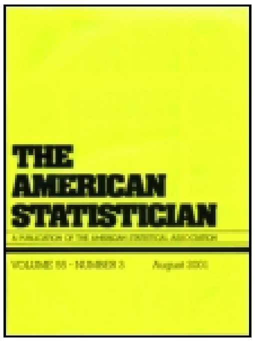

The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: http://www.tandfonline.com/loi/utas20

A graphical tool for interpreting regression coefficients of trinomial logit models

Flavio Santi, Maria Michela Dickson & Giuseppe Espa

To cite this article: Flavio Santi, Maria Michela Dickson & Giuseppe Espa (2018): A graphical tool for interpreting regression coefficients of trinomial logit models, The American Statistician, DOI: 10.1080/00031305.2018.1442368

To link to this article: https://doi.org/10.1080/00031305.2018.1442368

Accepted author version posted online: 09 Mar 2018.

A graphical tool for interpreting regression coefficients of trinomial logit models

Flavio Santi, Maria Michela Dickson and Giuseppe Espa*
Department of Economics and Management, University of Trento.

January 23, 2018

Abstract

Multinomial logit (also termed *multi-logit*) models permits the analysis of the statistical relation between a categorical response variable and a set of explicative variables (called *covariates* or *regressors*). Although multinomial logit is widely used in both the social and economic sciences, the interpretation of regression coefficients may be tricky, as the effect of covariates on the probability distribution of the response variable is non-constant and difficult to quantify. The ternary plots illustrated in this paper aim at facilitating the interpretation of regression coefficients and permit the effect of covariates (either singularly or jointly considered) on the probability distribution of the dependent variable to be quantified. Ternary plots can be drawn both for ordered and for unordered categorical dependent variables, when the number of possible outcomes equals three (*trinomial response variable*); these plots allow not only to represent the covariate effects over the whole parameter space of the dependent variable but also to compare the covariate effects of any given individual profile. The method is illustrated and discussed through analysis of a dataset concerning the transition of master's graduates of the University of Trento (Italy) from university to employment.

Keywords: multinomial logit model, ternary plot, vector field, marginal effect, generalised linear model

*We would like to thank Rocco Micciolo (*Department of Psychology and Cognitive Science, University of Trento*) and Marco Alfò (*Department of Statistical Sciences, Sapienza University of Rome*) who read a preliminary version of this paper and provided several valuable comments and suggestions.

ACCEPTED MANUSCRIPT

1 Introduction

The linear regression model (see e.g. Wooldridge, 2008)

$$y_i = \beta_1 + \beta_2 \, x_{2,i} + \ldots + \beta_p \, x_{p,i} + \varepsilon_i, \qquad i = 1, \ldots, n \tag{1}$$

enables the analysis of the statistical relation between the variable of interest y (usually referred to as *response variable*, *outcome*, or *predicted variable*) and a group of explanatory variables $x_1, \ldots, x_p$ (also called *regressors*, *covariates*, or *predictors*) within a given dataset.[1] Nevertheless, Equation (1) consistently models only response variables which can virtually take any (positive or negative) real value. In the event that this requirement cannot be met even approximately, model (1) should be restated so as to obtain an accurate and consistent mathematical description of the phenomenon to be studied.

A restatement is necessary, for example, when the response variable y is categorical and takes values in a set of qualitative labels $\{v^{(1)}, \ldots, v^{(M)}\}$ in which the number of elements M is finite ($M < \infty$). From a probabilistic point of view, categorical variables are modelled by means of the single-trial multinomial random variable (Johnson et al., 2005, 522-524), whereas their statistical relation to other variables (covariates) is analysed through multi-logit regression (Madsen and Thyregod, 2010; Agresti, 2002; Fahrmeir and Tutz, 1994).

In this paper we focus on the special case of trinomial response variables (that is, categorical variables where the number M of possible values is three), as they represent an important class of multinomial response variables both in case of ordinal and non-ordinal outcomes (Lipovetsky, 2015).

As it is illustrated in Section 2, the (non-ordinal) multi-logit (or *multinomial logit*) regression enables the statistical relation between the probability distribution of the response variable Y and a set of p predictors $x \in \mathbb{R}^p$ to be modelled as follows:

$$\log \left(\pi^{(m)}_{(x)} / \pi^{(1)}_{(x)} \right) = x^{\mathrm{T}} \beta^{(m)} \,, \tag{2}$$

where $\beta^{(m)} \in \mathbb{R}^p$ is the m-th vector of regressor coefficients, $\pi^{(m)}_{(x)} \equiv \mathbb{P}[Y = v^{(m)} \mid x]$, and $m = 2, 3$.

[1] Wooldridge (2008, ch. 2) discusses differences in terminology regarding names used for the variables on the left and the right hand side of (1).

From (2) it follows that (see Section 2):

$$\pi_{(x)}^{(m)} = \frac{\mathrm{e}^{x^{\mathrm{T}}\beta^{(m)}}}{1 + \mathrm{e}^{x^{\mathrm{T}}\beta^{(2)}} + \mathrm{e}^{x^{\mathrm{T}}\beta^{(3)}}}\,, \tag{3}$$

for $m = 1, 2, 3$, and, for notational convenience, $\beta^{(1)} \equiv 0 \in \mathbb{R}^p$.

Both Equation (2) and (3) demonstrates why the interpretation of regression coefficients of trinomial multi-logit models is complicated. In particular, Equation (2) allows us to interpret $\beta_j^{(m)}$ as the marginal effect of the j-th regressor on the logarithm of the relative probability $\pi^{(m)}/\pi^{(1)}$ which is often of little interest, whereas the functional form in (3) makes the relation between regressors and probabilities $\pi^{(m)}$ difficult to analyse.

Ordinal logit models pose analogous problems in assessing the effects of the covariates on the probability distribution of the dependent variable. Indeed, the interpretation of regression coefficients may still be tricky even though model specification considers only one coefficient β_j for each regressor j (see Section 3).

Several numerical indicators and graphical methods are usually adopted in order to ease the interpretation of multinomial logit estimates. The effect of continuous regressors is usually plotted on a Cartesian plane where the predicted probabilities are drawn as functions of the covariate values (see e.g. Agresti, 2002, ch. 7), nevertheless the graphical outcome may not be satisfactory in the case of discrete regressors, especially when it comes to dummy covariates. Numerical evaluation of covariate effects may be considered either in place of or in conjunction with graphical methods, as they permit the quantification (exact or approximate) of the covariate marginal effects (see e.g. Lipovetsky, 2011). However, both of these types of methods (graphical and numerical) can be used for studying the effect of a certain regressor, provided that values of the others are fixed. It follows that a single graph or a single numerical indicator may not be able to describe the relation between the covariate and the distribution probability of the dependent variable over the whole parameter space $\Theta = \{(\pi^{(1)}, \pi^{(2)}, \pi^{(3)}) \in [0,1]^3 \colon \pi^{(1)} + \pi^{(2)} + \pi^{(3)} = 1\}$.

This paper proposes a new graphical method for representing the effect of covariates on the probability distribution of a trinomial ($M = 3$) dependent variable Y over its whole parameter space Θ, so as to make the interpretation of regression coefficients estimates easier. The proposed method is based on a ternary plot (Bancroft, 1897) where each point represents a probability distribution (π_1, π_2, π_3) of Y, and the effect of the regressors is represented by a vector field drawn on the plot area. Each arrow of the vector field shows how the probability distribution (π_1, π_2, π_3) changes when the values of the regressors $x \in \mathbb{R}^q$ change.

The paper is organised as follows. Section 2 introduces the multi-logit model for the (unordered) trinomial response variable, discusses the problem of measuring the marginal effects of covariates and illustrates them within the trinomial-logit model. Section 3 briefly discusses the ordinal multinomial logit models and their covariates' effects as a special case of non-ordered categorical variables. In Section 4 ternary plots are illustrated and it is shown how a vector field can be drawn on a ternary plot. In Section 5 the method is applied to a real-data example regarding the employment status and the final degree score of Trento University of (Italy) master's graduates. Section 6 concludes the paper.

2 Non-ordinal trinomial logit model

Equations (2) and (3) characterise the multinomial logit model. The probabilistic framework where they derive from is briefly outlined in this section for trinomial dependent variables.

Let Y be a categorical random variable which takes values in $\{v^{(1)}, v^{(2)}, v^{(3)}\}$ with probabilities $\mathbb{P}[Y = v^{(m)}] = \pi^{(m)}$ for $m = 1, 2, 3$. If we define the indicator binary variable $D^{(m)}$ as:

$$D^{(m)} \equiv \begin{cases} 1 & \text{if } Y = v^{(m)} \\ 0 & \text{otherwise} \end{cases}$$

it is possible to express the probability of the realisation y of Y as follows:

$$\mathbb{P}[Y = y] = \prod_{m=1}^{3} \left(\pi^{(m)}\right)^{d^{(m)}} , \tag{4}$$

being $d^{(m)}$ the realisation of the random variable $D^{(m)}$.[2] The advantage of notation (4) is that it holds for any realisation $y \in \{v^{(1)}, v^{(2)}, v^{(3)}\}$.

[2]Note that $D^{(m)}$ is distributed as a Bernoulli random variable with probability $\pi^{(m)}$, that is $D^{(m)} \sim \mathcal{B}(\pi^{(m)})$. This result justifies the estimation approach of multi-logit models based on fitting $M - 1 = 2$ logit models. See Agresti (2002, pp. 272-274) for methodological details and Begg and Gray (1984) for estimators' properties.

The random vector $(D^{(1)}, D^{(2)}, D^{(3)})$ is distributed as a single-trial multinomial random variable having three possible outcomes (see e.g. Johnson et al., 2005, 522-524). It can be proved (Lehmann and Casella, 1998, 24-25) that $\eta_2 \equiv \log(\pi^{(2)}/\pi^{(1)})$ and $\eta_3 \equiv \log(\pi^{(3)}/\pi^{(1)})$ are the natural parameters of $(D^{(1)}, D^{(2)}, D^{(3)})$ and $\mathbb{R}^2$ is its natural parameter space.[3] It follows that the probability distributions of Y and $(D^{(1)}, D^{(2)}, D^{(3)})$ are fully identified by η_2 and η_3 (jointly considered) which can therefore be modelled by means of linear predictors $x^{\mathrm{T}}\beta^{(2)}$ and $x^{\mathrm{T}}\beta^{(3)}$ according to Equation (2).

Given the formal framework we have just outlined, Equation (3) and other formal properties of multi-logit regression can be easily derived. Agresti (2002) and Fahrmeir and Tutz (1994) give a detailed formal illustration of the multi-logit model.

The marginal effect of a regressor on the probability $\pi^{(m)}$ $(m = 1, 2, 3)$ can be computed in various ways according to the nature of the regressor itself. In case of continuous regressors, the derivative $\mathrm{d}\pi^{(m)}/\mathrm{d}x_k$ may be an appropriate measure of the marginal effect of the k-th regressor on $\pi^{(m)}$, however also the logarithmic derivative $\mathrm{d}\log(\pi^{(m)})/\mathrm{d}x_h$, or the relative change $\mathrm{d}\log(\pi^{(m)})/\mathrm{d}\log(x_k)$ may be appropriate. On the other hand, if x_k is a binary or (more generally) a discrete regressor, the finite difference $\pi^{(m)}_{(x+\Delta)} - \pi^{(m)}_{(x)}$, the incremental ratio $(\pi^{(m)}_{(x+\Delta)} - \pi^{(m)}_{(x)})/\Delta$, or the ratio $\pi^{(m)}_{(x+\Delta)}/\pi^{(m)}_{(x)}$ may be suitable measures of marginal effects, $\Delta \in \mathbb{R}^p$ being the unitary (or fractional) change of the k-th regressor.[4] Since all these measures can be computed through appropriate transformations of the others (see e.g. Lipovetsky, 2011, for an overview), in the following only $\mathrm{d}\pi^{(m)}_{(x)}/\mathrm{d}x_k$ and $\pi^{(m)}_{(x+\Delta)}/\pi^{(m)}_{(x)}$ are considered.

The first derivative of (3) with respect to the value of the k-th regressor (with $k = 1, \ldots, p$) is:

$$\frac{\mathrm{d}\pi^{(m)}_{(x)}}{\mathrm{d}x_k} = \left[\beta^{(m)}_k - \sum_{h=2}^{3} \beta^{(h)}_k \pi^{(h)}_{(x)}\right] \pi^{(m)}_{(x)}\,, \tag{5}$$

being $\beta^{(m)}_k \in \mathbb{R}$ the k-th component of vector $\beta^{(m)} \in \mathbb{R}^p$.

If the effect of a finite change $\Delta \in \mathbb{R}^p$ of one or more than one regressor is considered,

[3] For a formal definition of natural parameters and natural parameter space see Lehmann and Casella (1998, pp. 23-24).

[4] For instance, in case of a trinomial logit model with $p = 5$ regressors, if $k = 4$ and the fourth regressor is binary, then $\Delta = [0, 0, 0, 1, 0]^{\mathrm{T}}$.

few algebraic manipulations permit us to verify that:

$$\frac{\pi^{(m)}_{(x+\Delta)}}{\pi^{(m)}_{(x)}} = \left[1 - \sum_{h=2}^{3} \left(1 - \mathrm{e}^{\Delta^{\mathrm{T}}\beta^{(h)}}\right) \pi^{(h)}_{(x)}\right]^{-1} \mathrm{e}^{\Delta^{\mathrm{T}}\beta^{(m)}} ; \tag{6}$$

whereas, in the specific event that only one regressor (say the k-th) increases by one (that is $\Delta = \hat{e}_k$, being $\hat{e}_k \in \mathbb{R}^p$ the versor of the k-th axis in $\mathbb{R}^p$),[5] Equation (6) changes to:

$$\frac{\pi^{(m)}_{(x+\hat{e}_k)}}{\pi^{(m)}_{(x)}} = \left[1 - \sum_{h=2}^{3} \left(1 - \mathrm{e}^{\beta^{(h)}_k}\right) \pi^{(h)}_{(x)}\right]^{-1} \mathrm{e}^{\beta^{(m)}_k} .$$

Two issues emerge from Equations (5) and (6) and deserve particular attention.

Firstly, since Equation (6) holds also for changes $\Delta \in \mathbb{R}^k$ involving more than a single regressor (consider, for example $\Delta = [1, 0, 1, 0, 0]^{\mathrm{T}}$, where $k = 5$), it makes it possible to assess the joint effects of the regressors (that is, the effect of a change in two or more independent regressors), as well as changes in regressors with non-linear and interaction terms.

Secondly, both (5) and (6) are functions of vectors $\beta^{(2)}$ and $\beta^{(3)}$, of the probability distribution $(\pi^{(1)}, \pi^{(2)}, \pi^{(3)})$ in $x \in \mathbb{R}^p$, in addition to the marginal effect $\Delta \in \mathbb{R}^p$ (only for Equation (6)). This means that neither (5) nor (6) depend on the values of the regressors $x \in \mathbb{R}^p$ once the previous elements are considered. This property, which is shared by all generalised linear models (see e.g. Madsen and Thyregod, 2010, ch. 4), permits the effect of any finite or infinitesimal change of regressor values to be analysed once the probability distribution is known $(\pi^{(1)}, \pi^{(2)}, \pi^{(3)})$. This is independent of the actual values of regressors $x \in \mathbb{R}^p$. As a consequence, it is possible to study the effect of the regressors over the 2-dimensional parameter space $\Theta = \{(\pi^{(1)}, \pi^{(2)}, \pi^{(3)}) \in [0, 1]^3 : \pi^{(1)} + \pi^{(2)} + \pi^{(3)} = 1\}$ instead of the p-dimensional regressor space $\mathcal{A} \ni X$. This represents an enormous advantage, as typically $p > 2$.

Nonetheless, the aforementioned property does not hold if the regressors are not functionally independent. In particular, when either interaction terms or non-linear terms are

[5] The k-th versor $\hat{e}_k$ of a p-dimensional vector space is a vector of p zero elements except for the k-th which is equal to one. Thus, if $p = 4$ we see that $\hat{e}_1 = [1, 0, 0, 0]^{\mathrm{T}}$, $\hat{e}_2 = [0, 1, 0, 0]^{\mathrm{T}}$, $\hat{e}_3 = [0, 0, 1, 0]^{\mathrm{T}}$, $\hat{e}_4 = [0, 0, 0, 1]^{\mathrm{T}}$.

If the vector space is normed, versors have unit norm.

ACCEPTED MANUSCRIPT

included along with the linear terms, it is necessary to analyse the effect of the regressors over the space $\Theta \times \mathcal{B}$, $\mathcal{B} \subseteq \mathbb{R}^{p'}$ being a subspace of $\mathcal{A} \subseteq \mathbb{R}^p$ (for some $p' < p$) including all regressors which are functionally-dependent (Ibragimov, 2010) on regressors of interest.

Consider, for example, a trinomial logit model including $(p-1)$ functionally-independent regressors $X_1, \ldots, X_{p-1}$ and the interaction term $X_1 X_2$. The effect of X_1 on the probability distribution of the dependent variable cannot be assessed independently of X_2, thus a value x_2^* for X_2 should be fixed ($\mathcal{B} \subseteq \mathbb{R}$) and the effect of X_1 can be analysed over Θ for $X_2 = x_2^*$.

Analogous considerations hold when two or more regressors are included as different transformations of the same statistical variable. Consider, for example, the case where $X_1 = X$ and $X_2 = X^2$, X being a statistical variable. Also in such a situation, the effect of X on the probability distribution of the dependent variable should be analysed over Θ with respect to a fixed value $X = x^*$. Nonetheless, note that, when the effect of a change δx in X has to be assessed, the condition $\Delta = (\delta x,\, (\delta x)^2 + 2x\,\delta x)^{\mathsf{T}}$ must hold in order to guarantee the consistency of the analysis.[6]

3 Ordinal models

When the values of a categorical variable can be ordered according to specific criterion, the multi-logit model should be restated so as to consider such information. Several methodological approaches have been proposed in the literature with the aim of regressing ordered categorical responses over a set of covariantes; further discussion on the main modelling solutions can be found in Agresti (2002).

Here we only consider the *proportional odds model* (Agresti, 2002, pp. 275 ff.) for illustration purposes, however the method discussed in this paper can be easily extended to other multinomial regression models.

According to the notation used so far, Y is a categorical random variable which takes values in $\{v^{(1)}, v^{(3)}, v^{(3)}\}$ with probabilities $\mathbb{P}[Y = v^{(m)}] = \pi^{(m)}$ for $m = 1, 2, 3$. It is assumed that labels of Y can be ordered according to some relation (order) $\prec$ such that $v^{(1)} \prec v^{(3)} \prec v^{(3)}$. Within this framework, the probability $\mathbb{P}[Y = y]$ corresponds to (4), whereas

[6]Note that $(\delta x)^2 + 2x\,\delta x = (x + \delta x)^2 - x^2$.

the probabilities $\pi^{(m)}$ $(m = 1, 2, 3)$ are modelled through the following parametrisation:

$$\operatorname{logit}(\mathbb{P}[Y \leq v^{(m)} \mid x]) = \alpha_m - x^{\mathrm{T}}\beta \qquad m = 1, 2\,, \tag{7}$$

where α_m $(m = 1, 2)$ are constants, $\beta \in \mathbb{R}^{p-1}$ are covariate coefficients, and $x \in \mathbb{R}^{p-1}$ are regressors without the constant term.

From (7) it follows that (Agresti, 2002):

$$\pi^{(m)}_{(x)} = S^{(m)}_{(x)} - S^{(m-1)}_{(x)}\,. \tag{8a}$$

where $S^{(m)}_{(x)} \equiv \sum_{j=1}^{m} \pi^{(j)}_{(x)}$ is the cumulative probability and equals:

$$S^{(m)}_{(x)} = \frac{\mathrm{e}^{\alpha_m - x^{\mathrm{T}}\beta}}{1 + \mathrm{e}^{\alpha_m - x^{\mathrm{T}}\beta}}\,, \tag{8b}$$

$S^{(0)}_{(x)} \equiv 0$ being for notational convenience.

Equations (8) allow to compute the effect of the change $\Delta \in \mathbb{R}^{p-1}$ of regressors $x \in \mathbb{R}^{p-1}$ as follows:

$$\pi^{(m)}_{(x+\Delta)} = \frac{\mathrm{e}^{\Delta^{\mathrm{T}}\beta}\, S^{(m)}_{(x)}}{1 - (1 - \mathrm{e}^{\Delta^{\mathrm{T}}\beta})\, S^{(m)}_{(x)}} - \frac{\mathrm{e}^{\Delta^{\mathrm{T}}\beta}\, S^{(m-1)}_{(x)}}{1 - (1 - \mathrm{e}^{\Delta^{\mathrm{T}}\beta})\, S^{(m-1)}_{(x)}}\,. \tag{9}$$

In line with unordered multi-logit model, the effect of covariates on the probability distribution of the response variable only depends on the vector of change Δ, the covariate coefficients β, and the initial probability distribution $(\pi^{(1)}_{(x)}, \ldots, \pi^{(M)}_{(x)})$ (through $S^{(m)}_{(x)}$, $m = 1, 2, 3$).

4 Ternary plots

Ternary plots allows a triangular region to be represented as a result of some linear restriction on a bounded 3-dimensional space. In particular, it is convenient to use a ternary plot when it is necessary to graphically represent points whose position is identified by a system of three-dimensional coordinates. In any event, these coordinates must comply with a linear restriction (see e.g. Hennig et al., 2016; Graff, 2017, for a discussion on ternary plots in the context of cluster analysis and correspondence analysis respectively).

This is the case of the probabilities $(\pi^{(1)}, \pi^{(2)}, \pi^{(3)})$ of a trinomial random variable Y: each probability distribution of Y is uniquely identified by three probabilities which must

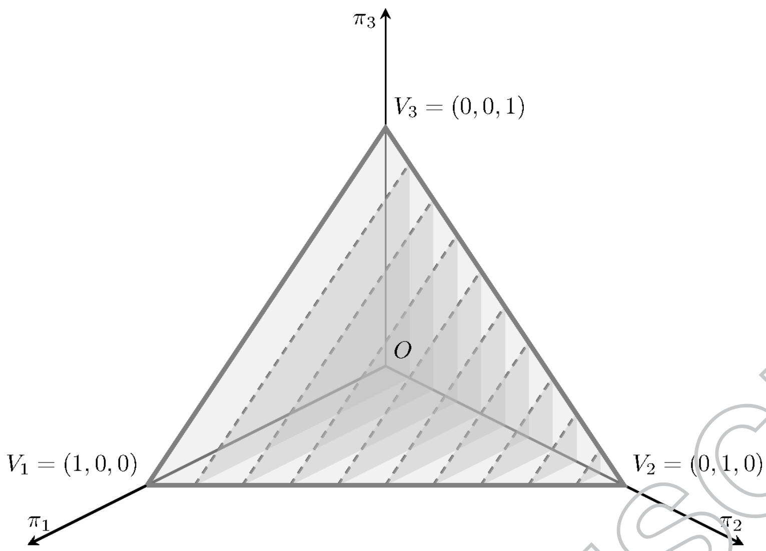

Figure 1: Parameter space Θ drawn in a 3-dimensional coordinate system. Each dashed line represents a set of points in Θ where π_2 has a constant value. The shaded triangles show the relation between the Cartesian and the ternary coordinates with regard to the component π_2.

always sum to one. It follows that the actual dimension of the parameter space of Y is 2, as noted in Section 2.

Figure 1 shows the symplex which represents the parameter space Θ in $\mathbb{R}^3$. The triangle having vertexes in $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ is the ternary plot which is used in this paper.

Figure 1 also shows how the component π_2 is represented in a ternary plot: the smaller is the distance between a point and the side (V_1V_3) opposite the vertex $V_2 = (0, 1, 0)$, the lower is the probability π_2. Moreover, points situated on a line which is parallel to the side V_1V_3 (dashed lines in Figure 1) share a constant value of π_2. Analogous conclusion can be drawn for components π_1 and π_3.

Figure 2 illustrates how to interpret the position of a generic point P on a ternary plot and shows how ternary plots are typically represented. Unlike Figure 1, only the triangle

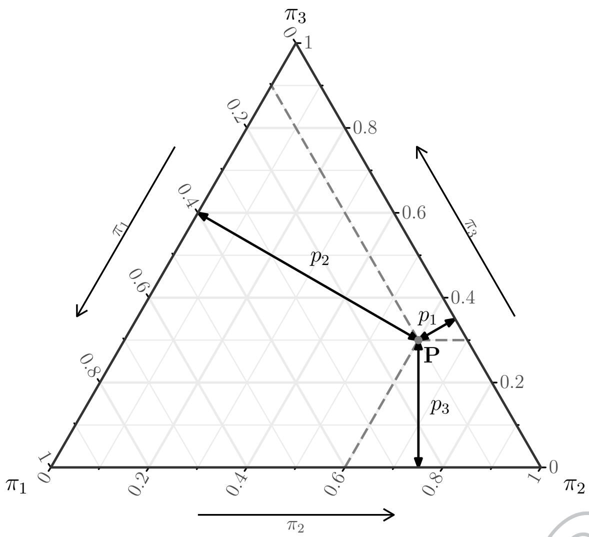

Figure 2: Coordinates of point $P = (p_1, p_2, p_3) = (0.1, 0.6, 0.3)$ (in grey) in a ternary plot. The lengths of the double arrows show the values of the ternary coordinates of P; the dashed lines show the value of the ternary coordinates on the reference axes.

$\triangle V_1V_2V_3$ is represented, whereas the original Cartesian coordinate system (π_1, π_2, π_3) is replaced by the names of the original axes near the corresponding vertex of the triangle. A grid of ternary coordinates and the associated values are added to the plot in order to improve the readability of the graph, and three arrows with the name and the direction of the ternary coordinate help in reading the values on ternary axes.

As discussed in Section 2, if the regressors of a multi-logit model are functionally independent, it is possible to study their effect on the probability distribution of the dependent variable over the parameter space $\Theta \in \mathbb{R}^2$. This means that it is possible to represent the effect of regressor changes by means of a vector field defined on a ternary plot.

To illustrate this point, consider, for example, a trinomial random variable with probability distribution $A = (0.1, 0.6, 0.3)$, and assume that Equation (6) permits a computation demonstrating that the finite change $\Delta \in \mathbb{R}^p$ of the regressors $x \in \mathbb{R}^p$ leads to a new distribution $B = (0.4, 0.4, 0.2)$. In this case, it is possible to represent the effect of the changes Δ as in Figure 3.

Figure 3 suggests that it is possible to consider a set of points $\{A_1, A_2, \ldots, A_K\}$ well

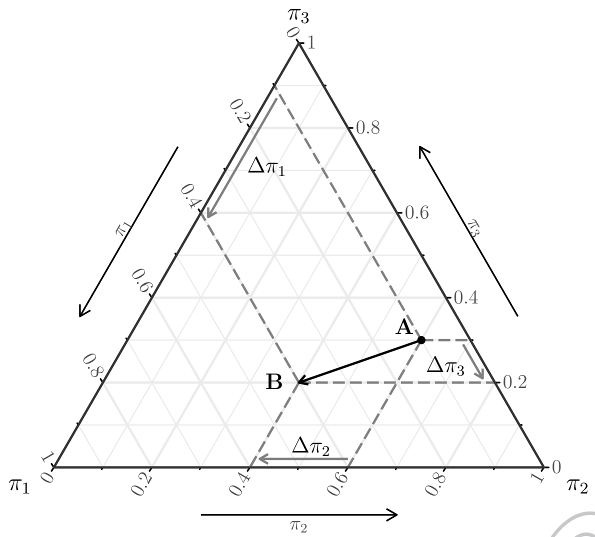

Figure 3: Decomposition of the change from the probability distribution $A = (0.1, 0.6, 0.3)$ to $B = (0.4, 0.4, 0.2)$ in terms of the changes of the coordinates π_1, π_2, π_3. Note that the total change of the probabilities equals zero ($\Delta\pi_1 + \Delta\pi_2 + \Delta\pi_3 = 0.3 - 0.2 - 0.1 = 0$) as the probabilities of both A and B must sum to one.

distributed over Θ, compute the distributions $\{B_1, B_2, \ldots, B_K\}$ resulting from the change $\Delta \in \mathbb{R}^p$ of the regressors, and draw the related arrows in a ternary plot. The resulting vector field can therefore provide a precise graphical representation of the covariate effect Δ on the probability distribution $(\pi^{(1)}, \pi^{(2)}, \pi^{(3)})$ over the entire parametric space Θ. An application of the method to real data is illustrated and discussed in Section 5.

In this respect it is worth pointing out that the method just illustrated aims solely at facilitating the interpretation of estimated regression coefficients, thus, any specification problem affecting the fitted regression will possibly alter the ternary plot too and the subsequent conclusions drawn from it.

This occurs, for example, in model misspecifications and multicollinearity, which typically have remarkable effects on coefficient point estimates in terms both of magnitude and sign. Ternary plots based on such estimates may exhibit a vector field where all or some arrows are headed in the wrong direction or are either too long or too short. In general, a ternary plot cannot reveal whether the underling regression model suffers from

Table 1: Multi-logit estimates (standard errors in parenthesis; p-values in brackets) for the employment situation of University of Trento's master's graduates one year after graduation. The reference category of the dependent variable is *employed*. The reference master's graduate implied by regressors is a *male* with *average* final grade, *average* duration, belonging to the *working class*.

	Dependent variable:					
	Trainee			Unemployed		
	(1)			(2)		
Constant	−2.521	(0.200)	[0.000]	−1.870	(0.160)	[0.000]
Female	0.378	(0.131)	[0.004]	0.449	(0.109)	[0.000]
Low final score	0.258	(0.151)	[0.089]	0.207	(0.127)	[0.103]
High final score	−0.345	(0.162)	[0.033]	−0.182	(0.130)	[0.163]
Short duration	0.298	(0.156)	[0.057]	−0.278	(0.130)	[0.033]
Long duration	0.024	(0.164)	[0.884]	−0.146	(0.126)	[0.248]
White-collar workers	−0.053	(0.172)	[0.759]	−0.025	(0.139)	[0.858]
Lower middle class	0.022	(0.185)	[0.908]	0.007	(0.151)	[0.962]
Upper middle class	−0.030	(0.189)	[0.877]	−0.238	(0.162)	[0.144]
Unclassified	−0.924	(0.736)	[0.209]	−0.147	(0.422)	[0.729]

misspecifications or multicollinearity, except through unexpected or anomalous covariate effects represented in the random field.

5 An application to real data

In this section, ternary plots are used to analyse the estimates of two trinomial logit models (ordered and unordered respectively) on a dataset regarding Trento University Master's graduates' transition from university to employment.

The dataset consists of 3282 observations about master's students graduated at the University of Trento (Italy) between 2009 and 2013. For each graduate we have information on gender, final degree score (classified as *low*, *average*, and *high*, according to the degree course distribution), duration of studies (classified as *short*, *average*, and *long*, according to the degree course distribution), in addition to the high school final score (which ranges between 60 and 100). A further variable which measures irregularity in the students' studies is included in the analysis as a categorical variable which takes values *low*, *average*, and *high*, according to the degree course distribution. Variable *irregularity* is positively associated with delays in passing exams, changes in degree programmes and other misalignments between the official exam schedule of the degree course and the actual student's curriculum.

All ordinal trinomial variables have been computed according to course-level statistics, in order to account for course-level heterogeneity. In particular, for each master's course the 33-th and the 67-th percentiles have been computed, and covariates' values have been classified according to their course-level percentiles.

The students' social class (classified as *working class*, *white-collar workers*, *lower middle class*, *upper middle class* and *unclassified* when it is not possible to determine graduates' social class) has also been included as a control variable.

Two models are fitted. In the former, the graduates' employment situation one year after graduation (unordered categorical variable) is used as response variable, distinguishing between those who are employed (have a paid job), those who are unemployed but are looking for a job, and those who are undertaking an apprenticeship. The latter model takes the final degree score (ordered categorical variable) as response variable.

The analysis is based on a dataset resulting from the linkage between responses to the yearly survey carried out by the Italian intercollegiate consortium Almalaurea (`www.almalaurea.it`) and the administrative data of the University of Trento.

The graduates' employment situation in the population is the following: 2580 employed (79%), 274 in a non-paid training work placement (8%), and 428 unemployed (13%). The coefficient estimates of the regression model are reported in Table 1.

As Table 1 shows, both the coefficients of the dummy variable *female* are positive and statistically significant at 1% level, and this seems to suggest that females face a lower probability of being employed one year after graduation than male graduates. Moreover, since the magnitude of the coefficient of category *unemployment* is larger than the coefficient

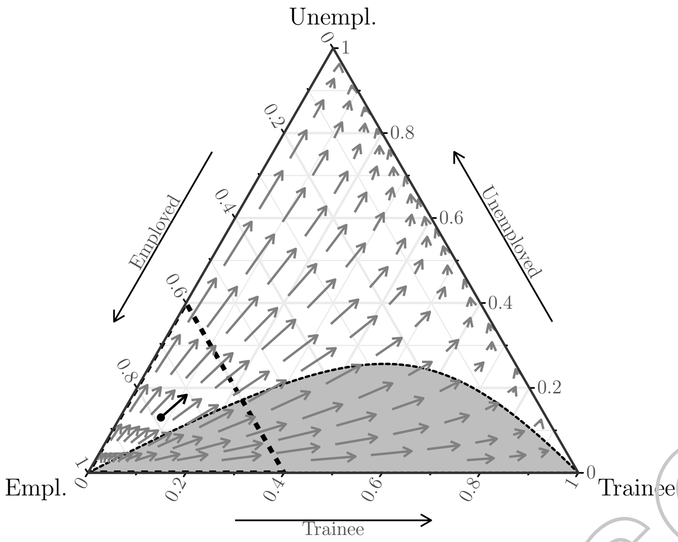

Figure 4: Effect of the dummy variable *female* ($\Delta = \hat{e}_2 \in \mathbb{R}^{10}$) on the probability distribution of the employment situation of master's graduates one year after graduation. The black arrow refers to the effect of the covariate *female* on the probability distribution of the population; the grey-coloured area covers points where the probability to be in a trainee work experience increases more than the unemployment risk; the triangle in dashed line shows the area of the ternary plot where points of the dataset are distributed.

of *trainee*, it may be expected that the reduction in female employment probability results mainly in an increased unemployment risk.

Figure 4 shows that such an interpretation is not completely true, as there are regions of the ternary plot (that is, regions of the parameter space Θ) where the covariate *female* mainly increases the probability of being in a training work experience instead of facing a higher unemployment risk. This happens when

$$\pi^{(3)}_{(x)} \leq \frac{1}{2} - \frac{1}{2}\,\frac{\beta^{(2)}_k - \beta^{(3)}_k}{\beta^{(3)}_k}\,\pi^{(2)}_{(x)} - \frac{1}{2}\sqrt{\left[1 - \frac{\beta^{(2)}_k - \beta^{(3)}_k}{\beta^{(3)}_k}\,\pi^{(2)}_{(x)}\right]^2 - 4\,\frac{\beta^{(2)}_k}{\beta^{(3)}_k}\,\pi^{(2)}_{(x)}\,(1 - \pi^{(2)}_{(x)})} \quad (10)$$

where, according to the notation of the previous sections, $\pi^{(1)}_{(x)}$, $\pi^{(2)}_{(x)}$, and $\pi^{(3)}_{(x)}$ respectively represent the probabilities of being employed, of being a trainee, and of being unemployed. In this case $k = 2$, as we consider the effect of second regressor (*gender*), being $\Delta = [0, 1, 0, \ldots, 0]^{\mathrm{T}} \in \mathbb{R}^{10}$.

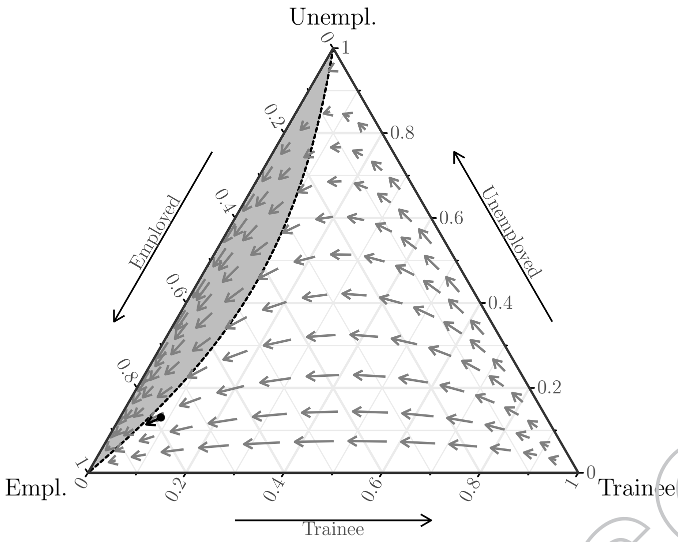

Figure 5: Effect of the dummy variable *high final score* ($\Delta = \hat{e}_4 \in \mathbb{R}^{10}$) on the probability distribution of the employment situation of master's graduates one year after graduation. The black arrow refers to the effect of the covariate *high final score* on the probability distribution of the population; the grey-coloured area covers points where the unemployment risk decreases more than the probability of being in a trainee work placement.

Inequality (10) identifies a convex set over Θ which is drawn in Figure 4 as a grey-coloured area. As it can be noted, region (10) represents a remarkable portion of the parameter space Θ, hence a sound interpretation of regression coefficients should not ignore its existence.

The inhomogeneous effect of the covariate *female* over the parameter space Θ derives from the modelling approach, and in particular from Equation (2), where the effects of covariates are included in terms of relative change (log-transform) within the relative probabilities (ratio between two probabilities).

A further issue regarding Figure 4 is worth noting. The dashed line triangle circumscribes the region where the points of the dataset are distributed, that is, the region of the parameter space of the dependent variable where the multi-logit regression has been fitted. Logically it may sometimes be more meaningful to analyse only specific area or even a smaller region if this results in a clearer graph (see e.g. Figures 7 and 8).

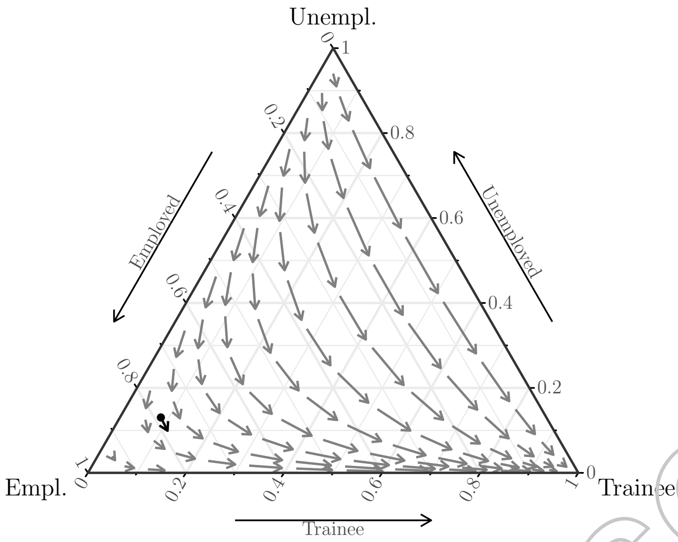

Figure 6: Effect of the dummy variable *short duration* ($\Delta = e_5 \in \mathbb{R}^{10}$) on the probability distribution of the employment situation of master's graduates one year after graduation. The black arrow refers to the effect of the covariate *short duration* on the probability distribution of the population.

Table 1 suggests that the final score has an asymmetric effect on the dependent variable: when it is lower than the average there is no evidence of a statistically significant effect, whereas a final score higher than the average is associated to a lower probability of taking up work placement. This second effect is particularly interesting and Figure 5 sums this up.

In addition, Figure 5 shows that an *high final score* mainly affects the probability of being employed rather than being unemployed. This is true over the whole parameter space Θ except for a (relatively) small region where the probability of accepting work placement is particularly low (see the grey-coloured area in Figure 5), suggesting that high final scores are mainly associated with low unemployment risk only when there are few opportunities to take a training work placement.

Duration affects the employment situation of master's graduates only when it is shorter than the average. In such cases, the unemployment risk diminishes, whereas the probability of taking up a work placement increases. Although both covariance coefficients are close to

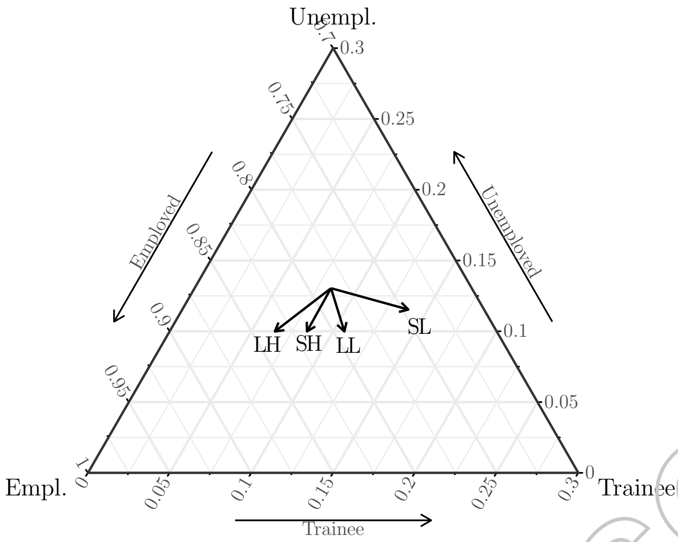

Figure 7: Comparison between four combined effects of dummy variables on duration and final degree score for the probability distribution in the population. Arrows represents the effects of *short duration* and *high final score* (SH, $\Delta = \hat{e}_5 + \hat{e}_4 \in \mathbb{R}^{10}$), *long duration* and *high final score* (LH, $\Delta = \hat{e}_6 + \hat{e}_4 \in \mathbb{R}^{10}$), *short duration* and *low final score* (SL, $\Delta = \hat{e}_5 + \hat{e}_3 \in \mathbb{R}^{10}$), *long duration* and *low final score* (LL, $\Delta = \hat{e}_6 + \hat{e}_3 \in \mathbb{R}^{10}$). Note that only a portion of the parametric space Θ has been represented.

0.3 in absolute value, the effect of the dummy variable *short duration* on the employment situation of master's graduates is rather small when the probability of being employed exceeds 0.7 (see Figure 6).

As discussed in Section 4, it is possible to analyse the effect of linear combinations of multiple covariates. This permits us to verify whether there is a trade-off between the final degree score and the swiftness of master's graduates in concluding their studies. Figure 7 shows the combined effects of the dummies on duration and final degree score; the four arrows showing how the population probability distribution changes when a master's student graduates in a short time with a high final grade (SH, $\Delta = \hat{e}_5 + \hat{e}_4 \in \mathbb{R}^{10}$), a long time with a high final grade (LH, $\Delta = \hat{e}_6 + \hat{e}_4 \in \mathbb{R}^{10}$), a short time with a low final grade (SL, $\Delta = \hat{e}_5 + \hat{e}_3 \in \mathbb{R}^{10}$), and a long time with a low final grade (LL, $\Delta = \hat{e}_6 + \hat{e}_3 \in \mathbb{R}^{10}$). The population values on employment rate, percentage of trainees, and unemployment rate

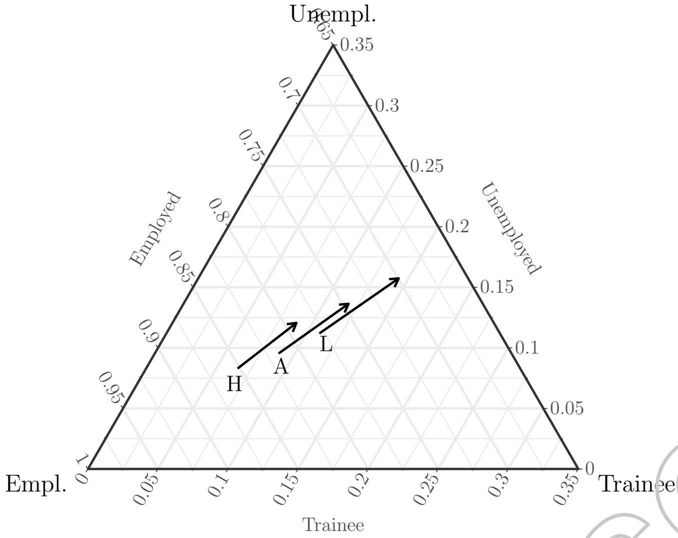

Figure 8: Effect of the dummy variable *female* ($\Delta = \hat{e}_2 \in \mathbb{R}^{10}$) on the probability distribution of the employment situation for three different profiles of working class male graduates who completed their studies earlier than their peers (*short duration*) and achieved different final degree scores: low (L), average (A) or high (H). Note that only a portion of the parametric space Θ has been represented.

are taken as a reference point.

Figure 7 suggests that there is no trade-off between duration and final degree score, as master's graduates who has completed their studies in a short time with a low final score (SL) face a higher unemployment risk, and a higher probability of undertaking a non-paid work placement than masters students graduating after a longer time with a higher final score (LH).

Moreover, Figure 7 suggests that the employment situation of master's graduates is related to the final degree score rather than duration of the studies (compare the lengths of segments LH-SH and LL-SL to lengths of LH-LL and SH-SL).

In line with the method illustrated regarding changes on a single covariate, it is possible to analyse the behaviour of any vector illustrated in Figure 7 (and also any difference amongst them) over the parameter space Θ.

A further possibility emerges by comparing the effect of (one or more) covariates amongst a group of individual profiles, i.e. specific configurations of covariates' values. An example, here, is seen in Figure 8, where the gender effect is represented for three different profiles of working class male graduates who completed their studies earlier than their peers (*short duration*). The profiles differ due to their final degree score, which may be low, average or high (respectively indicated as L, A, H).

As discussion on Figure 4 has already pointed out, the effect of gender on the employment situation is noteworthy and statistically significant, and the arrows in Figure 8 comply with that evidence. However, Figure 8 permits also to explicitly compare three different profiles, and to notice that the gender effect is large enough to make the employment situation of females graduated with a high final score (profile H) pretty similar to the employment situation of male graduates with a low final score (profile L).

The final degree score has been analysed as a response variable by means of an ordered multi-logit model. The final degree score has been categorised as *low*, *average* and *high* according to the course distribution; the distribution of the dependent variable in the population is: 987 low (30%), 1220 average (37%), and 1075 high (33%). The results of the regression are reported in Table 2.

As Table 2 shows, feminine gender is positively correlated with the final degree score, and Figure 9 represents the estimated relation between these two variables.

The interpretation of ternary plots does not change in case of ordered response variables. The only difference concerns the direction of arrows which always move either from the first to the last category or vice versa. This fact can be noted also in Figure 9 where the direction of arrows reflects the sign of the regression coefficient of the dummy variable on gender.

6 Conclusions

This paper proposes a graphical method for representing the effect of covariates (either singularly or jointly considered) on the conditional probability distribution in multinomial logit models where the response variable is an ordered or unordered categorical variable with three possible outcomes (trinomial). Unlike other graphical methods aiming at studying the effect of covariates in multinomial models (see e.g. Agresti, 2002), the method proposed in this paper permits to directly represent the variation in the probability distribution of the

Table 2: Multi-logit estimates (standard errors in parenthesis; p-values in brackets) for the final degree score of University of Trento's master's graduates. The response variable takes values *Low*, *Average*, and *High*. The reference master's graduate implied by regressors is a *male* with *average* final grade, *average* duration, belonging to the *working class*.

	Dependent variable:		
	Final score		
Female	0.244	(0.068)	[0.000]
Low irregularity	0.470	(0.090)	[0.000]
High irregularity	−0.371	(0.087)	[0.000]
Short duration	0.081	(0.087)	[0.351]
Long duration	−0.276	(0.089)	[0.002]
High school score	0.054	(0.003)	[0.000]
White-collar workers	0.095	(0.090)	[0.290]
Lower middle class	0.067	(0.098)	[0.494]
Upper middle class	0.128	(0.100)	[0.203]
Unclassified	0.050	(0.266)	[0.851]

dependent variable over the whole space of trinomial distribution functions. Furthermore, the graphical approach proposed is flexible enough to additionally analyse also the effect of combined and non-linear covariate effects using standard and easy-to-implement statistical tools.

References

Agresti, A. (2002). *Categorical Data Analysis* (2 ed.). Wiley.

Bancroft, W. D. (1897). A triangular diagram. *Journal of Physical Chemistry 1*, 403–10.

Begg, C. B. and R. Gray (1984). Calculation of polytomous logistic regression parameters using individualized regressions. *Biometrika 71*, 11–18.

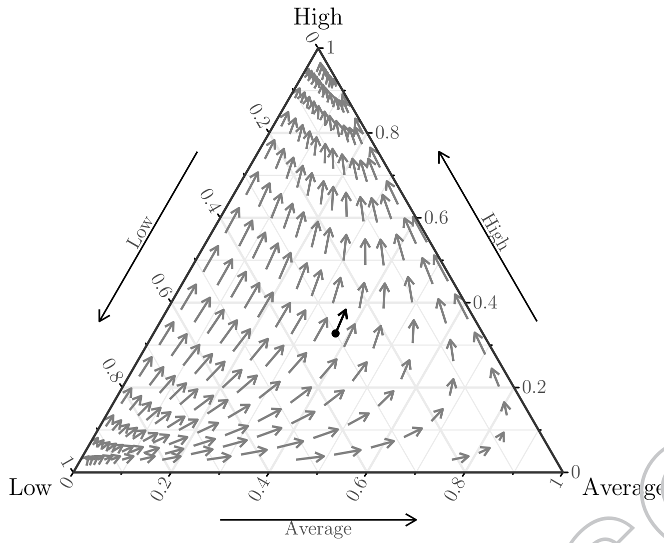

Figure 9: Effect of the dummy variable *female* ($\Delta = \hat{e}_2 \in \mathbb{R}^{10}$) on the probability distribution of the employment situation of master's graduates one year after graduation. The black arrow refers to the effect of the covariate *female* on the probability distribution of the population.

Fahrmeir, L. and G. Tutz (1994). *Multivariate Statistical Modelling Based on Generalized Linear Models*. Springer-Verlag.

Graff, J. F. (2017). Correspondence Analysis in Ternary Diagrams – A Graphical Approach to Improving the Validity of the Interpretation of Point Clouds. *Advances in Economics and Business 5*(9), 500–509.

Hennig, C., M. Meila, F. Murtagh, and R. Rocci (Eds.) (2016). *Handbook of Cluster Analysis*. CRC Press.

Ibragimov, N. H. (2010). *A Practical Course in Differential Equations and Mathematical Modelling*. Higher Education Press, World Scientific.

Johnson, N. L., A. W. Kemp, and S. Kotz (2005). *Univariate Discrete Distributions* (3 ed.). Wiley.

Lehmann, E. L. and G. Casella (1998). *Theory of Point Estimation* (2 ed.). Springer.

Lipovetsky, S. (2011). Conditional and multinomial logits as binary logit regressions. *Advances in Adaptive Data Analysis 3*, 309–324.

Lipovetsky, S. (2015). Trinomial response modeling in one logit regression. *Annals of Data Science 2*(2), 157–163.

Madsen, H. and P. Thyregod (2010). *Introduction to General and Generalized Linear Models*. CRC Press.

Wooldridge, J. M. (2008). *Introductory Econometrics: A Modern Approach* (4 ed.). South-Western College Publishing.

ACCEPTED MANUSCRIPT