A simple derivation of the Williams-Daly-Zachery theorem.

P O Lindberg

April 14, 2012

Abstract

In this note we give a simple elementary proof of the Williams-Daly-Zachery theorem for additive random utility (ARU) models. This theorem says that the choice probabilities in an ARU model are the partial derivatives of the expected achieved utility. This is proved under the condition that the choice probabilities are continuous. We further prove that the choice probabilities are continuous exactly when the probability of ties is zero. Finally we prove that the probability of ties is zero if the cdf of the random term (of the ARU model) is differentiable with locally bounded gradient.

1. Introduction.

The *Williams-Daly-Zachery* theorem is a result for *Additive Random Utility* (ARU) Models corresponding to Roy's Identity in consumer theory.

We consider a fairly standard setting for *Random Utility* (RU) Models (See e.g. Ben-Akiva and Lerman, 1985). We have a random vector $\mathbf{U} = (U_1, U_2, ..., U_n) \in \Re^n$ with joint cdf F_U. The U_i may be thought of as representing the (to the observer unknown) *utility* of *choice alternative i* for a randomly selected choice maker, and we are interested in the *choice probabilities*

$$\bar{p}_i =_{df} \Pr\{U_j \le U_i, j \ne i\}, \tag{1}$$

i.e. that alternative *i* delivers the maximum utility.

In an ARU model the utilities are given an additive structure:

$$U_i(\mathbf{v}) = v_i + X_i, \tag{2}$$

where the v_i are observable *population values* and the X_i are unobservable *individual values*, where the lack of knowledge is modeled by randomness.

Here we assume that these random terms $\mathbf{X} = (X_1, X_2, ..., X_n)$ have a joint c.d.f. *F*, such that the random variables X_i have finite expectations $\bar{X}_i =_{df} E(X_i)$.

ARU models have become workhorses in many areas of applied economics where discrete choices are considered, such as Choice of Residential Location (e.g. McFadden, 1978), Product Choice (e.g. McFadden, 1980), Labor Economics (e.g. Keane Todd and Wolpin, 2010), Transportation (e.g. Ben-Akiva and Lerman, 1985), and Product Differentiation (e.g. Anderson, de Palma and Thisse, 1992).

Let $\bar{U}(\mathbf{v}) =_{df} E(\max_i U_i(\mathbf{v})) = E(\max_i (v_i + X_i))$ be the *expected achieved utility* under given values the v_i, of a randomly selected, utility maximizing, decision maker. $\bar{U}(\mathbf{v})$ exists finitely, since the $\bar{X}_i$ do. Further let $\bar{P}_i(\mathbf{v}) = \Pr(v_i + X_i \le v_i + X_i, j \ne i)$, be the probability that alternative *i* delivers the maximum utility at **v**.

The *Williams-Daly-Zachery* (WDZ) "theorem" says that

$$\bar{P}_i(\mathbf{v}) = \frac{\partial \bar{U}(\mathbf{v})}{\partial v_i}. \tag{3}$$

It is an analogue in a discrete choice setting, to Roy's identity in consumer theory. It was termed the WDZ theorem in McFadden (1981), section 5.8. McFadden refers to Daly and Zachery (1978), and Williams (1977) as forerunners of the result. In McFadden (1981), the author gives a proof of the WDZ theorem assuming inter alias cdf's with densities for the random term, a proof he says is adapted after that of Daly and Zachary (1976). Anderson et al (1992), Ch. 3, gives an introduction to a "representative consumer" approach in a discrete choice setting. In particular they give a proof of the WDZ Thm. (their Lemma 3.2), assuming that the distribution of X has a density.

The WDZ theorem has been proved e.g. for generalizations of the GEV-models in (Lindberg, Erikson & Mattsson 1996), and for ARU models, where *F* has density by Fosgerau, Bierlaire and McFadden (2010).

In this note we give in section 2 a simple proof of the WDZ theorem under the condition that choice probabilities $\overline{P}_i(\mathbf{v})$ are continuous, and as a corollary, under the condition that *F* is continuously differentiable. In sec. 3, we review results on choice probabilities in RU models. These are applied o the ARU case in sec. 4. In particular we show that if the cdf of the random term is differentiable with locally bounded gradient, then the probability of ties is zero. Sec. 5 is devoted to the continuity of the choice probabilities w.r.t. change of one variable. In particular we show that the choice probabilities are continuous if there are no ties. In sec. 6 we show that the choice probabilities are continuous if the are continuous in each variable.

2. The Williams-Daly-Zachery theorem.

The theorem will be derived through a simple Lemma:

Lemma 1. Let $p_i =_{df} \overline{P}_i(\mathbf{v})$, and assume that v_i, the *i*-th population value changes from v_i to $v_i + \Delta v_i$, and let Δp_i be the corresponding change in $\overline{P}_i(\mathbf{v})$. Then, $\Delta \overline{U}$, the corresponding change in $\overline{U}(\mathbf{v})$, fulfils

$$p_i \Delta v_i \le \Delta \overline{U} \le (p_i + \Delta p_i)\Delta v_i \qquad (4)$$

Proof: Let us first consider the case $\Delta v_i > 0$, i.e. alternative *i* is improved, and divide the choice makers into groups according to their choices:

G_1, those who choose alternative *i* both before and hence also after the change of v_i,

G_c, those who change <u>to</u> alternative *i*, from some other alternative, and

G_2, those who don't choose alternative *i*, neither before nor after the change.

The expected changes in achieved utility for individuals in G_1 and G_2 (conditioned on being in that group) are Δv_i and 0 respectively. Let ΔU_c be the corresponding change in G_c.

The probability masses of these groups are respectively p_i, $\Delta p_i \ge 0$ and $1 - p_i - \Delta p_i$

Summing the effects on $\overline{U}(\mathbf{v})$ in the different groups we have

$$\Delta \overline{U} = p_i \Delta v_i + \Delta p_i \Delta U_c + 0. \qquad (5)$$

Further ΔU_c obviously fulfils $0 \le \Delta U_c \le \Delta v_i$. (6)

Thus, by (5) and (6),

$$p_i \Delta v_i \le \Delta \overline{U} \le p_i \Delta v_i + \Delta p_i \Delta v_i = (p_i + \Delta p_i)\Delta v_i, \text{ i.e. (4).}$$

Now let $\Delta v_i < 0$. Then $\Delta p_i \le 0$. Define the groups G_1 and G_2 in the way same as above, and G_c as those who switch <u>from</u> alternative *i*. Let the expected change in achieved utility for individuals in the groups be defined as above, i.e. Δv_i, ΔU_c and 0 respectively. Now, the probability masses of the groups are respectively $p_i + \Delta p_i$, $-\Delta p_i$ and $1 - p_i$.

Summing the effects in the different groups we have

$$\Delta \overline{U} = (p_i + \Delta p_i)\Delta v_i - \Delta p_i \Delta U_c + 0. \qquad (7)$$

Further ΔU_c obviously fulfils $\Delta v_i \le \Delta U_c \le 0$. (8)

Thus, by (7) and (8),

$p_i \Delta v_i + \Delta p_i \Delta v_i - \Delta p_i \Delta v_i \le \Delta \overline{U} \le (p_i + \Delta p_i)\Delta v_i$, i.e. (4) again. □

As a corollary we have

Proposition 1. Assume that the choice probability $P_i(\mathbf{v})$ is continuous at $\mathbf{v} = \overline{\mathbf{v}}$, then the partial derivative of $\overline{U}(\mathbf{v})$, the expected achieved utility, with respect to v_i exists at $\overline{\mathbf{v}}$, and fulfils the Williams-Daly-Zachery "relation"

$$\frac{\partial \overline{U}(\mathbf{v})}{\partial v_i} = P_i(\mathbf{v}).$$

Proof: Assume the setting of the proposition, let **v** change from $\overline{\mathbf{v}}$ to $\overline{\mathbf{v}} + \Delta v_i \mathbf{u}_i$, where $\mathbf{u}_i$ is the i-th unit vector, then, by the Lemma,

$p_i \Delta v_i \le \Delta \overline{U} \le (p_i + \Delta p_i)\Delta v_i$, (4)

where $\Delta p_i = P_i(\overline{\mathbf{v}} + \Delta v_i \mathbf{u}_i) - P_i(\overline{\mathbf{v}})$. First assume $\Delta v_i > 0$, then by division by Δv_i in (5) we get

$p_i \le \Delta \overline{U} / \Delta v_i \le p_i + \Delta p_i$. (9)

Similarly for $\Delta v_i < 0$, we get from (4) by division by Δv_i, since $\Delta v_i < 0$,

$p_i + \Delta p_i \le \Delta \overline{U} / \Delta v_i \le p_i$, (9'

Since by assumption $\Delta p_i \to 0$ when $\Delta v_i \to 0$, we get the desired conclusion, by (9) and (9').

□

So when is $P_i(\mathbf{v})$ continuous? One case is the following

Lemma 2. Suppose the c.d.f. *F* of **X** is continuously differentiable, then $P_i(\mathbf{v})$ is continuous.

Proof: Follows from Thm. 2 below. □

As a corollary we have

Corollary 1. Suppose the c.d.f. *F* of **X** is continuously differentiable, then $\overline{U}(\mathbf{v})$ is continuously differentiable and fulfils the Williams-Daly-Zachery "relation"

$$\nabla \overline{U}(\mathbf{v}) = P(\mathbf{v}),$$

where $P(\mathbf{v})$ is the vector of choice probabilities.

Proof. Under the conditions of the Corollary, $P_i(\mathbf{v})$ is everywhere continuous by Lemma 2. Thus by Prop.1, the partial derivatives $\frac{\partial \overline{U}(\mathbf{v})}{\partial v_i} = P_i(\mathbf{v})$ are continuous for all *i*. Thus, e.g. by Thm. 5.8 in Pugh (2002), $\overline{U}(\mathbf{v})$ is differentiable. Moreover $\nabla \overline{U}(\mathbf{v}) = P(\mathbf{v})$ and $\nabla \overline{U}(\mathbf{v})$ thus is continuous.

□

3. *Choice Probabilities in the RU case*.

In this section we will recall a choice probability result for RU models.

Thus suppose we have an RU model with random term $\mathbf{X} = (X_1, X_2,..., X_n) \in \Re^n$, whose cdf is *F*. In Lindberg (2012a) we derived the following result (framed in other words there). The *diagonal D* in the statement is the set $\{\mathbf{x} \in \Re^n \mid \mathbf{x} = \lambda \mathbf{1} \text{ for some } \lambda \in \Re\}$, where **1** is the "one-vector" $\mathbf{1} =_{df} (1,1,...,1)$.

Theorem 1. Suppose *F* is differentiable on the diagonal *D*, with locally bounded gradient, then

(i) $\bar{p}_i =_{df} \Pr\{X_j \leq X_i, j \neq i) = \int_{-\infty}^{\infty} F_i(x, x,..., x)dx$, (11)

where F_i is the partial derivative of *F* w.r.t. x_i, and

(ii) Probabilities of ties are 0, i.e. $\sum_i \bar{p}_i = 1$ (12)

Corollary 2. The conclusion of Thm 1 follows if *F* is continuously differentiable along the diagonal.

4. *Choice Probabilities in the ARU case.*

In the ARU case (2), each value of **v** gives an RU model, to which the results in the previous section can be applied. But we have to use that $\mathbf{U} = \mathbf{v} + \mathbf{X}$ has cdf $F(\mathbf{x} - \mathbf{v})$. Thus, for different values of **v** the diagonal "moves around". Thus we need *F* to be differentiable with locally bounded gradient in the whole space. We can now directly translate the results of the previous section.

***Proposition* 3**. Assume *F* differentiable with locally bounded gradient, then for all $\mathbf{v} \in \Re^n$

$P_i(\mathbf{v}) = \int_{-\infty}^{\infty} F_i'(x - v_1, x - v_2,..., x - v_n)dx$, and

$\sum_i P_i(\mathbf{v}) = 1$, i.e. the probability of ties is zero. □

This follows from Thm. 1. As a direct consequence we have

***Proposition* 3'**. If *F* is continuously differentiable, then for any **v** the probability of a tie is 0,

$P_i(\mathbf{v}) = \int_{-\infty}^{\infty} F_i'(x - v_1, x - v_2,..., x - v_n)dx$, and

$\sum_i P_i(\mathbf{v}) = 1$. □ (17)

5. *Continuity of the choice probabilities w.r.t one* v_i *in the ARU case*

As in the previous section, we assume that we have an ARU model $\mathbf{U} = \mathbf{v} + \mathbf{X}$, where the random term **X** has cdf $F(\mathbf{x})$. For the analysis we need to introduce the sets where maximum utility is achieved.

Thus let $\bar{A}^i(\mathbf{v}) =_{df} \{\mathbf{x} \in \Re^n \,|\, x_j + v_j \le x_i + v_i,\, j \ne i\}$, the set where alternative *i* gives the maximum utility and $A^i(\mathbf{v}) =_{df} \{\mathbf{x} \in \Re^n \,|\, x_j + v_j < x_i + v_i,\, j \ne i\}$, where alternative *i* delivers a unique maximum. Then we have

$$\bar{P}_i(\mathbf{v}) = \Pr\{\bar{A}^i(\mathbf{v})\}, \text{ and } P_i(\mathbf{v}) = \Pr\{A^i(\mathbf{v})\}, \qquad (18)$$

under the distribution given by *F*.

Further we define $A^i_{=}(\mathbf{v}) = \{\mathbf{x} \in \bar{A}^i(\mathbf{v}) \,|\, \mathbf{x} \notin A^i(\mathbf{v})\}$, where alternative *i* is maximal, but where there is a tie.

Further, since $\bar{P}_i(\mathbf{v}) = \Pr\{X_j + v_j \le X_i + v_i,\, j \ne i\} = \Pr\{X_j - X_i \le v_i - v_j,\, j \ne i\}$, we introduce $\mathbf{Y}^{(i)} = (Y_1^i, Y_2^i, ... Y_n^i)$ with components $Y_j^i = X_j - X_i$. Let the cdf of $\mathbf{Y}^{(i)}$ be $G^{(i)}$. Then we see that $\bar{P}_i(\mathbf{v}) = G^{(i)}(v_i\mathbf{1} - \mathbf{v})$. Since $G^{(i)}$ is a cdf, it is non-decreasing, and moreover continuous from the right in each argument. As a consequence we have:

***Lemma* 3**. $\bar{P}_i(\mathbf{v})$ is non-decreasing, and continuous from the right in v_i, and non-increasing, and continuous from the left in v_j, for $j \ne i$. □

To further analyze the continuity properties of the choice probabilities, let us study the relations between the $\bar{A}^i(\mathbf{v})$ and $A^i(\mathbf{v})$, when one component in **v** changes. First, we analyze the situation when only v_i varies.

Changing only v_i corresponds to adding multiples of $\mathbf{u}_i$ to **v**. We then obviously have

$$\bar{A}^i(\mathbf{v} - \delta\mathbf{u}_i) \subseteq A^i(\mathbf{v}) \subseteq \bar{A}^i(\mathbf{v}) \subseteq A^i(\mathbf{v} + \delta\mathbf{u}_i), \qquad (19)$$

i.e. $\bar{A}^i(\mathbf{v})$ and $A^i(\mathbf{v})$ are both "non-decreasing" as functions of v_i, and moreover they are interlaced as shown in (19). Eq. (19) further implies

$$\bar{P}_i(\mathbf{v} - \delta\mathbf{u}_i) \le P_i(\mathbf{v}) \le \bar{P}_i(\mathbf{v}) \le P_i(\mathbf{v} + \delta\mathbf{u}_i),$$

whence the $\bar{P}_i(\mathbf{v})$ and $P_i(\mathbf{v})$ are also non-decreasing and interlaced as functions of v_i.

Moreover, it is easily seen that if $\mathbf{x} \in A^i(\mathbf{v})$ then $\mathbf{x} \in \bar{A}^i(\mathbf{v} - \delta\mathbf{u}_i)$ for sufficiently small $\delta > 0$, which together with (19) implies that

$\bigcup_{\delta>0} \bar{A}^i(\mathbf{v} - \delta\mathbf{u}_i) = A^i(\mathbf{v})$, which in turn implies that

$\Pr\{\bar{A}^i(\mathbf{v} - \delta\mathbf{u}_i)\} \to \Pr\{A^i(\mathbf{v})\}$ as $\delta \downarrow 0$, and hence, by (18), that

$$\bar{P}_i(\mathbf{v} - \delta\mathbf{u}_i) \to P_i(\mathbf{v}) \text{ as } \delta \downarrow 0. \qquad (18)$$

Together with Lemma 3, these results give:

***Lemma* 4**. As a function of v_i, $\bar{P}_i(\mathbf{v})$ is continuous from the right. The limit from the left is $P_i(\mathbf{v})$. □

As a consequence we have:

***Proposition* 4**. If $\overline{P}_i(\mathbf{v})$ is discontinuous at $\bar{\mathbf{v}}$, as a function of v_i, then $\overline{P}_i(\mathbf{v})$ has a nonzero upward jump of size $\overline{P}_i(\bar{\mathbf{v}}) - P_i(\bar{\mathbf{v}}) = \Pr\{A^i_{=}(\bar{\mathbf{v}})\}$, and hence the probability of ties is positive.

$\overline{P}_i(\mathbf{v})$ is continuous at $\bar{\mathbf{v}}$, if and only if the probability of a tie is zero at $\bar{\mathbf{v}}$, and vice versa.

Proof: If $\overline{P}_i(\mathbf{v})$ is discontinuous at $\bar{\mathbf{v}}$, then the left limit $P_i(\bar{\mathbf{v}})$, differs from the right limit $\overline{P}_i(\bar{\mathbf{v}})$. Since $\overline{P}_i(\mathbf{v})$ is non-decreasing, the jump is positive, and by Lemma 4 of size $\overline{P}_i(\bar{\mathbf{v}}) - P_i(\bar{\mathbf{v}}) = \Pr\{\overline{A}^i(\bar{\mathbf{v}})\} - \Pr\{A^i(\bar{\mathbf{v}})\} = \Pr\{A^i_{=}(\bar{\mathbf{v}})\}$.

If $\overline{P}_i(\mathbf{v})$ is continuous at $\bar{\mathbf{v}}$, then the left limit $P_i(\bar{\mathbf{v}})$ equals from the right limit $\overline{P}_i(\bar{\mathbf{v}})$, whence $0 = \overline{P}_i(\bar{\mathbf{v}}) - P_i(\bar{\mathbf{v}}) = \Pr\{\overline{A}^i(\bar{\mathbf{v}})\} - \Pr\{A^i(\bar{\mathbf{v}})\} = \Pr\{A^i_{=}(\bar{\mathbf{v}})\}$, and we have zero probability of ties. □

We now study $\overline{P}_i(\mathbf{v})$ when only $v_j, j \neq i$, changes. The analysis is parallel to that for v_i, but a little more intricate. First we note that

$$\overline{A}^i(\mathbf{v} - \delta\mathbf{u}_j) \supseteq \overline{A}^i(\mathbf{v}) \supset \overline{A}^i(\mathbf{v} + \delta\mathbf{u}_j), \tag{19}$$

i.e. the $\overline{A}^i(\mathbf{v})$ are non-increasing as functions of $v_j, j \neq i$.

Since $\overline{P}_i(\mathbf{v})$ is continuous from the left as a function of $v_j, j \neq i$, we have to study $\overline{P}_i(\mathbf{v} + \delta\mathbf{u}_j)$ for $\delta \downarrow 0$.

If $\delta > 0$ and $\bar{\mathbf{x}} \in \overline{A}^i(\mathbf{v} + \delta\mathbf{u}_j) = \{\mathbf{x} \mid x_j + v_j + \delta \le x_i + v_i; x_k + v_k \le x_i + v_i, k \neq i, j\}$, then $\bar{\mathbf{x}} \in A^i_j(\mathbf{v}) =_{df} \{\mathbf{x} \mid x_j + v_j < x_i + v_i; x_k + v_k \le x_i + v_i, k \neq i, j\}$, i.e. the set where X_i is maximal and strictly better than X_j. Conversely, if $\bar{\mathbf{x}} \in A^i_j(\mathbf{v})$, then obviously $\bar{\mathbf{x}} \in \overline{A}^i(\mathbf{v} + \delta\mathbf{u}_j)$ for sufficiently small $\delta > 0$. Moreover the sets $\overline{A}^i(\mathbf{v} + \delta\mathbf{u}_j)$ are increasing as δ decreases. Thus $A^i_j(\mathbf{v}) = \bigcup_{\delta > 0} \overline{A}^i(\mathbf{v} + \delta\mathbf{u}_j)$, implying that $\lim_{\delta \downarrow 0} \overline{P}_i(\mathbf{v} + \delta\mathbf{u}_j) = \lim_{\delta \downarrow 0} \Pr\{\overline{A}^i(\mathbf{v} + \delta\mathbf{u}_j)\} = \Pr\{\bigcup_{\delta > 0} \overline{A}^i(\mathbf{v} + \delta\mathbf{u}_j)\} = \Pr\{A^i_j(\mathbf{v})\}$. This gives

***Lemma* 5**. As a function of $v_j, j \neq i$, $\overline{P}_i(\mathbf{v})$ is continuous from the left. The limit from the right is $\Pr\{A^i_j(\mathbf{v})\}$. □

As a preparation for the next result, let us introduce the notation $A \backslash B =_{df} \{\mathbf{x} \in A \mid \mathbf{x} \notin B\}$, and $A^i_{=,j}(\mathbf{v}) =_{df} \{x \mid x_j + v_j = x_i + v_i; x_k + v_k \le x_i + v_i, k \neq i, j\}$, i.e. the set where X_i is maximal, but there is a tie with X_j.

As a consequence of Lemma 5 we have:

***Proposition* 5**. If $\overline{P}_i(\mathbf{v})$ is discontinuous at $\bar{\mathbf{v}}$, as a function of $v_j, j \neq i$, then $\overline{P}_i(\mathbf{v})$ has a nonzero downward jump of size $\overline{P}_i(\bar{\mathbf{v}}) - \Pr\{A^i_j(\bar{\mathbf{v}})\} = \Pr\{\overline{A}^i(\bar{\mathbf{v}}) \backslash A^i_j(\bar{\mathbf{v}})\} = \Pr\{A^i_{=,j}(\bar{\mathbf{v}})\}$, and hence the probability of ties is positive.

If $\overline{P}_i(\mathbf{v})$ is continuous at $\bar{\mathbf{v}}$, as a function of v_j, $j \neq i$, the probability of a tie between X_i and X_j is zero at $\bar{\mathbf{v}}$, and vice versa

Proof: In the same way as for Lemma 4, but using that $\overline{A}^i(\bar{\mathbf{v}}) \setminus A^i_j(\mathbf{v}) = A^i_{=,j}(\mathbf{v})$. □

Applying Prop. 3 to Propositions 4 and 5, we get:

Corollary 3. In an ARU model with cdf *F* of the random term, the choice probabilities $\overline{P}_i(\mathbf{v})$ are continuous as both as functions of v_i and of v_j, $j \neq i$, if either

(i) the cdf *F* is continuously differentiable, or

(ii) the probabilities of ties are zero for any values of the population values **v.**

6. *Continuity of the choice probabilities in the ARU case.*

We can now put together the results of the previous sections

Theorem 2. Consider an ARU model, **U** = **v** + **X,** whose random term **X** has cdf *F*. If the choice probability function $\overline{P}_i(\mathbf{v})$ is continuous in each argument, then it is continuous as a function of $\mathbf{v}$. In particular, the $\overline{P}_i(\mathbf{v})$ are continuous, if

(i) *F* is continuously differentiable, or

(ii) the probabilities of ties are zero for any values of the population values **v.**

Proof: Assume $\overline{P}_i(\mathbf{v})$ is continuous as a function of any single argument, v_i, or v_j for $j \neq i$,.

Consider now an arbitrary $\bar{\mathbf{v}}$. Assume that (by renumbering) we study $\overline{P}_1(\mathbf{v})$.

Since $\overline{P}_1(\mathbf{v})$ is continuous as function of v_1, there is a $\delta_1 > 0$, such that $\overline{P}_1(\bar{\mathbf{v}}) - \varepsilon \leq \overline{P}_1(\bar{\mathbf{v}} + v_1\mathbf{u}_1) \leq \overline{P}_1(\bar{\mathbf{v}}) + \varepsilon$ for $-\delta_1 \leq v_1 \leq \delta_1$, and in particular $\overline{P}_1(\bar{\mathbf{v}} + \delta_1\mathbf{u}_1) \leq \overline{P}_1(\bar{\mathbf{v}}) + \varepsilon$.

Similarly, varying v_2 around $\bar{\mathbf{v}} + \delta_1\mathbf{u}_1$ there is a $\delta_2 > 0$, such that $\overline{P}_1(\bar{\mathbf{v}} + \delta_1\mathbf{u}_1) - \varepsilon \leq \overline{P}_{1i}(\bar{\mathbf{v}} + \delta_1\mathbf{u}_1 + v_2\mathbf{u}_2) \leq \overline{P}_1(\bar{\mathbf{v}} + \delta_1\mathbf{u}_1) + \varepsilon \leq \overline{P}_1(\bar{\mathbf{v}}) + 2\varepsilon$ for $-\delta_2 \leq v_2 \leq \delta_2$, and in particular

$$\overline{P}_i(\bar{\mathbf{v}} + \delta_1\mathbf{u}_1 - \delta_2\mathbf{u}_2) \leq \overline{P}_i(\bar{\mathbf{v}}) + 2\varepsilon$$

Continuing in this fashion for v_3, $v_4, \ldots, v_n$, we finally get

$$\overline{P}_1(\bar{\mathbf{v}} + \delta_1\mathbf{u}_1 - \textstyle\sum_{k=2}^{n}\delta_k\mathbf{u}_k) \leq \overline{P}_{1i}(\bar{\mathbf{v}}) + n\varepsilon, \text{ for appropriate } \delta_k > 0. \quad (20)$$

Similarly, working in the other direction, we get

$$\overline{P}_1(\bar{\mathbf{v}} - \delta_1'\mathbf{u}_1 + \textstyle\sum_{k=2}^{n}\delta_k'\mathbf{u}_k) \geq \overline{P}_1(\bar{\mathbf{v}}) - n\varepsilon, \text{ for appropriate } \delta_k' > 0. \quad (21)$$

Letting $\bar{\delta}_k \min(\delta_k, \delta_k')$, (20) and (21) imply, together with the monotonicity of $\overline{P}_1(\mathbf{v})$, that for any $v_k \in [-\bar{\delta}_k, \bar{\delta}_k]$,

$$\overline{P}_1(\bar{\mathbf{v}} + v_1\mathbf{u}_1 + \textstyle\sum_{k=2}^{n} v_k\mathbf{u}_k) \leq \overline{P}_1(\bar{\mathbf{v}} + \bar{\delta}_1\mathbf{u}_1 - \textstyle\sum_{k=2}^{n}\bar{\delta}_k\mathbf{u}_k) \leq \overline{P}_1(\bar{\mathbf{v}}) + n\varepsilon, \text{ and}$$

$$\overline{P}_1(\overline{\mathbf{v}} + v_1\mathbf{u}_1 + \sum_{k=2}^{n} v_k\mathbf{u}_k) \geq \overline{P}_1(\overline{\mathbf{v}} - \overline{\delta}_1\mathbf{u}_1 + \sum_{k=2}^{n} \overline{\delta}_k\mathbf{u}_k) \geq \overline{P}_1(\overline{\mathbf{v}}) - n\varepsilon\,,$$

Which in turn implies that $\overline{P}_1(\mathbf{v})$ is continuous, which proves the first part of the Theorem.

The second part follows from the first part since by Cor. 3, all $\overline{P}_i(\mathbf{v})$ are componentwise continuous if *F* is continuously differentiable, or if the probability of ties is zero. □.

Acknowledgement. This paper has benefitted from comments by Dr M Fosgerau.

References

Anderson, S.P., A. de Palma and J.-F. Thisse (1992), *Discrete Choice Theory of Product Differentiation*, The MIT Press, Cambridge, MA.

Ben-Akiva, M. and S.R. Lerman (1985), *Discrete Choice Analysis*, The MIT Press, Cambridge, MA.

Daly, A. and S. Zachery (78), Improved multiple choice models, in D. Henscher and Q. Dalvi (eds.) *Identifying and Measuring the Determinants of Mode Choice*, Teakfields, London

Fosgerau, M., D. McFadden and M. Bierlaire (10), Choice probability generating functions, Report TRANSP-OR 100605, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne.

Keane, M., P. Todd and K.I. Wolpin (2010), The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications, *Handbook of Labor Economics* , eds. Card and Ashenfelter, Volume 2, Elsevier, p. 332-461.

Lindberg, P.O., A. Eriksson and L.-G. Mattsson, Invariance of achieved utility in random utility models, *Environment & Planning* A **27**(1), 121-142.

McFadden, D. (1974), Conditional logit analysis of choice behavior, in Zaremba (ed.), *Frontiers in Econometrics*, 75-96, Academic Press, New York.

McFadden, D. (1978), Modeling the Choice of Residential Location, in Karlqvist et al. (eds.), *Spatial Interaction Theory and Planning Models*, 75-96, North Holland, Amsterdam.

McFadden, D. (1980), Econometric Models for Choice Among Products, J. Business **53**, S13-S29.

McFadden, D. (1981), Econometric Models of Probabilistic Choice, in Manski and McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, 198-272, MIT Press, Cambridge, Mass.

Pugh, C.C. (2002), *Real Mathematical Analysis*, Springer.

Williams, H.C.W.L. (1977), On the Formulation of Travel Demand Models and Economic Measures of User Benefit, *Environment & Planning* A **9**(3), 285-344.