Meta-Preferences: What Are They? And Are They Different From Ordinary Preferences?

Bernhard Wieland

1. Introduction

Over the last two decades economists have increasingly become interested in the topic of self-control. Within public choice Buchanan has been part of this development in his attempt to "model the individual for constitutional analysis" (1975, 1979, 1985). According to Buchanan a person chooses in a quasi existentialistic fashion what kind of person he or she wants to be. An individual must choose at some point in time ".... a life plan, a sequence of actions, that he hopes will ensure that his experiences are 'interesting', 'good', 'rewarding', and/or 'happy' " (1985: 69). Not choosing will also be a choice. Buchanan models this choice by assuming that the individual "constructs himself " as a sequence of persons through periods $t_0, t_1, \ldots, t_n$, recognizing the temporal dependency of the later on the earlier "persons". Each "person" in $t_k$ is modeled as a preference ordering $\preceq_k$ in $t_k$. Thus, choosing a life plan or profile of behaviour amounts to choosing among sequences of possible preference orderings. But in order to be able to make such choices a person must possess a preference ordering of higher order, that is a preference for preferences, or a meta-preference ordering. This preference relation of second order may induce many of the "rules" that people impose voluntarily on their behaviour because people try to approximate the sequence of chosen preference orderings by restraining voluntarily their choice sets. Such rules may be regular working habits or strict rules with respect to eating, drinking or borrowing. Similar ideas have been developed by other authors like Sen (1977), Thaler/Shefrin (1981), or Etzioni (1986) to name just a few.

As one would expect the use of meta-preferences in economic theory has been criticised. Broadly speaking the proponents of criticism can be divided into two types. The first type of critics, the revisionists, argue that everything that can be explained with the help of meta-preferences can be explained within a mono-utility framework as well. Indeed, as numerous
writers have shown, the mono-utility framework is sufficiently malleable to "explain" almost every kind of behaviour. From a methodological point of view it is doubtful, however, whether this is a virtue. The second type of critics are the radicals. They claim that economic theory should discard the notion of preference ordering altogether. For them the criticism that can be applied to the mono-utility model holds for the multiple utility framework as well. In their view it is unavoidable to include elements from psychology and other social sciences if one wants to arrive at a realistic model of human behaviour. As an intermediate position one may consider the view that people have multiple utilities but that there is no hierarchy among these utilities. (For an overview over this debate see Brennan (1989), Lutz (1993) Brennan (1993).)

In this paper I want to direct attention to a different criticism. I shall show that under a well known rationality assumption (Richter's Strong Congruence Axiom) meta-preferences can be reduced to just another preference ordering of the underlying set of alternatives X. That is, if there are n preference orderings \( \leq_1, \leq_2, \ldots, \leq_n \) on X and if there is a meta-preference ordering \( \preceq \) on the set \( \{\leq_1, \leq_2, \ldots, \leq_n\} \) then there exists a further preference ordering \( \leq_{n+1} \) on X that explains all the individual's observed choices. If this result is true, then meta-preferences would seem to be superfluous as a theoretical tool. In accordance with the methodological principle that "entities are not to be multiplied without necessity" (Ockham's razor) their use should be avoided.

Another consequence of the just mentioned result is that for economic analysis it is irrelevant, whether people give themselves rules or not. All that is necessary for economic analysis is that their behaviour can be modeled as if they acted according to a certain preference ordering. Whether they first choose this preference ordering in a quasi pre-constitutional personal state and whether they achieve adherence to this preference ordering by acting through self-imposed rules is of no relevance for economic analysis in general or constitutional economics in particular. This objection would seem to apply not to Buchanan's variety of meta-preferences but to all economic models of human behaviour that use higher order preferences.
In saying this I do not want to deny the importance of meta-preferences for other fields like sociology or psychology. It may also be true, that subjectively people behave very much in the way described by Buchanan. All that I am claiming is that economics can (and should) dispense with meta-preferences.

It is clear, however, that these very radical conclusions depend on one's willingness to accept the rationality assumption (or something similar) that is the basis of the above mentioned result. Many authors (like Sen, for instance) who take a very critical position on rationality axioms in general, would probably decline to do this. For those, however, who accept economic rationality in the traditional sense (and I count Buchanan among them) the conclusion just cited seems to pose a challenge.

2. Reducing meta-preferences to ordinary preferences

When dealing with meta-preferences it is natural to ask about the source of these meta-preferences. As far as I can see it is possible to classify meta-preferences into two categories according to their sources.

In the first category meta-preferences rank ordinary preferences according to purely intrinsic criteria. Examples are the aesthetic value of a preference ordering or the value of its source itself, for instance, divine revelation. This first category employs criteria of ranking that have nothing to do with the outcomes of the various ordinary preference orderings. The preference orderings have certain characteristics which make them more or less valuable in themselves.

In the second category the preference orderings are ordered according to the results they generate. If preference ordering A generates "better" optima than preference ordering B it is meta-preferred to B. (This distinction between the two categories of meta-preference orderings is somewhat reminiscent of Kant's distinction between hypothetical and categorical imperatives in his Critique of Practical Reason, though one should not press the analogy too much.) It is intuitive to conjecture that meta-preferences in the second sense are logically redundant, because it should be possible to mimick them by an ordinary preference ordering on the underlying space of alternatives. This is trivial in the case where the meta-preference ordering is complete, that is, where all preference orderings are comparable. In that case, and
when there is only a finite number of preference orderings one of these preference orderings is "the best" and accordingly the individual's behaviour is fully described by this best preference ordering. Knowledge of the meta-preference ordering gives no additional information. This case, however, is uninteresting. The interesting case arises when several of the preference orderings being ranked are not comparable. This means that the individual will probably alternate between these preference orderings in different situations. In some situations action according to preference ordering A is preferred in others the individual will act according to preference ordering B.

However, if it is outcomes of the various preference orderings that counts one could ask again if it is not possible to construct an ordinary preference ordering that rationalizes or mimicks the meta preference ordering. The question therefore is: Are meta preference orderings really something new? Or are they simply equivalent to a further preference ordering on the underlying set of choice alternatives?

In this section I shall demonstrate that under a certain consistency condition on the meta-preference ordering meta-preferences can indeed be reduced to ordinary preferences. Obvious candidates for such a condition are axioms of the "revealed preference type". In fact, it will be shown in this section that assuming Richter's Strong Congruence Axiom for the choice function generated by a meta-preference ordering allows to reduce this meta-preference ordering to a normal preference ordering on the underlying set of alternatives X.

Let us start with a given set of alternatives X and n preference orderings on X, which are assumed to be complete and transitive. Define $Z = \{\leq_1, \leq_2, \ldots, \leq_n\}$ and assume that there exists a meta-preference ordering $\S$ on Z which is reflexive and transitive but not complete. (If $\S$ were complete the whole problem would be trivial, because then the individual would always act according to the "best" preference ordering in Z.) Let us further assume that the system of admissible budget sets $\mathfrak{I} \subseteq \mathcal{P}(X)$ is finite and that the budget sets themselves are finite too. ( $\mathcal{P}(X)$ is the power set of X.)

It was argued above, that establishing a ranking among preference orderings is equivalent to establishing a preference among the results which these preference orderings generate. This argument can be modeled in the following fashion: It was assumed that $\mathfrak{I}$ is finite. Therefore
\[ \exists \text{can be written as } \{B_1, B_2, \ldots, B_m\}. \text{ For each } B_i \text{ and for each } \leq_j \text{ there is a maximal element } b_{ij} \text{ in } B_i \text{ which is maximal with respect to } \leq_j. \text{ Define the set of all these optima by } B = \{b_{11}, b_{12}, \ldots, b_{1n}, \ldots, b_{m1}, \ldots, b_{mn}\}. \text{ Note that there may be several elements in } B \text{ such that } b_{ji} = b_{jk} \text{ or even } b_{ji} = b_{lk}. \text{ Now define the following system } \mathcal{R} \text{ of subsets of } B:\]

\[ \mathcal{R} = \{\{b_{ji}, b_{jk}\} : b_{ji} \in B, j = 1, \ldots, m, \text{ and } i, k = 1, \ldots, n, i \neq k}\}. \]

This is the system of all pairs of optima, which are located in the same budget set (for all budget sets).

The postulate that the meta-preference ordering must somehow reflect an ordering of the optima that are generated by the underlying preference orderings can now be modeled by assuming the existence of a choice function \( C : B \rightarrow \mathcal{P}(B) \), such that

\[ \leq_i \otimes \leq_j \Leftrightarrow \forall k = 1, \ldots, m: b_{kj} \in C(b_{ki}, b_{kj}). \]

This condition says that if one compares in each set \( B_k \) the \( \leq_i \)-optimum with the \( \leq_j \)-optimum the meta-preference ordering \( \otimes \) entails that the \( \leq_j \)-optimum is preferred.

Now we impose on the choice function \( C \) Richter’s Strong Congruence Axiom.

**Strong Congruence Axiom:** Let \( X \) be a set of alternatives and let \( \mathcal{N} \subseteq \mathcal{P}(X) \) be a system of admissible budget sets. Assume that there exists a choice function \( C : \mathcal{N} \rightarrow \mathcal{P}(X) \). For \( x, y \in X \) define \( y \leq x \) if and only if there exists a \( Z \in \mathcal{N} \) such that \( x \in C(Z) \) and \( y \in Z \). Define \( y \leq^* x \) if and only if there exists a finite sequence of elements \( a_1, a_2, \ldots, a_n \) in \( X \) such that \( y \leq a_1 \leq a_2 \leq \ldots \leq a_n \leq x \). With these definitions \( C \) is said to satisfy the Strong Congruence Axiom if and only if the following implication holds for every \( Z \in \mathcal{N} \) and every pair \( x, y \in X \): \( y \leq^* x \) and \( y \in C(Z) \) and \( x \in Z \) implies \( x \in C(Z) \).
Intuitively speaking this axiom ensures consistence across choice situations. If there exist \( n+1 \) choice situations such that in situation 1 alternative \( a_1 \) has been chosen over \( x \), in situation 2 alternative \( a_2 \) has been chosen over \( a_1 \), ...., and in situation \( n+1 \) alternative \( y \) is chosen over \( a_n \), then in any choice situation that involves a direct comparison between \( x \) and \( y \) alternative \( y \) must be chosen over alternative \( x \).

Let us apply the Strong Congruence Axiom to the system \( \mathfrak{H} \) defined above. This system only contains sets with two elements. Therefore the relation \( \leq^* \) becomes relevant only when there are elements \( b_{lh} \) which are optimal in other sets than \( B_l \) as well, so that for instance \( b_{lh} = b_{jk} \). This means that the \( \leq_{lh} \)-optimum in budget set \( B_l \) is also the \( \leq_{lk} \)-optimum in some budget set \( B_j \). Let us now assume, for instance, that a series of comparisons across sets has established that \( b_{jk} \leq^* b_{lz} \). In such a case the Strong Congruence Axiom requires that there be no "preference reversal" among \( b_{lz} \) and \( b_{jk} \) in the set \( B_l \). If \( b_{jk} \) is chosen from \( B_l \) then \( b_{lz} \) must be chosen from \( B_l \) too.

Applied to our context this Axiom postulates that the ranking of \( \leq_l \)-optima which the meta-ordering establishes via the choice function must be consistent across budget sets. The relation \( \leq^* \) refers to indirect comparisons across budget sets and the implication stated in the axiom demands that these indirect comparisons are consistent.

If the Strong Congruence Axiom applies for a certain choice function \( C \) then a well known theorem by Richter (1966) establishes that \( C \) can be rationalized by a complete and transitive preference relation \( \leq \) on \( X \) in the sense that for every \( A \in \mathfrak{H} \) it is true that \( C(A) = \{ x: x \in A, y \leq x, \text{ for all } y \in A \} \) (Note that \( \leq \) is not identical with the ordering \( \leq \) which occurs in the definition of the Strong Congruence Axiom.)

Applying this result to our choice function on the system \( \mathfrak{H} \) establishes the existence of a complete and transitive preference ordering \( \leq' \) on \( B \). We can extend this preference ordering from \( B \) to \( X \) in the following way:

\( B \) is finite. \( \leq' \) is complete and transitive. Therefore there is a minimum in \( B \) with respect to \( \leq' \). Denote this element by \( a \). Now choose one arbitrary element \( b \in X \setminus B \) and define a new preference relation \( \leq'' \) on \( X \) in the following way:
(1) For all \(x, y \in B\): \(y \leq x\) if and only if \(y \leq' x\).

(2) \(b <' a\).

(3) For all \(x \in X \setminus B\), \(x \not\in b\): \(x \sim b\).

It is easily checked that this preference relation is complete and transitive. Thus we have found a preference relation on \(X\) which is equivalent to the meta-preference ordering \(\leq\) on \(Z\) in the sense that both rankings order the \(\leq_1\)-optima in the same way.

We therefore have proved:

**Proposition:** Let there be a set \(Z\) of \(n\) preference orderings \(\leq_1, \leq_2, \ldots, \leq_n\) on a set of alternatives \(X\) and assume that there exists a meta-preference ordering \(\leq\) on \(Z\) which is reflexive and transitive but not complete. Assume further that the system of admissible budget sets \(\mathcal{I} \subseteq \wp(X)\) is finite and that the budget sets themselves are finite too. (\(\wp(X)\) is the power set of \(X\).) The following holds: If the meta-preference ordering orders the \(n\) preference orderings according to their optima and if the choice function generated by the meta-preference ordering on the system of all pairs of optima exhibits consistency in the sense of Richters Strong Congruence Axiom then there exists a preference ordering \(\leq_{n+1}\) which rationalizes this choice function.

It is clear that the result that was just established will not impress somebody who rejects rationality axioms like the Strong Congruence Axiom altogether. Sen (1993), for instance, argues that axioms like Richter's congruence axiom are far too strict and too simplistic. For Sen every choice situation may be different. So it reflects a simplistic psychology to impose a uniform rationality postulate on the chooser that is supposed to hold across choice situations.

The following section exhibits an example (based on Sen (1993)) which corresponds to this kind of view. In this example meta-preferences *cannot* be reduced to ordinary preferences precisely because they violate consistency across choices. As will be immediately apparent the behaviour depicted in this example is far from "irrational". On the contrary, this behaviour is quite common and quite "reasonable".
3. An example where meta-preferences are not reducible

Imagine you are sitting at a dinner table together with a group of other people. A basket with fruit is passed around. When the basket reaches you, there are two possibilities: (1) Only two fruit are left in the basket, an apple and an orange or (2) only one fruit is left.

In situation (1) you have three options:
- x = you take nothing
- y = you take the apple
- z = you take the orange

In situation (2) you have the choice between x and y (or z, depending on precisely which fruit is left over).

We can model this situation by specifying the basic set of alternatives as $X = \{x, y, z\}$. The admissible subsets taken from $X$ are $\{x,y\}$, $\{x,z\}$, and $\{x,y,z\}$ itself.

Now consider three preference orderings on $X$. $\leq_1$ is the preference ordering of a shy and well educated person. She always prefers to take nothing but if she could choose among the apple and the orange she would choose the orange. Thus $y < z < x$. The second preference ordering $\leq_2$ is the preference ordering of an uncompromising egoist who does not care what good manners dictate. For him $x < z < y$. The third preference ordering $\leq_3$ is a compromise between the first two rankings: $y < z < x$. This person will behave according to good manners as long as the apple is concerned but not with respect to the orange.

Now assume that all these preference orderings are at your disposal. You must choose according to which ordering you want to act. Let us assume, moreover, that you possess a meta-preference ordering which tells you that $\leq_2$ is better than $\leq_3$ and that $\leq_1$ is better than $\leq_3$, but which does not establish any order between $\leq_1$ and $\leq_2$. Thus the "compromising" preference ordering is dominated by both the altruistic and the egotistic preference orderings. At the same time, however, the meta-preference ordering does not discriminate between the
altruistic and the egotistic preference ordering. This meta preference ordering is reflexive and transitive, but not complete.

Now you observe yourself making the following choices: From \{x,y\} you choose x; from \{x,z\} you choose x; from \{x,y,z\} you choose y. That is, as long as there is only one fruit left you act according to \leq_1. But as soon as two fruit are left you switch to \leq_2. When "enough" fruit are left for the others you act as an egoist. Otherwise you act altruistically.

The decisive question now is the following: Is it possible to rationalize your choice behaviour (and therefore your meta preference ordering) by a fourth preference ordering \leq_4 on X? If this were the case it would be logically redundant to introduce a meta-preference ordering of \leq_1, \leq_2, \leq_3 in the first place.

It is clear, however, that your choice behaviour cannot be rationalized by a reflexive, complete and transitive preference ordering on X. We have \(C(\{x,z\}) = \{x\}\) and \(C(\{x,y,z\}) = \{z\}\). This violates a fundamental rationality property called Basic Contraction Consistency or Chernoff's Condition or Condition \(\alpha\) in the literature (see Sen 1993). If this property is violated there is no consistent preference ordering rationalizing the observed choices.

One has to conclude therefore that without additional assumptions meta preference orderings are not logically redundant. They are really "meta".

4. Discussion and conclusions

In closing should note some limitations of the result in section 3. First of all, the result has probably not been proved in the most general form possible. I have used Richter's Strong Concruence Axiom as basis for my proof. The next logical step would be to ask, if the result cannot be proved for weaker rationality assumptions. If so, the case for avoiding meta-preferences in economic theory would be strengthened. If on the other hand the rationality assumptions under which the result can be proved are considered to be very strong or very unrealistic, the opposite holds. But then a dilemma arises: If the rationality conditions are considered to be too strong to give a realistic picture of human behaviour then it must be asked what is to be substituted for these rationality assumptions. Sen, for instance, explicitly rejects the use of axioms like the Strong Congruence Axiom. For Sen choices like the ones
described in the example of the last section (which is derived from Sen) are frequent. Every choice situation may be different. So it reflects a simplistic psychology to impose a uniform rationality postulate on the chooser that is supposed to hold across choice situations.

One asks oneself, however, what should be substituted for the rationality postulates. The idea of meta-preferences is not very useful for analysis as long as one is not able to specify how and why a certain meta-preference ordering leads an individual to act according to preference ordering \( \preceq i \) in a certain situation and not according to preference ordering \( \preceq j \).

Maybe that psychology can help. It seems, however, that psychologists reject the whole model of human choice developed by economists completely. If one is prepared to resort to psychology at all, why not be consequent and throw the whole economic model of preference orderings overboard, instead of using "hybrid" notions like the notion of meta-preferences?

Thus either way the use of meta-preferences in economic theory seems questionable.

But what of Buchanan's argument, that meta-preferences can explain why people give themselves rules? It seems quite plausible that at some point in their lives people take stock and contemplate the question what type of person they are at the moment what type of person they would like to become instead. It also seems plausible that people impose rules on their behaviour in order to force themselves to behave like the person they want to become.

There is also a further aspect of this process which Buchanan explores mainly in his essay "Natural and Artifactual Man" (1979). This is the aspect that the possibility of self-selection and self-education is perhaps the only way we can give meaning to the notion of individual responsibility.

I have no intention to question the importance of these ideas. I have (unfashionable) doubts, however, that economic analysis has the right tools to deal with these age old problems. It seems to me that the result in section 3 of this paper gives economists the possibility to stick to their turf and to leave these problems to philosophers, psychologists or sociologists. It may be true that people give themselves rules. But then (as was shown) they act as if they simply followed a normal preference ordering. Thus, economists have no need to consider rules as a separate entity. In my view economists who are prepared to accept rationality postulates like
Richter's Strong Congruence Axiom should apply Ockhams razor and weed meta-preferences out of economic theory.

**Appendix: Meta-Preferences as preferences over "meta-commodities".**

In this paper I have dealt with meta-preferences only in the sense of a ranking of rankings. One might think of a second concept of meta-preferences, however, which up to now is not very prominent in the literature, but wich Buchanan briefly mentions in *The Reason of Rules* (Chapter 5, Part I, Section II) although without discussing it in more detail.

...This concept considers meta-preferences as a ranking of very basic commodities, that is, commodities which are even more basic than Gary Becker's famous Zs. Buchanan calls them the "ultimate Zs" (1985: 69). It is these ultimate Zs that make a life plan "intersting", "good", "rewarding" or "happy".

Buchanan is not very specific about these ultimate Zs. In order to fix ideas I propose to illustrate the ultimate Zs with the help of Jeremy Bentham's (1789) list of fifteen simple pleasures and simple pains. (For the following proof the precise content of the ultimate Zs is irrelevant however.) Bentham's fifteen pleasures are the pleasures of senses, riches, address, friendship, good reputation, power, piety, benevolence, malevolence, knowledge, memory, imagination, hope, association and relief of pain. The fifteen simple pains are the negative counterparts of the fifteen pleasures. Most entries in this "inventory" are self-explanatory (like the pleasures of the senses, of riches, of friendship, good reputation, and power). Others deserve a word of explanation. The pleasure of address consists in the feeling of dexterity that somebody has, for instance, when he plays a musical instrument well or when he is coping with a difficulty. Pleasures of benevolence or malevolence arise when we see that people which we like are happy or when we see that people which we dislike are unhappy. Pleasure of piety is felt when somebody believes that he leads a life that is agreeable to god. The pleasure of knowledge is the pleasure that we feel when we learn new things or make intellectual discoveries. The pleasure of memory arises when past pleasures are remembered.
Pleasures may also arise when pleasures are imagined (pleasure of imagination) are hoped for (pleasure of hope) or associated with certain objects (pleasure of association).

The 15 basic pleasures and pains contained in Bentham’s list can be experienced in different degrees. Let us assume that these degrees are measurable and let us denote degrees or quantities of the fifteen simple pleasures and pains by \( b_1, \ldots, b_{15} \), and quantities of the fifteen simple pains by \( b_{16}, \ldots, b_{30} \). I shall call the set of all possible tuples \((b_1, \ldots, b_{30})\) the Bentham space and one particular tuple \((b_1, \ldots, b_{30})\) a Bentham bundle. Thus, a Bentham bundle is a certain combination of feelings (pains as well as pleasures), which are experienced to the degree \( b_1, \ldots, b_{30} \). I propose to take these Bentham bundles as a model of Buchanan’s "ultimate Zs".

All pleasures and pains that are not elements of Bentham's inventory are combined pleasures. This means that several of the simple pains and pleasures act together to produce this particular pleasure. A self-cooked meal may create pleasures of the senses, of riches (if it uses expensive ingredients) and of companionship (if it is prepared and consumed with another person). Every combined pleasure can be described by a Bentham bundle and every Bentham bundle is a combined pleasure.

To a certain extent one can think of Bentham bundles in the way one thinks about bundles of market goods in elementary consumer theory. In particular one may assume that there are Bentham bundles which are indifferent to each other. Bentham himself is known to have said: "Quantity of pleasure being equal, pushpin is as good as poetry". (Mill, 1968, 54.) Using the notion of Bentham bundles this can be interpreted in the sense that the Bentham bundle generated by pushpin is indifferent to the Bentham bundle generated by poetry. Likewise somebody may be willing to trade off a Bentham bundle that contains more sensual pleasure for another Bentham bundle that contains more pleasure of, say, power. Thus we may define indifference sets in the Bentham space. But after this step is taken the similarity with basic consumer theory ends. Each Bentham bundle contains pleasures as well as pains. One cannot therefore assume that in general agents prefer "more to less". That is, we cannot assume that \( b_i \leq b'_i \) for all \( i = 1, 2, \ldots, n \) entails that the Bentham bundle \((b_1', \ldots, b'_n)\) is preferred to \((b_1, \ldots, b_n)\). (Here \( \leq \) refers to the normal ordering of the real numbers.) We cannot rescue
this condition by simply assigning a negative sign to the $b_1, \ldots, b_{30}$. It may well be that some persons do not prefer Bentham bundles which give them more pleasures and less pains at the same time.

It may be objected that this is not true for "normal" people. This objection, however, raises the question of what constitutes normality. Psychologists tell us that it is quite "normal" that people do not always follow the principle "less pain more pleasure". Ronald Coase has stated that "there is no reason to suppose that most human beings are engaged in maximizing anything unless it be unhappiness, and even this with incomplete success." (Coase 1988, 4).

Probably everybody orders the bundles in the Bentham space according to his personal philosophy, his education, experiences, neuroses, custom, social norms etc.

For this reason I shall assume that only two properties of the usual preference orderings can be transferred to meta-preferences, namely completeness and transitivity. Thus meta-preferences in the sense of Bentham are considered to be a complete and transitive preference ordering over the Bentham space. In addition, I assume that the Bentham space is finite. Thus agents can only distinguish between Bentham bundles which are "different enough". This means that there are certain thresholds of perception which must be reached before an agent realizes that two bundles are different. Furthermore I shall assume that the space of commodities and the space of market goods are also finite.

What is the logical relation between meta-preferences in this sense and commodity preferences (that is, preferences over the Beckerian $Z$s)? In order to deal with this question one has to realize that for every Bentham bundle $b = (b_1, \ldots, b_{30})$ there exists a set $C(b)$ of commodity bundles $z = (z_1, \ldots, z_m)$ each of which is able to generate $b$. Thus $C(b) = \{z: z$ generates $b\}$. Denote the set of all commodity bundles $z = (z_1, \ldots, z_m)$ by $X$ and the system of all sets $C(b)$ by $\mathcal{C}$. This construction allows us to identify Bentham bundles with sets of commodity bundles. It was argued that it is plausible to assume that each individual orders Bentham bundles by a complete and transitive preference ordering. Denote this meta preference ordering by $\preceq^*$ and the preference ordering on the underlying commodity space $X$ by $\preceq$. 
It seems plausible that all commodity bundles which generate the same Bentham bundle should be indifferent to each other. Moreover, if two commodity bundles generate two different Bentham bundles which, however, are indifferent with respect to the meta preference ordering \( \preceq^* \) then these two commodity bundles should be indifferent with respect to \( \leq \) too. If these two postulates are accepted the meta preference ordering \( \preceq^* \) induces a preference ordering \( \preceq \) on the commodity space \( X \) by the definition \( x \prec y \) if and only if there exists \( C(b) \) and \( C(b') \) such that \( x \in C(b) \), \( y \in C(b') \), and \( b \prec^* b' \). The following proposition states that \( \leq \) is unique in the sense of being the only preference ordering on the commodity space which is compatible with \( \preceq^* \).

**Proposition:** Let there be given meta-preferences \( \preceq^* \) on the Bentham space which are are complete and transitive. Define \( x \prec y \) if and only if there exists \( C(b) \) and \( C(b') \) such that \( x \in C(b) \), \( y \in C(b') \), and \( b \prec^* b' \). Then the ordering \( \leq \) on \( X \) defined in this way is unique. Moreover, it is true that (1) \( z \sim z' \) for all \( z, z' \in C(b) \), for all \( b \), and (2) that \( z \sim z' \) if \( z \in C(b) \) and \( z' \in C(b') \) with \( b \sim^* b' \).

**Proof:** Start from the definition of \( \prec \) and define \( \sim \) in the usual way by \( x \sim y \) if and only if neither \( x \prec y \) nor \( y \prec x \). This leads to the condition \( x \sim y \) if and only if there exist Bentham bundles \( b \) and \( b' \) such that \( b \sim^* b' \) and \( x \in C(b) \) and \( y \in C(b') \). The \( C(b) \), however, are disjoint by assumption. (One commodity bundle cannot generate two different Bentham bundles.) This proves conditions (1) and (2) in the proposition. It is also easily checked that \( \leq \) is complete and transitive. To establish uniqueness assume that there is a second preference ordering \( \prec', \prec' \neq \prec \), satisfying the condition \( x \prec' y \) if and only if there exists \( C(b) \) and \( C(b') \) such that \( x \in C(b) \), \( y \in C(b') \), and \( b \prec^* b' \). \( \prec' \neq \prec \) implies that there exists at least one pair of commodity bundles \( z \) and \( w \) such that \( z \prec' w \) but not \( z \prec w \). This is equivalent to \( z \prec' w \) and \( w \leq z \). Therefore there must be Bentham bundles \( b \), \( b' \), \( a \), \( a' \), such that \( w \in C(b) \), \( z \in C(b') \), and \( b' < b \) and \( w \in C(a) \), \( z \in C(a') \), and \( a \leq a' \). It was assumed, however, that the \( C(b) \) are disjoint. Therefore we must have \( a \sim b \) and \( a' \sim b' \). Substituting this into \( b' < b \) leads to \( a' < a \), a contradiction to \( a \leq a' \). Q.E.D.
The proposition which was just proved establishes that there is only one preference ordering on the commodity space that is compatible with a given set of meta-preferences on the Bentham space. Because meta-preferences on the Bentham space can be translated in a unique fashion into preferences on the commodity space not much is gained by assuming meta-preferences in the first place.

References