NOTES ON STOCHASTIC CHOICE

Andreu Mas-Colell

Universitat Pompeu Fabra

Abstract

A general formalism on stochastic choice is presented. Tje Rationalizability and Recoverability (Identification) problems are discussed. For the identification issue parametric examples are analyzed by means of techniques of mathematical tomography (Random transforms)

Key words : Stochastic Choice, Recoverability, Identification, Random transform

JEL CODES : D11, D81, D83

PROLOGUE (2005)

The notes on Stochastic choice that follow were presented at a meeting held in San Sebastian in June of 1983 and organised by Salvador Barberà. It was research in progress that, alas, was never pursued. But it seems, by its subject, a most indicated contribution to a volume to honour Ket Richter. Obviously, I have the hope, but not the certainty, that something is still of interest in them. Or simply that there will be something to catch the sharp analytical eye of Ket. With my best regards to Ket, a model for us all of how theory should be done, here they go. I have corrected some obvious inaccuracies and, occasionally, tightened some looseness of language. I have also added some references (in particular, Falmagne,1978, Fishburn,1998, Barberà and Pattanaik,1986, McFadden and Richter,1991, McFadden,2004, are very relevant to the subject matter of these notes) and taken into account the remarks of a referee(whom I thank). Otherwise the text is as in 1983.

I. **A General Formalism**

A very general setting for the stochastic choice problem can be described thus (see also Manski, 1977). There is given as data:

(1) A set of alternatives X. It is convenient to think of X as finite.

(2) A set of "budgets" $\mathcal{B} \subset 2^X$. Put $Y = \prod_{B \in \mathcal{B}} B$. A point of Y is a selection of an alternative in every budget. Denote by $\mathcal{M}$ the probability measures on Y.

(3) A set of admissible statistics $f_j : \mathcal{M} \to R, j \in J.$

(4) A set of observed values $a_j, j \in J$ of the statistics.

As an example, the usual stochastic choice problem corresponds to the above where the admissible statistics in (3) are the marginal distributions. [Precisely: f is admissible if and only if it is of the form $f(v)=\int \psi(y)dv$ where $\psi(y)$ is the projection on one variable]. Even more restricted, if the admissible statistics are the mean of every marginal we have as data a sort of aggregate demand. Another situation falling in the above setting would be one where for every alternative x we are given the probability that x is chosen for some budget, etc.

Denote by $\mathcal{P}$ the set of linear orders on X.
Every probability measure μ on $\mathcal{P}$ induces a probability measure ν_μ on Y by the rule $\nu_\mu(A)=\mu\{\succ\in\mathcal{P}:$ denoting by $x(B)$ the $\succ-$maximal element on $B\in\mathcal{B}$ we have $\{x(B)\}_{B\in\mathcal{B}}\in A\ \}$.

That is to say, ν_μ is the measure generated on Y by the choice vectors induced by preferences.

We then have two problems:

Rationalizability problem: A stochastic choice situation (described by (1)-(4) above) can be rationalized if there is a probability measure μ on $\mathcal{P}$ such that $f_j(\nu_\mu)=a_j$ for every j.

Which conditions must the data of the problem satisfy in order for a rationalization to exist?

Recoverability (or identification) problem: Assuming that the data are rationalizable, when is the rationalization unique?

Remark: Strictly speaking there is still a third problem, previous to the rationalizability one and vacuously non-restrictive in the usual stochastic choice model. It could be called the compatibility problem, namely, under which conditions there is a probability measure ν on Y such that $f_j(\nu) = a_j$ for every j.

II.- **A Particular Case**

After so much generality I become very concrete. I concentrate for the rest of the Notes on the particular case where there is a distinguished alternative, denoted 0, every $B \in \mathcal{B}$ includes 0 and for each $B \in \mathcal{B}$ there is an admissible statistic which is the probability that 0 is not selected in B. In other words, the data of the problem is an array $p(B)$, $B \in \mathcal{B}$, to be interpreted as asserting that given B the probability that 0 be the preferred element is $1 - p(B)$. We always put $p(\{0\}) = 0$.

Define the equivalence relation $\approx$ on $\mathcal{P}$ by $\succ \approx \succ'$ iff $"x \succ 0 \Leftrightarrow x \succ' 0"$. Obviously, if $\succ \approx \succ'$ then the data of the problem will never be able to distinguish between $\succ$ and $\succ'$. Therefore, the rationalizability and, above all, the recoverability problem should properly be posed with respect to $\mathcal{P}^* = \mathcal{P}/\approx$. Note that for the elements of $\mathcal{P}^*$ the

transitivity requirement has no strength. Avoiding the transitivity issue is the main advantage of analyzing the particular case of a distinguished alternative.

I briefly discuss three subcases that differ by the nature of the admissible □. Take X finite, with $\#X = n+1$.

(a) $\mathcal{B} = \{B \in 2^X : 0 \in B\}$

The rationalization problem for this subcase has been extensively treated and is completely solved. See Falmagne, 1978, Barberà and Pattanaik, 1986, Cohen and Falmagne, 1990, Barberà, 1991..

Every preference in $\mathcal{P}^*$ can be identified with a set $B \in \mathcal{B}$, i.e. B is the set of alternatives at least as good as 0. Then a probability on $\mathcal{P}^*$ can be identified with a list $0 \le \pi(B) \le 1$, $\sum_{B \in \mathcal{B}} \pi(B) = 1$. If π ( ) rationalizes p ( ) then we must have $1 - p(B) = \sum_{\substack{A \in B \\ A \cap B\{0\}}} \pi(A)$ for every $B \in \mathcal{B}$. Therefore, p ( ) can be rationalized if and only if the following recursion process (see Barberà and Pattanaik, 1986)yields a probability measure. Put first. $\pi(\{0\}) = 1 - p(X)$. Suppose now that $\pi(C)$ has been computed for any C up to size $m+1$. Put then $\pi(B) = 1 - p((X \setminus B) \cup \{0\}) - \sum_{C \subset B} \pi(C)$ for *B* of size *m* + 2

Obviously, this recursion process gives us a complete list $\pi(B)$, $B \in \mathcal{B}$. Also, $\sum_{B \in \mathcal{B}} \pi(B) = 1$ by construction. Therefore, π is a

probability measure, i.e. in admissible rationalization, if and only if $\pi(B)$ is non-negative for all B. Those are the conditions obtained in the above references. Note that if $p(\ )$ is rationalizable then the rationalization is unique and can be recovered by the previous recursion.

Recoverability, i.e. uniqueness, is not surprising in view of the fact that one gets from π to p by a linear transformation and that there are as many equations (one for each B) as unknowns (one for each B).

(b) $\mathcal{B} = \{B \in 2^X : 0 \in B,\ \#B = 2\}$

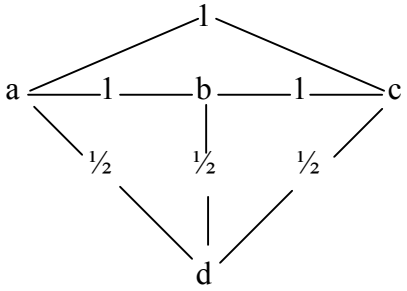
This is in a sense the polar opposite to subcase (a). Here we only have the outcome of the pairwise matching of 0 against every $x \neq 0$. We write $p(\{0,x\}) = p(x)$.

It is obvious that in this subcase, where there is much less information than in (a), any $p(\ )$ can be rationalized. Indeed, any $p(\ )$ can be looked at as a point in $[0,1]^n$. Every extreme point of this convex set is of the form $p(\ ) \in \{0,1\}$ and can be rationalized (by ordering $"x \succ 0 \Leftrightarrow p(x) = 1"$). Therefore, the entire $[0,1]^n$ can be rationalized.

The counterpart to the above pleasant fact is that the preference distribution cannot be recovered. This is clear counting equations (n) and unknowns $(2^n$, one for every $B \in 2^X,\ 0 \in B)$.

(c) Intermediate subcases

To get clean results for families □ intermediate between subcases (b) and (a)is hard. Consider, for example, the subcase $\square = \{B \in 2^X : 0 \in B \text{ , } \#B = 3 \}$. It is a good exercise to verify that for the instance represented in the figure below (with $x = 0, a, b, c, d$) there is no rationalizing preferences.



From now on I limit myself to subcase (b), i.e. our data is the probability $p(x)$ of any $x \in X$ winning against 0. The common fact in the next two sections is that restrictions are imposed on underlying permissible preferences. In section III I study a rationalization problem with a convexity hypothesis on preferences. In section IV I sketch and discuss an analytic treatment of the recoverability problem.

III.- **Rationalizability with convex preferences**

With 0 a distinguished alternative in X we are given, for every $x \in X$, $x \neq 0$ a number $0 \leq p(x) \leq 1$ which is interpreted as the

probability of *x* winning over 0. We have seen (subcase (b) in II) that p can always be rationalized by a distribution μ on $\mathcal{P}$. In applications, however, it may be important that μ give positive weight only to preferences satisfying some restrictions.

Suppose, for example, that $X^* = X \setminus \{0\}$ is a subset of a linear space. Say $X^* \subset R^m$. Then we may be interested in rationalizing by members of the set of convex (or, more precisely, convex-compatible) preferences, i.e. $\mathcal{P}_C = \{\succ \in \mathcal{P}$: if "if $A \subset X^*$, $x \succ 0$ for every $x \in A$, and $y \in$ (convex hull $A) \cap X^*$" then $y \succ 0\}$.

It is no longer true that any $p(\,.\,)$ can be rationalized by a μ concentrated on $\mathcal{P}_C$. The problem of characterizing the set of admissible $p(\,.\,)$ seems pretty hard indeed. But for the simple case, i.e. $m = 1$ (the set X^* lies in the real line) the solution is fairly trivial.

Let $X^* \subset R$. Put $X^* = \{x_1, ..., x_n\}$ where $x_i > x_j$ for $i > j$. Denote $p(i) = p(x_i)$.

Proposition: The function $p : X^* \to [0,1]$ can be rationalized by a μ on $\mathcal{P}_C$ if and only if $p_1 + \sum_{i=2}^{i=n} \max\{0, p_i - p_{i-1}\} \leq 1$.

Remark: Presumably the proposition can be extended to the case where $X^* \subset R$ is compact. The general statement would then be along the lines: "The function $p : X^* \to [0,1]$ can be rationalized by a μ on $\mathcal{P}_C$

if and only if it is of bounded variation and has variation norm $\leq 1''$.

Proof of the Proposition:

(1) Necessity: Identifying sets with preferences let $\mathcal{B}_C$ be the set of convex preferences. For every $i = 1,...,n$ denote $\mathcal{B}_i = \{B \in \mathcal{B}_C : x_{i-1} \notin B, x_i \in B\}$. These sets constitute, by the convexity hypothesis, a partition of $\mathcal{B}_C$. So, if π is a probability measure concentrated on $\mathcal{B}_C$ we have $\sum_{i=1}^{n} \pi(\mathcal{B}_i) \leq 1$. Suppose now that π generates p. Then $\pi(\mathcal{B}_1) = p_1$. Consider any $i > 1$. We have $p_i = \pi(\{B \in \mathcal{B}_C : x_i \in B\})$. But $\{B \in \mathcal{B}_C : x_i \in B\} = \mathcal{B}_i \cup \{B \in \mathcal{B}_C : x_i \in B$ and $x_j \in B$ for some j less than i) . This is a disjoint union and, by convexity, the second set is a subset of $\{B \in \mathcal{B}_C : x_{i-1} \in B\}$ which probability is p_{i-1}. Therefore $p_i \leq \pi(\mathcal{B}_i) + p_{i-1}$, or $\max\{0, p_i - p_{i-1}\} \leq \pi(\mathcal{B}_i)$. Hence, $p_1 + \sum_{i=2}^{n} \max\{0, p_i - p_{i-1}\} \leq 1$ and necessity is established.

(2) Sufficiency: We shall actually show that: "There is always a π such that $\sum_{B \neq \phi} \pi(B) = p_1 + \sum_{i=2}^{n} \max\{0, p_i - p_{i-1}\}$". So, let the bracketed statement be an induction hypothesis on n. It is obviously true for $n = 1$. Let it be true for $n-1$. In particular for the set $\{x_1,...,x_{n-1}\}$, i.e. there is a probability measure π on $\mathcal{B}_C^{n-1} = \{B \in \mathcal{B}_C : x_n \notin B\}$ such that:

(a) $\sum_{\substack{B\in\mathcal{B}_C^{n-1}\\ B\neq\phi}}\pi(B)=p_1+\sum_{i=2}^{n-1}\max\{0,p_i-p_{i-1}\}\le 1$, and

(b) for every $i\le n-1, p_i=\sum_{\substack{x_i\in B\\ B\in\mathcal{B}_C^{n-1}}}\pi(B)$.

Now we extend π to X^* as follows. Let $q_n = \min\{p_{n-1}, p_n\}$.

For any $B \in \mathcal{B}_C^{n-1}$ such that $x_{n-1}\in B$ consider the rule $B\to B\cup\{x_n\}$. Under this rule transfer a probability weight q_n from $\{B\in\mathcal{B}_C^{n-1}: x_{n-1}\in B\}$ to $\mathcal{B}_C$. If $q_n = p_n$ then we are done: the equality in (a) has not been altered and (b) also holds for $i=n$. If $q_n = p_{n-1} < p_n$ then we in addition transfer a probability weight $p_n - p_{n-1}$ from the set ϕ to the set $\{x_n\}$. This can be done because by the induction hypothesis $\pi(\phi)=1-\left(p_1+\sum_{i=2}^{n-1}(p_i-p_{i-1})\right)\ge p_n-p_{n-1}$. Then again the equality in (a) remains and (b) has been extended to $i=n$. This concludes the induction step.

Remark: As it should be expected if the condition of the proposition holds then the admissible probability on preferences need not be unique. Suppose that $X^*=\{1,2\}$ and $p_1=\frac{1}{3}$, $p_2=\frac{1}{3}$. Then two admissible π are "$\pi(\{1\})=\pi(\{2\})=\frac{1}{3}$, $\pi(\phi)=\frac{1}{3}$" and "$\pi'(\{1,2\})=\frac{1}{3}$, $\pi'(\phi)=\frac{2}{3}$". The π obtained by construction in the proof of the proposition would be π' in this example, namely, it is the one that maximizes the probability that 0 be the overall maximin,

i.e. $\pi(\phi)$. The construction of the proof seems to indicate that this maximizing probability measure is unique.

IV.- **Analytic treatment of the recoverability problem**

We keep studying the distinguished alternative case. We now take X to be an Euclidean space R^n. The distinguished alternative is the origin 0. The function $p: X \to [0,1]$ gives the probability $p(x)$ that x wins against 0. For convenience, p is left undefined at 0.

For an analytic treatment it is important (or, at least, convenient) that the set of admissible preferences be somehow restricted to depend on a finite number of parameters. So, we assume that we have given a parameter set Q which, to make life simple, we identify with some Euclidean space R^m. For every parameter value $q \in Q$ preferences are expressed by a utility function $\mathrm{U}(x,q)$, normalized to equal zero whenever x = 0. It is assumed that $\mathrm{U}: R^n \mathrm{x}\, R^m \to R$ is a "nice" function (continuous, differentiable, analytic,…).

Given a probability measure μ on Q a probability choice function $P: X \to [0.1]$ is generated as follows:

$p(x) = \mu\{q : \mathrm{U}(x,q) > 0\} = \int_{\{q:\mathrm{U}(x,q)>o\}} f(q)dq$ where the second equality applies only if M has a density f. From now on we shall assume that all μ we deal with have densities which, moreover, are sufficiently nice

(say of class C^∞ and equal to zero outside of a compact set, or, at least, "rapidly decreasing").

The recoverability problem is then: Assuming that p is generated as above, can f be uniquely recovered from p (in the class of "nice" densities)?

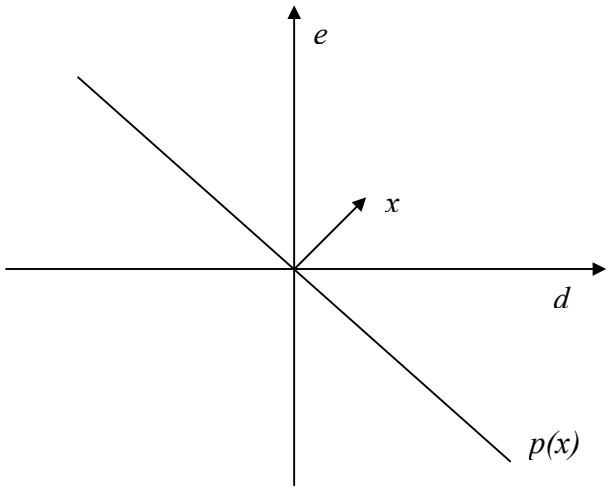
Remark: Strictly speaking the recoverability problem should be posed only for the family of indifference curves passing through the origin because this is all the information that p uses. But, in the parameterized world we are now working in, recovering the indifference curve will usually be tantamount, (i.e. except perhaps for a normalizing parameter) to recovering the entire utility function. So, I do not worry about the distinction.

For the remaining of this section I discuss an extended example with $n = 2$ and $\mathrm{U}(x,q)$ a general quadratic: $\mathrm{U}(x,q) = ax_1^2 + bx_2^2 + cx_1x_2 + dx_1 + ex_2$. So, without further a priori restrictions we have five parameters, i.e. $m = 5$. I consider a sequence of three subcases, which differ by the type of a priori restriction imposed.

Example 1: Take $a = b = c = 0$ as a priori restrictions.

(The same qualitative features of the example are obtained with other combinations of three zero restrictions, eg. $c = d = e = 0$, or $b = d = e = 0$). In this case $\mathrm{U}(x,q)$ reduces to $\mathrm{U}(x,q) \equiv \mathrm{U}(x,d,e) = dx_1 + ex_2$.

This model is not identified. Take, for example, $p(x)=\frac{1}{2}$ for all x. Any symmetric density f on the $d-e$ plane will generate p because, for any x, $\{d,e: dx_1+ex_2>0\}$ is just the half space above the hyperplane with normal x and the integral of a symmetric density on a half space is $\frac{1}{2}$. See the figure.



Example 2: As in example 1, $a = b = c = 0$. But suppose now that in addition there is another restriction in the form of a non-homogeneous linear equation. For instance, $d+e=1$. The origin of this restriction could be, for example, a normalization convention.

Then the model is obviously identified because given any underlying density f we can use p to compute the distribution function of f on the line defined by $d+e=1$ on the $d-e$ plane.

Observe also that f can be recovered by using only the information contained in the p function in any arbitrarily small neighbourhood of zero.

After discussing two more examples I shall present, in the next section, a recoverability proposition for arbitrary m and n which generalizes Example 2.

Example 3: $a=1$, $b=1$, $c=0$.

In this case for given d,e the indifference curves of the utility function $\mathrm{U}(x,e,d)=x_1^2+x_2^2+dx_1+ex_2$ are concentric circles around the vector $\left(-\frac{d}{2},-\frac{e}{2}\right)$. i.e., x is preferred to 0 according to if $\left(-\frac{d}{2},-\frac{e}{2}\right)$ is closer to x than to 0.

So, in the obvious way we can identify the variable and the parameter space and think of densities f as being defined on the x space itself (think of the parameter as the peak of the preferences). Note that $p(x)$ is the integral of f on the half space of vectors to the side that includes x of the line perpendicular to x and cutting the segment $[0,x]$ in its middle point (this is the half space of vectors closer to x than to 0):

[Remark: The similarities of this with the well known majority voting model are intended].

Now a mathematical disgression.

Let S^1 be the 1-dimensional sphere in two dimensional Euclidean space. Given f we can define a function $\psi : S^1 \mathrm{x}\, R \to R$ by letting $\psi(v,t)$ be the integral of the f function on the line (more generally, affine subspace) $\{y : v.y = t\}$ endowed with the usual Lebesgue measure. In Fourier analysis this function (as well as its obvious higher dimensional generalizations $\psi : S^{n-1} \mathrm{x}\, R \to R$) is known as the Radon transform of f and, not surprisingly, it is useful in things like X-ray reconstruction. The fact is that there is an inversion formula such that if f is "nice" then starting with $\psi(v,t)$ we recover f.

The inversion formula is particularly simple for the case at hand where f is defined on the plane and the Radon transform (also called in this case the X-ray transform) evaluates integrals on lines. For any x and $s > 0$ let $n_s(x)$ be the average value of $\psi(v,t)$ on lines which are at a distance s from x, i.e. $n_s(x) = \frac{1}{2\pi} \int_{S^1} \psi(v, v.x + s) dv$.

Then it turns out that if f is continuous and has a compact support $f(x)$ can be recovered by the formula $f(x) = \frac{1}{\pi} \int_0^\infty \frac{dn_s(x)}{s}$ where the integral is in the sense of Stieltjes. More precisely, and integrating the above formula by parts:

$$f(x) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \left(\frac{\eta_\varepsilon(x)}{\varepsilon} - \int_\varepsilon^\infty \frac{n_s(x)}{s^2}\, ds \right)$$

See Shepp and Kruskal (1978), Helgason, (1980), or Dym and McKean (1972) for these Fourier analysis techniques. Their relevance for recoverability problems in economics has been noted in another

context by Ph. Dybvig and A. McLennan. I would also like to thank A. Grunbaum for the mathematical references.

Back to Example 3. The relevance of the mathematical disgression to our problem is that the Radon Transform of the density f can be computed from the choice probabilities $p(x)$. As it is clear from the geometric discussion:

$$\psi(v,\, t) \;=\; -\,\frac{\partial}{\partial t}\; p(2tv)$$

(Strictly speaking the above applies to $t \neq 0$. For $t = 0$ just let $\psi(v,0) = \lim_{t\to 0}\psi(v,t)$)

Summing up: the model of example 3 has the recoverability property. Note however that, in contrast to example 2, it is now essential to use all the information contained in $p(x)$. Restricting onerself to a small neighbourhood of 0 will not do.

Remark: Given an arbitrary p ( ) we can compute $\psi(v,t)$ as above by means of the inversion formula to get a $f(x)$. That f be a well defined (and "nice") density function (i.e. $f(x) \geq 0$ and $\int f(x)dx = 1$) is, therefore, the necessary and sufficient condition for rationalizability within the restrictions of Example 3. What one gets, unfortunately, is not precisely a transparent condition.

Example 4: This is not a quadratic but a cubic example:

$\mathrm{U}(x,q) = x_2 - ax_1^3 - bx_1^2 - cx_1$ For given a,b,c the equation

$x_2 = ax_1^3 - bx_1^2 - cx_1$ yields a non-linear indifference curve through the origin. Actually, I have no idea if this model is identified or not.

Since we only have two variables but three parameters the guess is that it is not but…

In the next section I present the promised generalization of Example 2.

V. A generalization of Example 2.

Let's go back to the original set-up of Section IV with

$$\mathrm{U} : R^n \mathrm{x}\, R^m \to R$$

Suppose first that U takes the additive form:

$\mathrm{U}(x,q) = g_1(x)q_1 + ... + g_m(x)q_m = g(x).q$. This covers all the polynomial cases and, pushed to the limit, could cover all the analytic utility functions. If q^3 lies in the segment $[q^1, q^2] \subset R^n$ then $\mathrm{U}(.,q^3)$ is intermediate between $\mathrm{U}(.,q^1)$ and $\mathrm{U}(.,q^2)$ (or, rather, their preference relations are) in the sense used by Chichilnisky and Grandmont. In fact, one could wonder if for $m \geq 3$ the concept of intermediate preferences provides a characterization of the above additive form We assume that the function $g : R^n \to R^m$ is C^1.

Suppose that in the space of parameter R^m there are some a priori given identifying restrictions in the form of a system of s linear equations:

$$\underset{s\mathrm{x}m}{B}\ q - c = 0$$

The density f is supported in the set of solutions to the above system. Hence, it is in the nature of the problem that solutions exists.

Proposition: A sufficient condition for the model to be identified, i.e., for every nice f to be recoverable, is that:

$$rank\begin{bmatrix} \underset{n\text{x}m}{(\partial g(0))^T} & \underset{n\text{x}1}{0} \\ \underset{s\text{x}m}{B} & \underset{s\text{x}1}{-c} \end{bmatrix} = m+1$$

Moreover, only the values of p ( ) on a neighbourhood of 0 matter.

Sketch of proof: Denote

$L = \partial g(0)(R^n)$,

$M = \{q : Bq = \alpha c \text{ for some } \alpha\}$

$N = \{q : Bq = c\}$.

The three are subspaces of R^m (N is affine). It is a simple exercise to verify that if the rank condition is satisfied then the dimension of L is not smaller than the dimension of M and, in fact, that the projection of L on M is onto.

We now argue that any affine half space in N i.e. any set of the form $A = \{q \in N : q.y < \beta\}$, $y \in R^m$, can be realized by taking a y belonging to L and putting $\beta = 0$. Indeed, we can first realize A in the form of $A = \{q \in N : q.z < \bar{q}.z\}$, where $\bar{q} \in A$ and z belongs to the translate of N to the origin. If the rank condition is satisfied then $c \neq 0$. So, $0 \notin N$ and therefore $\{q \in N : q.z = \bar{q}.z\}$ spans a hyperplane in M. By the observation in the previous paragraph this hyperplane is realized for some $y \in L$. This y does the job.

Appealing now (with some care) to the Implicit Function Theorem we conclude that any affine half space in N can be realized in the form $\{q \in N : q.g(x) < 0\}$ for an arbitrarily small x.

Because the density function f lies in N and $p(x) = \int_{\{q \in N : q.g(x \le 0)\}} f(x)dx$ we can finally recover f from p by using the Fourier analysis techniques discussed in Example 3. This ends the sketch of proof.

Example 3 shows that the rank condition is sufficient but not necessary for identification. The ability to use any x not limited to a neighbourhood of the origin, may make up for insufficient variation of g at 0. Nevertheless, it can be presumed (?) that a more general condition will again revolve on a counting of effective parameters versus independent directions of variations of $g(x)$.

Remark: The entire analysis of this section uses only the information contained in $p(x)$, i.e. only on the pairwise comparison that include the origin. It stands to reason that if more information was available, eg. on all pairwise comparisons, then fewer identifying restrictions would suffice.

References: Barberà, S. and P. Pattanaik (1986): Falmagne and the rationalizibility of Stochastic Choices in Terms of Random Orderings, Econometrica, 54, 707-716.

Barberà, S. (1991): Rationalizable Stochastic Choice over Restricted Domains, in Chipman, *et.al.*, 203- 217.

Chipman,J., D. McFadden, and M.Richter, eds, (1991) Preferences, Uncertainty and Rationality, Westview Press.

Cohen, M. and J. Falmagne, (1990): Random Utility Representation of Binary Choice Probabilities: A New Class of Necessary Conditions, Journal of Mathematical Psychology, 34, 88-94.

.Dym, H. and H. McKean (1972): Fourier Series and Integrals, Academic Press.

Falmagne, J. (1978) A Representation Theorem for Finite Random Scale Systems, Journal of Mathematical Psychology, 18, 52-72.

Fishburn, P. (1998) Stochastic Utility, in S.Barberà, P. Hammond and C. Seidl, eds,: Handbook of Utility Theory, Kluwer, 273-320.

Helgason, S. (1980): The Radon Transform. Burkhauser.

McFadden, D.: Revealed Stochastic Preference: A

Synthesis, University of California, Berkeley, 2004.

McFadden, D. and M. Richter (1991): Stochastic Rationality and Revealed Stochastic Preference, in Chipman *et.al.*, 166-186.

McLennan, A. (1991): Binary Stochastic Choice, in Chipman *et.al.*, 187-202.

Shepp, L. and J. Kruskal: Computerized Tomography, the new X-ray technology. American Mathematical Monthly, 1978.