A multivariate nonlinear analysis of tourism expenditures

Enrico Foscolo

Free University of Bozen-Bolzano
School of Economics and Management

December 6th, 2013

Acknowledgments

Based on a joint work with **M. Disegna** and **F. Durante** (Free University of Bozen–Bolzano)

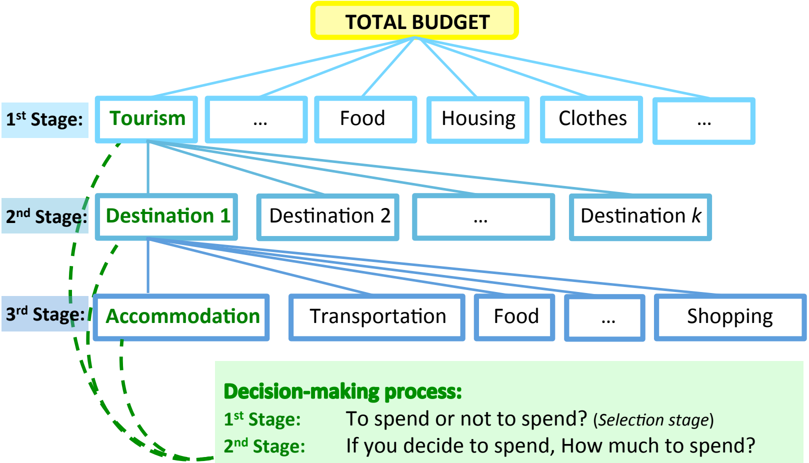
Disegna, M., Durante, F., Foscolo, E. (2013). A multivariate nonlinear analysis of tourism expenditures. Available at BEMPS (working papers): `http://www.unibz.it/en/economics/research/workingpapers/default.html`

We thank FUB, Banca di Trento e Bolzano, and Intesa Sanpaolo

STTOBS - South Tyrolean Tourism OBServatory

Turismo in Alto Adige:
un'esperienza olistica per il benessere di corpo, mente e spirito!

Visitors' expenditures



Decision-making process:

1st Stage: To spend or not to spend? (*Selection stage*)

2nd Stage: If you decide to spend, How much to spend? (*Outcome stage*)

(Brida et al., 2013; Kim et al., 2010; Alegre et al., 2010; Alegre & Pou, 2004)

(The three-stage budgeting process proposed by Syriopoulos & Sinclair, 1993)

The main question

Focusing on Selection stage,

How can we deal with dependence among single decisions?

From a probabilistic point of view . . .

Let Y_j be a dichotomous variable describing the decision to spend ($Y_j = 1$) or not ($Y_j = 0$) in the j-th expenditure category

How to define $\mathbb{P}\big(Y_1 = y_1, \ldots, Y_J = y_J \,|\, \boldsymbol{X}\big)$, $y_j \in \{0, 1\}$?

Our solution (roughly speaking . . .)

Fitting univariate Logit regressions first

⇓

Grouping them by means of a **Copula function**

What Copulas are?

Multivariate distribution functions whose univariate margins are uniformly distributed on $[0, 1]$

The best alternative (at the present) to several classical multivariate distribution functions such as Normal, *t* Student, or Pareto distributions

What Copulas are?

Multivariate distribution functions whose univariate margins are uniformly distributed on $[0, 1]$

The best alternative (at the present) to several classical multivariate distribution functions such as Normal, *t* Student, or Pareto distributions

Why Copulas? (I)

- Allow to model discrete as well as continuous random variables
- Flexibility in describing extremal dependence in a convenient way

Why Copulas? (I)

- Allow to model discrete as well as continuous random variables
- Flexibility in describing extremal dependence in a convenient way

Why Copulas? (II)

- Permit the combination of any univariate marginal distributions that do not need to come from the same family

Able to accommodate various decision-making rules

- Encompass a number of existing multivariate models, and provide a framework for generating many more

Able to reflect various behavioral interactions in a consistent way

Why Copulas? (II)

- Permit the combination of any univariate marginal distributions that do not need to come from the same family

Able to accommodate various decision-making rules

- Encompass a number of existing multivariate models, and provide a framework for generating many more

Able to reflect various behavioral interactions in a consistent way

Why Copulas? (II)

- Permit the combination of any univariate marginal distributions that do not need to come from the same family

Able to accommodate various decision-making rules

- Encompass a number of existing multivariate models, and provide a framework for generating many more

Able to reflect various behavioral interactions in a consistent way

Why Copulas? (II)

- Permit the combination of any univariate marginal distributions that do not need to come from the same family

Able to accommodate various decision-making rules

- Encompass a number of existing multivariate models, and provide a framework for generating many more

Able to reflect various behavioral interactions in a consistent way

The Copula-based logit choice model

$$\mathbb{P}\left(Y_1 = y_1, \ \ldots, Y_J = y_J \mid \boldsymbol{X}\right) = \\ = \psi\left\{F_1\left(y_1 \mid \boldsymbol{X}_1\right), \ldots, F_J\left(y_J \mid \boldsymbol{X}_J\right); C\right\}$$

- $C \in \mathcal{C}_{\boldsymbol{\theta}}$
- $F_j\left(y_j \middle| \boldsymbol{X}_j\right) = \begin{cases} 0 & y_j < 0 \\ \frac{1}{1+\exp\left(\boldsymbol{X}_j^{\top} \boldsymbol{\beta}_j\right)} & 0 \leq y_j < 1 \\ 1 & y_j \geq 1 \end{cases}, \quad \boldsymbol{X}_j \subseteq \boldsymbol{X}$
- see Song (2007) for a definition of $\psi(\cdot)$

Go to the formula

The Copula-based logit choice model

$$\mathbb{P}\left(Y_1 = y_1, \ \ldots, Y_J = y_J \mid \boldsymbol{X}\right) = \\ = \psi\left\{F_1\left(y_1 \mid \boldsymbol{X}_1\right), \ldots, F_J\left(y_J \mid \boldsymbol{X}_J\right); C\right\}$$

- $C \in \mathcal{C}_{\boldsymbol{\theta}}$
- $F_j\left(y_j \middle| \boldsymbol{X}_j\right) = \begin{cases} 0 & y_j < 0 \\ \frac{1}{1+\exp\left(\boldsymbol{X}_j^\top \boldsymbol{\beta}_j\right)} & 0 \leq y_j < 1 \\ 1 & y_j \geq 1 \end{cases}, \quad \boldsymbol{X}_j \subseteq \boldsymbol{X}$
- see Song (2007) for a definition of $\psi\left(\cdot\right)$

Go to the formula

How to estimate the Copula-based logit choice model? (I)

Two-step estimator (known as **the inference functions for margins method**)

(1) Fitting marginal parametric distributions first; i.e., obtaining the maximum likelihood estimates $\hat{\boldsymbol{\beta}}_1, \ldots, \hat{\boldsymbol{\beta}}_J$

(2) Fitting the parametric Copula for fixed margins afterwards; i.e., obtaining the maximum (composite) likelihood estimate $\hat{\boldsymbol{\theta}}$

Asymptotic properties have been investigated by Joe (1997, 2005)

How to estimate the Copula-based logit choice model? (I)

Two-step estimator (known as **the inference functions for margins method**)

(1) Fitting marginal parametric distributions first; i.e., obtaining the maximum likelihood estimates $\hat{\boldsymbol{\beta}}_1, \ldots, \hat{\boldsymbol{\beta}}_J$

(2) Fitting the parametric Copula for fixed margins afterwards; i.e., obtaining the maximum (composite) likelihood estimate $\hat{\boldsymbol{\theta}}$

Asymptotic properties have been investigated by Joe (1997, 2005)

How to estimate the Copula-based logit choice model? (I)

Two-step estimator (known as **the inference functions for margins method**)

(1) Fitting marginal parametric distributions first; i.e., obtaining the maximum likelihood estimates $\hat{\boldsymbol{\beta}}_1, \dots, \hat{\boldsymbol{\beta}}_J$

(2) Fitting the parametric Copula for fixed margins afterwards; i.e., obtaining the maximum (composite) likelihood estimate $\hat{\boldsymbol{\theta}}$

Asymptotic properties have been investigated by Joe (1997, 2005)

How to estimate the Copula-based logit choice model? (II)

In practice, different Copula families may be fitted to the same set of univariate margins

A criterion is need in order to select a proper Copula among the possible different functions

$\Downarrow$

Perform the likelihood-ratio test for nested models due to Vuong (1989)

How to estimate the Copula-based logit choice model? (II)

In practice, different Copula families may be fitted to the same set of univariate margins

A criterion is need in order to select a proper Copula among the possible different functions

$\Downarrow$

Perform the likelihood-ratio test for nested models due to Vuong (1989)

The method at work

Dataset

"**International Tourism in Italy**": annual survey, Bank of Italy (2011)

Observation

550 international visitors who have visited any city in South–Tyrol for "tourism, holiday, leisure"

Univariate Logit regression results

	Accommodation	Transportation	Shopping
Food & beverages expenditure	0.001** (0.001)	0.001** (0.001)	0.001* (0.001)
Expenditure for other services	0.001* (0.001)	0.001** (0.001)	0.001 (0.001)
German–speaking	−1.049** (0.48)	−1.078*** (0.35)	0.473** (0.22)
Retired	−1.447*** (0.3)	−1.158*** (0.26)	−0.038 (0.24)
Alone traveler	−1.273*** (0.33)	−0.895*** (0.28)	−0.285 (0.29)
Safety (Level of satisfaction)	0.700*** (0.12)	0.808*** (0.14)	0.424*** (0.08)
Constant	−2.796*** (1.02)	−4.145*** (1.19)	−3.272*** (0.73)

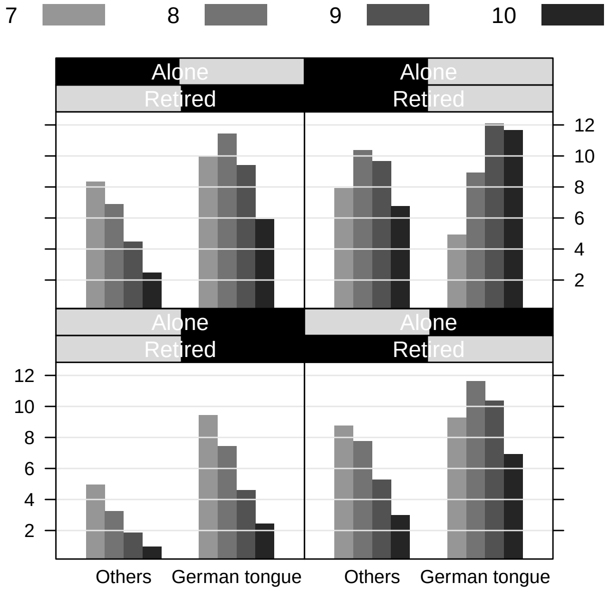
$\star \ldots p < 10\%$, $\star\star \ldots p < 5\%$, $\star\star\star \ldots p < 1\%$

The core

14 Copula models ⇒ the **Normal Copula**

$$\hat{\boldsymbol{\theta}} = \left\{ \underset{(0.012)}{0.729^{\star\star\star}}, \underset{(0.023)}{0.193^{\star\star\star}}, \underset{(0.001)}{0.246^{\star\star\star}} \right\}$$

Which differences? ⏩ Go to the output



Concluding remarks (I)

Aim

study the interactions in tourists' decisions

Given a set of explanatory variables, the Copula-based logit choice model is able to predict the probability that the tourist is likely to spend in two or more categories

Concluding remarks (II)

Case study

Data collected by Bank of Italy in 2011 throughout the survey entitled “International Tourism in Italy” are analyzed

Main finding independence among expenditures underestimates the probability of spending in all categories

Implications

specific marketing campaign that offer a combination of different services (meals, lodging, shopping, etc.) according to tourists’ profile

What now?

Money, of course!

In other words, we leave a thorough investigation of the second stage (the outcome stage) and **its integration with the selection phase** to future research

What now?

Money, of course!

In other words, we leave a thorough investigation of the second stage (the outcome stage) and **its integration with the selection phase** to future research

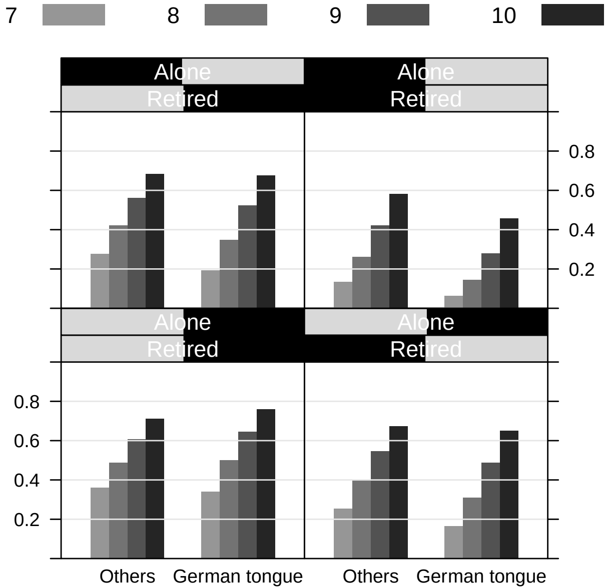
Song (2007)

$$\mathbb{P}\left(Y_1 = y_1, \ldots, Y_J = y_J \mid \boldsymbol{X}\right) =$$

$$= \sum_{j_1=0}^{y_1} \ldots \sum_{j_J=0}^{y_J} (-1)^{\sum_{k=1}^{J}(j_k + y_k)} C\left\{F_1\left(j_1 \mid \boldsymbol{X}_1\right), \ldots, F_J\left(j_J \mid \boldsymbol{X}_J\right)\right\}$$

Go Back

The Normal-based output ◂ Go Back



The Independence-based output

Go Back