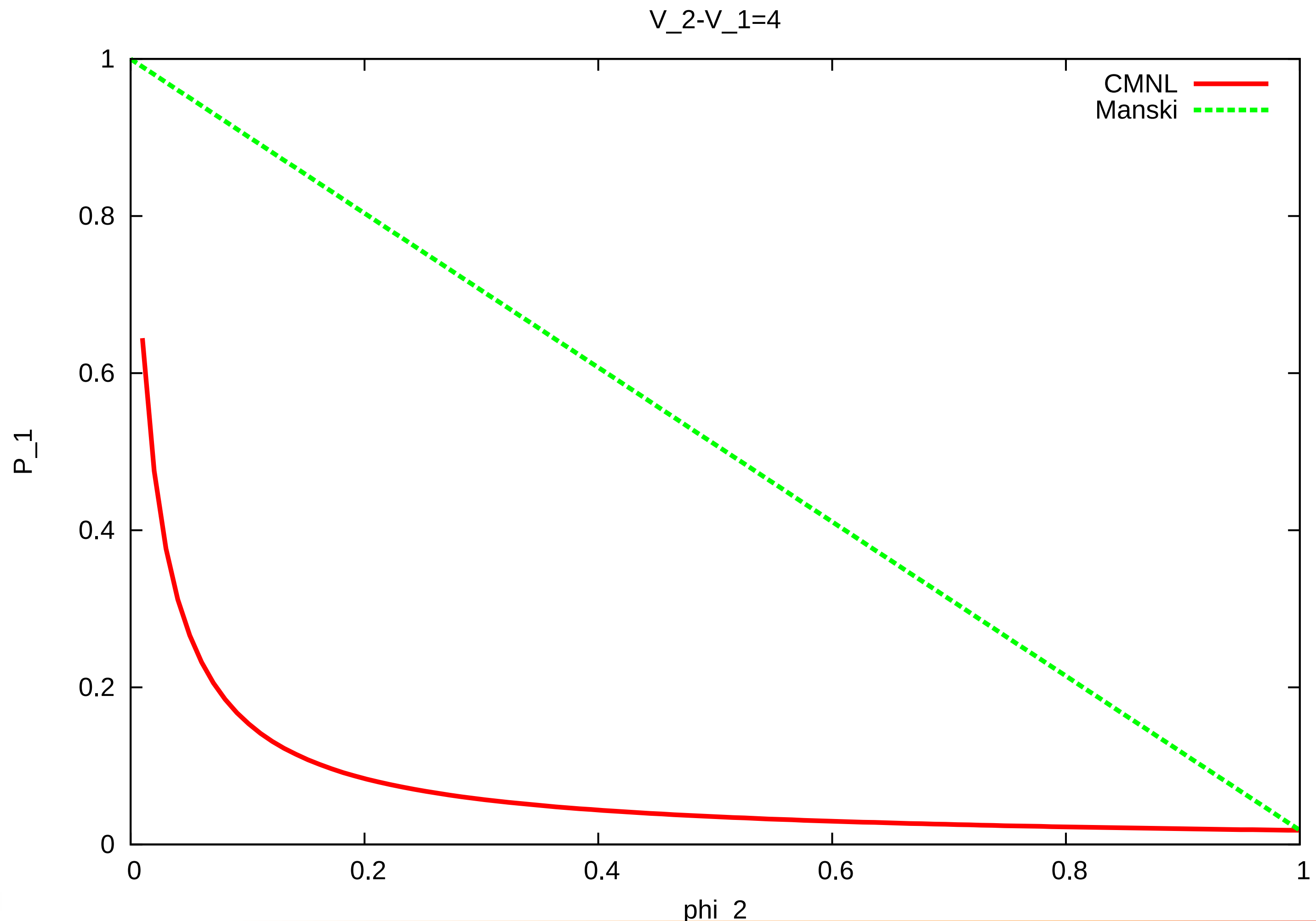
Comparison of CMNL and Manski

Alt. 2 is dominant



Comparison of CMNL and Manski

Alt. 2 is even more dominant



Comparison of CMNL and Manski

- CMNL underestimates the choice probability for alternative 1
- When alt. 1 is dominant, it makes no difference if it is preferred because of a high utility, or if because 2 is not even considered.
- When alt. 2 is dominant, the CMNL may be completely off
- Clearly, the model parameters could be adjusted to attenuate that error

Synthetic data

- Swissmetro data set, 5607 observations
  1. Driving a car (CAR)
  2. Regular train (TRAIN)
  3. Swissmetro, the future high speed train (SM)
- Exogenous variables come from the data set
- Synthetic choice set
  - TRAIN and SM always available
  - CAR available depending on travel time

$$\phi_{\text{CAR}} = \frac{1}{1 + \exp(\omega(TT_{\text{CAR}}/60 - a))}$$

- Synthetic choice

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Synthetic data

Postulated model

Parameter	Value	Car	Train	Swissmetro
$\mathrm{ASC}_{\text{CAR}}$	0.3	1	0	0
ASC_{SM}	0.4	0	0	1
β_{cost}	-0.001	Cost (CHF)	Cost (CHF)	Cost (CHF)
β_{tt}	-0.001	In veh. travel time (minutes)	In veh. travel time (minutes)	In veh. travel time (minutes)
β_{he}	-0.005	0	Headway (minutes)	Headway (minutes)
a	3	Consideration threshold of car (hours)		
ω	1,2,3,5,10	Consideration dispersion of car		

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Synthetic data

- 100 choice data sets are simulated for each value of ω
- Results:
  - mean of each parameter over 100 estimations
  - t-test against the true value, based on the empirical std. deviation.

Estimation results for Manski's model

real ω value		1		2		3		5	
parameter	real value	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
ASC_{SM}	0.4	0.396	0.044	0.399	0.010	0.405	0.053	0.401	0.017
β_{cost}	-0.01	-0.010	0.283	-0.010	0.001	-0.010	0.179	-0.010	0.052
β_{he}	-0.005	-0.005	0.241	-0.005	0.010	-0.005	0.048	-0.005	0.082
β_{time}	-0.01	-0.01	0.074	-0.010	0.050	-0.010	0.049	-0.010	0.003
a	3	2.963	0.019	3.008	0.118	3.000	0.100	2.998	0.081
ω	see top	1.003	0.028	2.014	0.079	3.066	0.210	5.095	0.170

Estimation results for CMNL model

real ω value		1		2		3		5	
parameter	real value	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
ASC_{SM}	0.4	0.565	2.013 *	0.550	2.375 *	0.536	1.804	0.506	1.485
β_{cost}	-0.01	-0.008	4.825 *	-0.008	3.580 *	-0.009	2.309 *	-0.009	1.182
β_{he}	-0.005	-0.005	0.202	-0.005	0.151	-0.005	0.071	-0.005	0.120
β_{time}	-0.01	-0.007	3.929 *	-0.008	3.645 *	-0.008	2.813 *	-0.009	2.316 *
a	3	2.186	1.753	2.656	3.073 *	2.773	3.762 *	-2.869	3.305 *
ω	see top	1.043	0.239	2.094	0.403	3.118	0.431	5.238	0.424

(* indicates an insignificant parameter)

Synthetic data

- Manski model performs well, as expected
- CMNL may significantly bias the estimates
- The more deterministic the constraint, the better the CMNL

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Conclusion

- CMNL is not adequate to model the choice set generation
- It is a model on its own, derived from semi-compensatory arguments
- Its complexity is linear in the number of alternatives, while Manski's model is exponential.
- Research question: how can we modify the CMNL to be a better approximation of Manski's model.