Statistica Sinica Preprint No: SS-2016-0325.R2

Title	Discrete Choice Models for Nonmonotone Nonignorable Missing Data: Identification and Inference
Manuscript ID	SS-2016-0325.R2
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202016.0325
Complete List of Authors	Eric J. Tchetgen Tchetgen Linbo Wang and BaoLuo Sun
Corresponding Author	Eric J. Tchetgen Tchetgen
E-mail	etchetgen@gmail.com
Notice: Accepted version subject to English editing.	

CDF of (R, L).

Lemma 1 clarifies what the identification task entails, because under assumption (1), $f(L|R=1)$ is just-identified, and therefore $f(R, L)$ is nonparametrically just-identified only if one can just-identify $\text{Odds}_r(L)$ for all r. Below we describe a sufficient condition for identification under the discrete choice model of the nonresponse process .

3 Identification

3.1 The discrete choice nonresponse model

The DCM associates with each realized nonresponse patt[illegible] underlying utility function $U_r = \mu_r(L) + \varepsilon_r$, where $\{\varepsilon_r : r\}$ are i.i.d. with cumulative distribution function F_ε, and $\mu_r(L)$ encodes the dependence of a person's utility on L (McFa[illegible], Train, 2009). Some common choices of F_ε include the extreme value distribution (further discussed below) and the normal distribution, although in principle any CDF [illegible] specified. It is then assumed that a person's observed response pattern maximizes her utility, that is $R = \arg\max_r \{U_r : r\}$. Together, these assumptions imply that for each [illegible]

$$\Pi_r = \pi_r(L) = [illegible] = \int \prod_{s \neq r} F_\varepsilon(\Delta\mu_{rs}(L) + \varepsilon)\, dF_\varepsilon(\varepsilon), \tag{2}$$

where $\Delta\mu_{rs}(L) = \mu_r(L) - \mu_s(L)$ captures the dependence on L of a difference in utility in comparing [illegible]on's choice between nonresponse patterns r and s, see Train (2009). The integral in (2) is generally not available in closed form for most choices of F_ε (with the notable exception of the extreme value distribution, see Section 2.2), but can easily be evaluated by numerical integration using say, Gaussian quadrature. Two interesting observations about equation (2) are worth noting. Although not immediately apparent from the expression in the display, equation (2) gives rise to a proper probability mass function, that is $\sum_r \pi_r(l) = 1$ for all values of l and for any choice of F_ε. This remarkable result is a direct consequence of utility maximization as a formal principle for generating multinomial probabilities $\{\pi_r : r\}$. A second interesting observation is that only differ-