Nonparametric identification of static multinomial choice models

Otero, Karina V.

The Pennsylvania State University

2016

Nonparametric identification of static multinomial choice models

Karina Otero*

May 2016

Abstract

This paper proposes a new nonparametric identification strategy for static multiple choice models with random heterogeneity in unobservables. The strategy relies on functional properties of the sub-utilities and the distribution of the unobservables, a known payoff function for the "outside option" and exclusion restrictions for all but one alternative. This new strategy does not transform the multiple choice model into a set of binary models, does not need "special regressors", additive separability on observables or differentiability conditions. Some ideas for this new identification strategy are borrowed from Theorem 2 in Matzkin (1993) that intends to identify all the sub-utility functions but one and also the distribution of the shocks in differences. However, the proof of this published theorem is incorrect and (to the best of my knowledge) this paper is the first literature pointing this out and providing a new proof of a different version of the theorem after modifications of its assumptions.

Keywords: Nonparametric identification, Markov decision processes, discrete choice.

JEL Classification: C14, C35, C51.

I Introduction

Discrete choice models have been broadly applied in economics, some parametric examples that involve models of this type are in Berry, Levinsohn, and Pakes (1995), Chintagunta and Honore (1996), McFadden (1973) and Nevo (2001). These type of models are particularly useful when there is a finite number of mutually-exclusive alternatives and economic agents (or nature) pick each one with certain probability, depending on the attributes of each alternative. This paper focuses on nonparametric static multiple choice models with choice-specific and individual random heterogeneity in unobservables. The identification strategies for this type of models allow to uniquely determine the unknown objects using observations of the choices, features of the alternatives and characteristics of the agents. These strategies can be designed

Comments are welcome at karnv.o@gmail.com.

*Ph.D. in Economics, Pennsylvania State University. I am grateful to Andres Aradillas-Lopez for his valuable observations. All errors are my own.

to learn about unknown objects under different assumptions. The solution of a particular identification problem guarantees necessary uniqueness conditions to estimate the objects consistently. In this paper, there are not assumptions on the parametric functional forms of both, the utility functions and the joint distribution of the unobservables, i.e. the "shocks".

One of the earliest works that relax parametric assumptions for the distribution of the unobservables of this type of models is Manski (1975). An identification strategy for similar multinomial models known as "special regressor" was proposed by Lewbel (2000). In Lewbel's paper the utilities are assumed to be additive separable in particular observable terms, the "special regressors", which must be all distinct in the sense that none of them can be a deterministic function of some other. It is also assumed that the coefficients of these special regressors are the same and the identification of the distribution of the shocks is achieved in the limit. This is possible because when the special regressors become infinitely large for all alternatives except two, the multiple choice model can be reduced to a binary choice model. Briesch, Chintagunta, and Matzkin (2010) develop extensions of the model in McFadden (1973) and nonparametrically identify utility functions under assumptions of independence between the observable and unobservable variables and arguments based on the "special regressors" strategy. Berry and Haile (2009) study the identification of the joint distribution of indirect utilities under large support conditions, an additive separability restriction on preferences[1] and monotonicity restrictions.[2] Chiappori and Komunjer (2009) propose an identification strategy that departs from the "special regressor" assumption and is based on exclusion restrictions. The strategy requires differentiability conditions such that when a group of regressors are excluded from relevant conditional distribution functions, a number of restrictions on the partial derivatives of these functions can be derived. In particular, this method requires twice continuously differentiable sub-utility functions with everywhere non vanishing partial derivative with respect to observables that satisfy the exclusion restrictions. In addition, the second order partial derivatives of the conditional density of the unobserved choice-specific characteristics must be linearly independent.[3][4]

This paper nonparametrically identifies the sub-utility functions and the distribution of the unobserved random heterogeneity under some conditions weaker than the assumptions in the literature. The strategy relies on a strictly increasing distribution of the shocks in differences, continuity of the sub-utility functions and availability of at least 1 exclusion restriction for all but 1 alternative. Moreover, it is assumed that one of the choice-specific "sub-utility" or "payoff" functions, the one for the "outside option", is known and its range is the real line. This is a noticeable difference with respect to the identification strategies available so far and allows to use information that could be available for the "outside option", for instance this function could be estimated in advance. Thus, the standard assumption that the utility for the "outside option" is 0 under any circumstances is not compatible. Note that this new identification strategy does not transform the multiple choice model into a set of binary choices, does not need "special regressors", additive separability on observables or differentiability conditions.

Some ideas for this new identification strategy are borrowed from Theorem 2 in Matzkin (1993), T2M, which intends to identify all the sub-utility functions but one and the distribution of the shocks in

[1]for a group of observables and unobservables.

[2]A broader review on identification strategies can be found in Matzkin (2013).

[3]This requirement rules out the normal distribution.

[4]Other assumptions in Chiappori and Komunjer (2009) include location assumptions on the sub-utilities and their derivatives over certain subsets, and unobserved choice-specific characteristics with 0 expectations conditional on the available instruments.

differences under assumptions on their functional properties, range and level attained over certain subsets. The T2M is an interesting example of ideas for a nonparametric identification strategy, but its proof is incorrect and (to the best of my knowledge) this paper is the first literature pointing this out and providing a proof of a different version of the T2M after modifications to its assumptions.

The next section presents a stylized version of the multiple choice model and the Theorem 2 as stated in Matzkin (1993). The following section revise the assumptions and prove a new theorem to achieve the identification of both, the sub-utility functions and the distribution of the unobserved random heterogeneity in differences. The last section summarizes the conclusions.

II The Model

Consider a setup in which agent i has a vector $s_i \in S \subset \mathbb{R}^L$ of socioeconomic characteristics and chooses an alternative j among the $J < \infty$ available alternatives.[5] Each alternative $j \in A = \{1, .., J\}$ provides certain level of utility to the agent depending on a vector of observable attributes, $z_{ij} \in Z_j \subset \mathbb{R}^{K_j}$, which may be different for each agent. The utility level attained by agent i is

$$U_i(j) = V^*(j, s_i, z_{ij}) + \varepsilon_{ij}$$

with unobservable random utility ε_{ij} and an unknown real function $V^*(j, \cdot) : (S \times Z_j) \to \mathbb{R}$ common to all agents. The subindex i is omitted now to simplify the notation given that the setup is the same for all agents.

Assume that the distribution of the unobservable random vector $(\varepsilon_1, ..., \varepsilon_J)$ is independent of the vector of observable characteristics (s, z) and let $P_j^*(\cdot)$ be the cumulative distribution of $\varepsilon^{(j)} = (\varepsilon_1 - \varepsilon_j, \dots, \varepsilon_J - \varepsilon_j) \in \mathbb{R}^{J-1}$. Then, the agent with characteristics s chooses alternative $j \in A$ if

$$V^*(j, s, z_j) + \varepsilon_j > V^*(k, s, z_k) + \varepsilon_k \text{ for } k \neq j$$

The J-dimensional vector $z = (z_1, \dots, z_J)$ denotes attributes of the J alternatives, with $z \in Z = \prod_{j=1}^{J} Z_j$. Let $V^*(s, z) = (V^*(1, s, z_1), \dots, V^*(J, s, z_J))$, such that $V^* : (S \times Z) \to \mathbb{R}^J$. For each $j \in A$, denote

$$V^{*(j)}(s, z) = (V^*(j, s, z_j) - V^*(1, s, z_1), \dots, V^*(j, s, z_j) - V^*(J, s, z_J)) \in \mathbb{R}^{J-1},$$

and the probability of choosing alternative j under the independence assumption becomes $P_j^*\left(V^{*(j)}(s, z)\right)$. Let $\Pr(j|(s, z))$ be the observable probability that an agent with socioeconomic characteristics s choose alternative $j \in A$ when the attributes are z. Then, equation (1) holds.

$$\Pr(j|(s, z)) = P_j^*\left(V^{*(j)}(s, z)\right). \tag{1}$$

Denote by P^* the set of distributions $\left\{ P_j^*(\cdot) \middle| j \in A \right\}$. Let $\mathbf{V}$ denote a set of functions $V : (S \times Z) \to \mathbb{R}^J$, such that the domain of the jth coordinate of V, i.e. $V(j, \cdot)$, is $S \times Z_j$. Let $\mathbf{P}$ denote a set

[5] This section follows the notation in Matzkin (1993), but note that an additional dimension such as markets could be incorporated.

whose generic element is a collection $\left\{ P_j^* (\cdot) \middle| j \in A \right\}$ of distribution functions of $\varepsilon^{(j)}$. Assume that $(V^*, P^*) \in (\mathbf{V} \times \mathbf{P})$. Denote $V = V^*$, if for all $j \in A$, $V(j, s, z_j) = V^*(j, s, z_j)$ a.s. with respect to the probability measure of (s, z), denoted by G. In addition, $P = P^*$ whenever for all $j \in A$, $P_j(\alpha \,|\, (s, z)) = P_j^*(\alpha \,|\, (s, z))$ a.e. with respect to the Lebesgue measure in $\mathbb{R}^{J-1}$ and a.s. with respect to the probability measure G. The identification result is obtained when the pair (V^*, P^*) can be uniquely recovered from the conditional choice probabilities $\Pr(j | (s, z))$ and other observables.

Definition 1. (V^*, P^*) is identified in $(\mathbf{V} \times \mathbf{P})$ if for all $(V, P) \neq (V^*, P^*)$ there is some set $D \subset (S \times Z)$ such that $G(D) > 0$ and for some alternative j and all $(s, z) \in D$: $\Pr(j | (s, z); V, P) \neq \Pr(j | (s, z); V^*, P^*)$.

Theorem 2 in Matzkin (1993) intends to provide sufficient conditions for the identification in definition 1. However, its proof is not correct because intuitively, the combination of location assumptions and exclusion restrictions do not allow for sufficient variability to actually identify P^*. To see this formally, theorem 1 presents the Theorem 2 as stated in Matzkin (1993) followed by the details of the gap in its proof.

Theorem 1 (Theorem 2, Matzkin, 1993)**.** *Suppose that the following assumptions are satisfied:*

Assumption A. *For all $V \in \mathbf{V}$, V is a continuous function.*

Assumption B. *There exists a vector of socioeconomic characteristics $\bar{s} \in S$, a subset of attributes $\bar{Z}_1 \subset Z_1$, and some $\bar{z}_j \in Z_j$ for $j = 2, \ldots, J$ such that*

i. *for all $V \in \mathbf{V}$, $V(1, \bar{s}, \bar{Z}_1) = \mathbb{R}$.*
ii. *for all V, $V' \in \mathbf{V}$ and all $z_1 \in \bar{Z}_1$, $V(1, \bar{s}, z_1) = V'(1, \bar{s}, z_1)$,*
iii. *for all $V \in \mathbf{V}$ and all $j = 2, \ldots, J$, $V(j, \bar{s}, \bar{z}_j) = 0$.*

Assumption C. *There exists $z_1^* \in Z_1$ and $\gamma \in \mathbb{R}$ such that for all $s \in S$ and all $V \in \mathbf{V}$, $V(1, s, z_1^*) = \gamma$.*

Assumption D. *For all $P \in \mathbf{P}$ and all $j \in A$ there exists a function $P_j : \mathbb{R}^{J-1} \to \mathbb{R}$ such that for all $(s, z) \in (S \times Z)$, $P_j(\cdot | (s, z)) = P_j(\cdot)$.*

Assumption E. *For all $P \in \mathbf{P}$ and all $j \in A$, P_j is continuous.*

Assumption F. *For all $P \in \mathbf{P}$ and all $j \in A$, P_j is strictly increasing.*

Assumption G. *The support of G is $S \times Z$.*

Then, (V^*, P^*) is identified within $(\mathbf{V} \times \mathbf{P})$

Section III proposes a modification to the assumptions in Theorem 1 and provides a complete proof. But before moving forward, it is relevant to point out why the arguments in the proof of Theorem 1 developed in Matzkin (1993) are not sufficient to actually prove this theorem.

The main arguments to prove theorem 1 in Matzkin (1993) are as follows. Suppose that $(V, P) \in (\mathbf{V} \times \mathbf{P})$ and for all $j \in A$,

$$P_j(j | (s, z); V, P) = P_j(j | (s, z); V^*, P^*) \text{ a.s.} (G).$$

Then by assumption D,
$$P_j\left(V^{(j)}(s,z)\right) = P_j^*\left(V^{*(j)}(s,z)\right) \text{ a.s.} (G).$$
Since P_j, P_j^*, V and V^* are continuous, it follows by Assumption G that
$$P_j\left(V^{(j)}(s,z)\right) = P_j^*\left(V^{*(j)}(s,z)\right) \ \forall (s,z) \in (S \times Z).$$
In particular,
$$P_1\left(V^{(1)}(s,z)\right) = P_1^*\left(V^{*(1)}(s,z)\right) \ \forall (s,z) \in \{\bar{s}\} \times \bar{Z}_1 \times \prod_{j=2}^{J} \{\bar{z}_j\}. \tag{2}$$

By assumption B item iii, recall that for all $V \in \mathbf{V}$ and all $j = 2, \ldots, J$, $V(j, \bar{s}, \bar{z}_j) = 0$. Therefore:
$$\begin{aligned} V^{(1)}(\bar{s}, z_1, \bar{z}_2, \ldots, \bar{z}_J) &= (V(1,\bar{s},z_1) - V(2,\bar{s},\bar{z}_2), \ldots, V(1,\bar{s},z_1) - V(J,\bar{s},\bar{z}_J)) \\ &= (V(1,\bar{s},z_1), \ldots, V(1,\bar{s},z_1)) \end{aligned}$$
Let $\alpha = (\alpha_1, \ldots, \alpha_{J-1}) \in \mathbb{R}^{J-1}$. By assumption B item ii and provided that $(V^*, P^*) \in (\mathbf{V} \times \mathbf{P})$ for all $z_1 \in \bar{Z}_1$,
$$V^{(1)}(\bar{s}, z_1, \bar{z}_2, \ldots, \bar{z}_J) = V^{*(1)}(\bar{s}, z_1, \bar{z}_2, \ldots, \bar{z}_J) = \alpha^*.$$
where $\alpha^* = (\alpha_1^*, \ldots, \alpha_1^*)$. Hence, by (2),
$$P_1(\alpha^*) = P_1\left(V^{(1)}(\bar{s}, z_1, \bar{z}_2, \ldots, \bar{z}_J)\right) = P_1^*\left(V^{*(1)}(\bar{s}, z_1, \bar{z}_2, \ldots, \bar{z}_J)\right) = P_1^*(\alpha^*).$$
Then Matzkin (1993) p. 160 concludes that $P_1 = P_1^*$, but this cannot be always true because α^* does not span the whole space. If it were true that $P_1 = P_1^*$ then the proof could continue pointing out that for all $j \in A$, $\varepsilon^{(j)}$ is a linear transformation of $\varepsilon^{(1)}$, thus P_j would be equal to P_j^* for all $j \in A$ and therefore $P = P^*$. For the final arguments, see Matzkin (1993), p. 160.

It follows from assumption B items ii and iii that α^* is not a typical general vector $\alpha \in \mathbb{R}$, so it is not possible to identify the whole P_1^* and therefore P^* remains unknown. Section III proposes an alternative theorem and develops its proof.

III Revision of assumptions in Theorem 1 and nonparametric identification

Consider the following modifications of the assumptions in section II. Each alternative has its own corresponding attributes such that after considering them, the attributes of the other $j \neq 1$ alternatives are irrelevant. Alternative $j = 1$ can be interpreted as an "outside option" or an alternative of reference for which there is more available information. The attributes of alternative $j = 1$ are allowed to directly affect the utility levels attained by other alternatives. Note that the introduction of exclusion restrictions for $j > 1$ is indeed a generalization of the original setup and the proof of theorem 2 also holds under the original one, with full exclusion restrictions for all $j \geq 1$. The attributes can be correlated but they are not allowed to be deterministic functions of others. Assumption H redefines the sub-utility functions according to the updated assumptions.

Assumption H. *For all $j > 1$, $V(j, \cdot) : S \times Z_1 \times Z_j \to \mathbb{R}$ and for $j = 1$, $V(1, \cdot) : S \times Z_1 \to \mathbb{R}$.*

There is some combination of values for the socioeconomic characteristics and the attributes of alternatives $j \neq 1$ under which the sub-utility functions are 0 within a subset, independently of the values attained by the attributes of the outside option. Besides, under particular socioeconomic characteristics and for each given value of the attributes of the outside option within a subset, the range of the sub-utility functions is the real line. Assumption I replaces assumption B to incorporate these changes.

Assumption I. *There exists a vector of socioeconomic characteristics $\bar{s} \in S$, a subset of attributes $\bar{Z}_1 \subset Z_1$, and observable attributes $\bar{z}_j \in Z_j$ for all $(j = 2, \dots, J)$ such that*

i. *for all $V \in \mathbf{V}$, $V\left(1, \bar{s}, \bar{Z}_1\right) = \mathbb{R}$.*
ii. *for all V, $V' \in \mathbf{V}$ and all $z_1 \in \bar{Z}_1$, $V\left(1, \bar{s}, z_1\right) = V'\left(1, \bar{s}, z_1\right)$,*
iii. *for all $V \in \mathbf{V}$ and all $j = 2, \dots, J$,$V\left(j, \bar{s}, \bar{Z}_1, \bar{z}_j\right) = 0$.*
iv. *for all $V \in \mathbf{V}$, all $z_1 \in \bar{Z}_1$ and all $j = 2, \dots, J$, $V\left(j, \bar{s}, z_1, Z_j\right) = \mathbb{R}$*

The sub-utility function for the outside option is known over different subsets and cannot be 0 everywhere by assumption I item i. Assumption J replaces assumption C which is usefult to identify the sub-utilities in levels on the whole space.

Assumption J. *For all $(s, z_1) \notin \{\bar{s}\} \times \bar{Z}_1$, and all V, $V' \in \mathbf{V}$, $V\left(1, s, z_1\right) = V'\left(1, s, z_1\right)$.*

Theorem 2 states a revised version of theorem 1.

Theorem 2. *Suppose that assumptions A and D to J are satisfied:*

Then, (V^*, P^*) is identified within $(\mathbf{V} \times \mathbf{P})$

To prove Theorem 2, first consider the following lemma that states sufficient conditions under which the arguments of a particular multivariate function are equal.

Lemma 1. *Let $P_j : \mathbb{R}^J \to \mathbb{R}$, $a, b_j \in \mathbb{R}$ for $j = 0, \dots, J$, $\overrightarrow{a}_0 = (a, \dots, a)'$, $\overrightarrow{a}_j = \overrightarrow{a}_1 = (-a, 0, \dots, 0)'$ for $j > 0$, and*

$$\overrightarrow{b}_j = (b_j - b_0, \dots, b_j - b_{j-1}, b_j - b_{j+1}, \dots, b_j - b_J)',$$

for any $0 < j < J$, with the corresponding changes for $j = 0, J$.

Denote $(\lambda_1, \dots, \lambda_J) \gtrsim (\gamma_1, \dots, \gamma_J)$ when for all $j > 0$, $\lambda_j \geq \gamma_j$, and for at least a $k \in \{1, \dots, J\}$, $\lambda_k > \gamma_k$. Let $m \gtrsim m'$ imply that $P_j\left(m\right) \neq P_j(m')$ for any j. Then, $P_j(\overrightarrow{b}_j) = P_j\left(\overrightarrow{a}_j\right)$ for all $j = 0, \dots, J$, implies that $b_0 - b_j = a$ for all $j = 1, \dots, J$.

Proof. (Lemma 1)

Suppose by contradiction that $b_0 - b_j \neq a$ for some j and without loss of generality (wlog) let $j = 2$. If $\max_j(b_0 - b_j) \leq a$ and $b_0 - b_2 < a$ then $P_0(\overrightarrow{b}_0) \neq P_0\left(\overrightarrow{a}_0\right)$ by assumption. If $\max_j(b_0 - b_j) > a$ and wlog $b_0 - b_1 = \max_j(b_0 - b_j)$, then $P_1(\overrightarrow{b}_1) \neq P_1(\overrightarrow{a}_1)$ because $b_1 - b_j \leq 0$ for $j > 2$ and $b_1 - b_0 < -a$. Therefore, it cannot be true that $b_0 - b_2 \neq a$ when $P_j(\overrightarrow{b}_j) = P_j\left(\overrightarrow{a}_j\right)$ for all $j = 0, \dots, J$. ∎

To gain some intuition about the proof of Theorem 2, note that the sub-utility of the outside option is known over $\bar{s} \in S$ and a subset of attributes $\bar{Z}_1 \subset Z_1$. Then, when the attributes of alternatives $j > 1$

reach levels $\bar{z}_j \in Z_j$ for all $(j = 2, \ldots, J)$, assumption I item iii eliminates the uncertainty on V^*. This combination helps to identify P^* on the corresponding subespace. Assumption I item iv becomes useful to identify sub-utilities outside of $\prod_{j=2}^{J} \{\bar{z}_j\}$ because given an arbitrary unknown $V^*(k, \bar{s}, z_1, z_k)$ there is some $\{z_j\}_{j \neq 1,k}$ that place a proposed V on the subespace where P^* was identified. By lemma 1 we can argue that the arguments of the functions must be equal, and by assumption I item ii those arguments become known. Repeating the procedure it is possible to identify V^* on a larger subespace, i.e. $\{\bar{s}\} \times \bar{Z}_1 \times \prod_{j=2}^{J} Z_j$. This subespace is sufficient to span the domain of P, so P^* can be identified. The sub-utilities in differences are identified by invertibility of P^* and finally assumption J is useful to identify the sub-utilities in levels on the whole space.

Proof. (Theorem 2)

Suppose that $(V, P) \in (\mathbf{V} \times \mathbf{P})$ and for all $j \in A$,

$$P_j(j|(s,z); V, P) = P_j(j|(s,z); V^*, P^*) \text{ a.s.} (G).$$

Then by assumption D,

$$P_j\left(V^{(j)}(s,z)\right) = P_j^*\left(V^{*(j)}(s,z)\right) \text{ a.s.} (G).$$

Since P_j and P_j^* are continuous and V and V^* are continuous, it follows by assumption G that

$$P_j\left(V^{(j)}(s,z)\right) = P_j^*\left(V^{*(j)}(s,z)\right) \; \forall (s,z) \in (S \times Z).$$

In particular,

$$P_1\left(V^{(1)}(s,z)\right) = P_1^*\left(V^{*(1)}(s,z)\right) \; \forall (s,z) \in \{\bar{s}\} \times \bar{Z}_1 \times \prod_{j=2}^{J} \{\bar{z}_j\}. \tag{3}$$

Recall that by assumption I item iii we know that for all $V \in \mathbf{V}$ and all $j = 2, \ldots, J$, $V\left(j, \bar{s}, \bar{Z}_1, \bar{z}_j\right) = 0$. Therefore, for any $z_1 \in \bar{Z}_1$,

$$\begin{aligned} V^{(1)}(\bar{s}, z_1, \bar{z}_2, \ldots, \bar{z}_J) &= \left(V(1, \bar{s}, z_1) - V(2, \bar{s}, z_1, \bar{z}_2), \ldots, V(1, \bar{s}, z_1) - V(J, \bar{s}, z_1, \bar{z}_J)\right) \\ &= \left(V(1, \bar{s}, z_1), \ldots, V(1, \bar{s}, z_1)\right) \end{aligned}$$

Moreover, by assumption I items i to iii, $V\left(1, \bar{s}, \bar{Z}_1\right) = V^*\left(1, \bar{s}, \bar{Z}_1\right) = \mathbb{R}$ and for $j > 1$,

$$V^{(j)}\left(\bar{s}, \bar{Z}_1, \bar{z}_2, \ldots, \bar{z}_J\right) = \left(-V\left(1, \bar{s}, \bar{Z}_1\right), 0, \ldots, 0\right)$$

Consider first case 1, in which for all $j \in A$ and all $(z_1, \bar{z}_{-1}) \in \bar{Z}_1 \times \{\bar{z}_j\}_{j=2}^{J}$

$$\begin{aligned} P_j(j|(\bar{s}, z_1, \bar{z}_{-1}); V^*, P^*) &= P_j\left(V^{(j)}(\bar{s}, z_1, \bar{z}_{-1})\right) \\ &= P_j\left(V^{*(j)}(\bar{s}, z_1, \bar{z}_{-1})\right) \end{aligned}$$

implies that $P_j = P_j^*$, so P^* is identified on the space spanned by the sub-utility functions on $\{\bar{s}\} \times \bar{Z}_1 \times \prod_{j=2}^{J} \{\bar{z}_j\}$.

Consider now case 2, for all $j \in A$, $z_1 \in \bar{Z}_1$ and some $z_{-1} \notin \{\bar{z}_j\}_{j=2}^{J}$,

$$P_j\left(j|\left(\bar{s}, z_1, z_{-1}\right); V^*, P^*\right) = P_j\left(V^{(j)}\left(\bar{s}, z_1, z_{-1}\right)\right).$$

In case 2, w.l.o.g., first proceed identifying $V^*\left(2, \bar{s}, z_1, z_2\right)$ as follows. Given assumption I item iv, pick for all $j \neq 1, 2$ some $\hat{z}_j \in Z_j$, $\tilde{z}_1 \in \bar{Z}_1$ and $\bar{z}_{-1} \in \{\bar{z}_j\}_{j=2}^{J}$ such that for all $j \in A$

$$P_j\left(j|\left(\bar{s}, \tilde{z}_1, \bar{z}_{-1}\right); V^*, P^*\right) = P_j^*\left(V^{(j)}\left(\bar{s}, z_1, z_2, \hat{z}_3, \ldots, \hat{z}_J\right)\right).$$

Case 1 guarantee that $P_j\left(j|\left(\bar{s}, \tilde{z}_1, \bar{z}_{-1}\right); V^*, P^*\right) = P_j^*\left(V^{*(j)}\left(\bar{s}, \tilde{z}_1, \bar{z}_{-1}\right)\right)$ for all $j \in A$, so

$$P_j^*\left(V^{*(j)}\left(\bar{s}, \tilde{z}_1, \bar{z}_{-1}\right)\right) = P_j^*\left(V^{(j)}\left(\bar{s}, z_1, z_2, \hat{z}_3, \ldots, \hat{z}_J\right)\right). \tag{4}$$

By lemma 1, equation (4) implies $V^{*(j)}\left(\bar{s}, \tilde{z}_1, \bar{z}_{-1}\right) = V^{(j)}\left(\bar{s}, z_1, z_2, \hat{z}_3, \ldots, \hat{z}_J\right)$ for all $j \in A$, so $V\left(2, \bar{s}, z_1, z_2\right) = V^*\left(1, \bar{s}, z_1\right) - V^*\left(1, \bar{s}, \tilde{z}_1\right)$. Note that $V^*(1, \cdot)$ is identified under assumption I item ii, so $V^*\left(2, \bar{s}, z_1, z_2\right)$ is also identified for any arbitrary $z_1 \in \bar{Z}_1$ and $z_2 \neq \bar{z}_2$. Using the same argument, all levels $V^*(j, \cdot)$ are identified for arbitrary $z_1 \in \bar{Z}_1$ and $z_{-1} \notin \{\bar{z}_j\}_{j=2}^{J}$.

Combining case 1 and 2, $V^{*(j)}(\cdot)$ is identified for all j on $\{\bar{s}\} \times \bar{Z}_1 \times \prod_{j=2}^{J} Z_j$, and provided assumption I items i and iv, P^* is identified. Since P^* is invertible (by assumption F and lemma 1) and identified, then the mapping $V^{*(j)}(\cdot)$ is identified for all j and $(s, z) \in (S \times Z)$. Finally, by assumption J the levels, $V^*(j, \cdot)$ are also identified for all j and $(s, z) \in (S \times Z)$. ■

Under full exclusion restrictions, i.e. when the attributes of alternative $j = 1$ do not affect the utility levels of other alternatives $j > 1$, the analogue of theorem 2 holds and assumption J can be replaced by assumption C. Moreover, under full exclusion restrictions, if alternative $j = 1$ does not depend on s, $V^*(1, \cdot)$ can be identified on Z_1 after identifying P^* and $\{V_j^*(\cdot)\}_{j=2}^{J}$. Therefore, in this later case assumptions C and J are not needed.

IV Conclusions

This paper proves a new identification strategy for nonparametric static multiple choice models with choice-specific and individual random heterogeneity in unobservables. The literature has proposed some nonparametric identification strategies for static multiple choice models (see for instance Berry and Haile 2009, Briesch, Chintagunta, and Matzkin 2010, Chiappori and Komunjer 2009, Lewbel 2000 and Matzkin 2013) but the available methods rely on "special regressors" arguments, identification at infinity, parametric assumptions for the utility functions and/or differentiability conditions of the unknown objects.

The identification strategy proved in this paper allows for nonparametric functional forms of both, the utility functions and the joint distribution of the unobservables in differences. The strategy mainly relies on functional properties of the sub-utility functions, properties of the distribution of the unobservables and the availability of at least 1 exclusion restriction for all but 1 alternative. Besides, it is assumed that the choice-specific "payoff" function for the "outside option" is known. Thus, the standard assumption that the utility for the "outside option" is always 0 is ruled out.

Some ideas for this new identification strategy are borrowed from Theorem 2 in Matzkin (1993), which intends to identify all the sub-utility functions but one and the distribution of the shocks in differences. The proof of Theorem 2 in Matzkin (1993), is incorrect and (to the best of my knowledge) this paper is the first literature pointing this out and providing a proof of an alternative version of the T2M after modifications to its assumptions.

References

BERRY, S., AND P. HAILE (2009): "Identification of discrete choice demand from market-level data," *Manuscript, Yale University*.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, pp. 841–890.

BRIESCH, R. A., P. K. CHINTAGUNTA, AND R. L. MATZKIN (2010): "Nonparametric discrete choice models with unobserved heterogeneity," *Journal of Business & Economic Statistics*.

CHIAPPORI, P., AND I. KOMUNJER (2009): "On the nonparametric identification of multiple choice models," *Manuscript, Columbia University*.

CHINTAGUNTA, P. K., AND B. E. HONORE (1996): "Investigating the effects of marketing variables and unobserved heterogeneity in a multinomial probit model," *International Journal of Research in Marketing*, 13(1), 1–15.

LEWBEL, A. (2000): "Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables," *Journal of Econometrics*, 97(1), 145–177.

MANSKI, C. F. (1975): "Maximum score estimation of the stochastic utility model of choice," *Journal of econometrics*, 3(3), 205–228.

MATZKIN, R. L. (1993): "Nonparametric identification and estimation of polychotomous choice models," *Journal of Econometrics*, 58(1–2), 137–168.

——— (2013): "Nonparametric identification in structural economic models," *Annu. Rev. Econ.*, 5(1), 457–486.

MCFADDEN, D. (1973): "Conditional logit analysis of qualitative choice behavior," *Frontiers in Econometrics*, pp. 105–142, bibtex: mcfadden_conditional_1973.

NEVO, A. (2001): "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 69(2), 307–342.