Discussion Paper No. 1025*

CASE–BASED CONSUMER THEORY

by

Itzhak Gilboa+

and

David Schmeidler++

February 1993

* We are grateful to many people for conversations, comments and references. In particular, we would like to thank Eva Gilboa, Ehud Kalai, Morton Kamien, Elon Kohlberg, Hanan Polansky and participants of seminars at Tel-Aviv University and the Universitat Autònoma de Barcelona. Special thanks are due to Enriqueta Aragones for many specific comments on an earlier version of this paper.

+ KGSM-MEDS, Northwestern University, Leverone Hall, Evanston, IL 60208.

++ Departments of Economics and of Statistics, Tel Aviv University, Tel Aviv 69978, Israel; Department of Economics, Ohio State University, Columbus, OH 43210-1172.

Abstract

The neo-classical theory of consumer behavior, while a powerful tool, suffers from some well-known flaws. Specifically, it assumes a highly, often unrealistically rational utility-maximizing consumer and sheds little light on the dynamic nature of consumption decisions.

In this paper we make some preliminary steps towards an alternative theory of consumer choices, restricted to the case of repeated "small" decisions. We assume that the consumer is choosing among products, rather than bundles, and that (s)he is a "case-based decision maker". In particular, such a consumer is not necessarily "optimizing" and may be "satisficing" in the sense of March and Simon (1958).

The aggregation of choices among products implicitly defines a choice of a "bundle". It turns out that if the "aspiration level" of the consumer is relatively low, (s)he tends to be satisficed and may choose a "corner" solution, which is not necessarily "utility maximizing" in the classical sense. If, however, the aspiration level is relatively high, the consumer keeps switching among the products, and their relative frequencies converge to an interior point in the bundles space, as suggested by the classical theory under the assumption of convex preferences.

Furthermore, in our model we find that the "utility" of a product is closely related to the (limit) relative frequency with which it is consumed. Thus this model offers a new definition of a product's "utility", as a cardinal measure of desirability.

To study the effect of changes in market conditions, we propose to incorporate a product's price directly into its utility. According to this view, the consumer does not consider the "utility" of each product (or bundle) as separate from the budget constraint. Rather, the fact that a certain product is expensive is implicitly assumed to alter the experience of consuming it. It follows that the consumer's reaction to price changes is "immediate", and does not require to (implicitly) solve the new optimization problem. Moreover, such a consumer

may well respond to price changes without necessarily maximizing his/her utility subject to the budget constraint.

Next we study the relationship between the weight attached to the price in the (linear) evaluation of a product and the budget constraint. We show that, under certain reasonable conditions, the total expenditure is a decreasing function of this parameter and it follows that there exists a unique value for this parameter which balances the consumer's budget.

Finally, we introduce the notion of the "potential" of the utility function, which is akin to the neo-classical utility function. We model substitution and complementarity effects, which are related to the cross derivatives of the potential.

1. Introduction

The (neo-)classical theory depicts a consumer as a rational agent, who is confronted with known products, prices and budget constraint. The theory suggests that such a consumer has a utility function over the space of product bundles, and that (s)he would choose a feasible bundle so as to maximize this function.

A literal interpretation of the theory would lead us to consider agents who spend their time in supermarkets' aisles with pocket calculators, computing the utilities of various consumption bundles. Since this account is somewhat at odds with everyday observations, most economic theorists would prefer the more metaphorical interpretation, according to which consumers only behave *as if* they solved an optimization problem. True to the logical-positivistic tradition, they would argue that a theoretical construct such as "utility" gains its meaning only through the way it is used to explain and predict observable economic phenomena, and that the reasonability of the utility-maximization paradigm should only be judged based on its behavioral implications.

While this interpretation is not as preposterous as the literal one, it is still not quite convincing as a general theory of consumer behavior. For one, it is seldom the case that one knows all the available products and their prices. The classical way to deal with this difficulty, namely analyzing the problem as decision under uncertainty within the Bayesian paradigm, is hardly satisfactory since it postulates that the decision maker behaves as if (s)he knew all the conceivably available products, had beliefs on their simultaneous materialization (including their cost), and maximized expected utility in this set-up. The amount of knowledge required to conduct this optimization consciously is daunting. On the other hand, it is never quite clear how would a consumer, who does not possess this knowledge, come to behave as if (s)he did.

Yet even when the products, prices and budget may be assumed known, it is somewhat unlikely that consumers will decide to purchase the bundle which is, indeed, "optimal" according to their utility. In a one-shot problem this is, of course, more than likely: it is tautological. However, it is far from obvious why we may assume that a single utility function may explain consumer choices under various circumstances. Indeed, the theory provides an axiomatic derivation of the utility (to be maximized) from a binary preference relation (or a choice correspondence.) Furthermore, the axioms such a relation has to satisfy all seem rather plausible *when a preference relation in considered.* But the very notion of preferences over product bundles is a theoretical idealization. People choose among products, not among bundles. It may be theoretically convenient to map a consumer's choice between, say, two products at a store to the implied choice between the corresponding bundles (containing these products and the rest of the consumer's assets), but convenience does not justify the assumption that this implied preference relation is nicely behaved. Differently put, the neo-classical theory appears to be very reasonable when formulated in its own language. However, this language can hardly be considered very intuitive.

Another (and to an extent related) weakness of classical consumer theory is that it considers the consumer's choice problem in isolation, as a history-independent optimization problem. In reality, however, consumers' choices are likely to be context-dependent; the effect of price changes, for instance, can hardly be assumed independent of current consumption. This point is doubly relevant in view of the previous one: since the objects of consumption are products, bundles may be said to be "consumed" by aggregation of product consumption over time. Yet the dynamics of consumption is hardly modeled in the neo-classical theory.

Many authors have criticized economic theory in general, and the neo-classical model in particular, along similar lines, mostly for the rationality assumptions it relies on. Perhaps the best-known and most influential attack on the "rational agent" paradigm is Simon's (1957) notion of "bounded rationality" and the suggestion of March and Simon (1958) that rather than optimizing, people are merely "satisficing". According to this theory, a decision maker has an "aspiration level", and as long as his/her choices attain that level (say, of "utility"), he/she will not even attempt to further optimize the decision made.

While "bounded rationality" models have been studied quite extensively in economic and game theory, it appears that none of the theories suggested is general and simple enough to serve as a foundation for economic theory at large. Ironically, many of the models, which attempt to incorporate computation costs into the optimization problem, end up being even more computationally demanding than the original models they were supposed to refine.

In Gilboa-Schmeidler (1992) we propose a new theory of decision making, whether under uncertainty or not, which submits that a decision maker chooses an act which is "best" according to its past performance in similar decision problems. We dub it "Case-Based Decision Theory" (CBDT)[1]. As opposed to classical decision theory, CBDT does not assume, neither explicitly nor implicitly, that the decision maker "knows" the utility of any choice apart from those which were actually experienced. The formal model of case-based decisions naturally gives rise to the notions of "satisficing decisions" and "aspiration levels". Furthermore, since past "cases" are at the heart of CBDT, it is naturally history- or context-dependent.

At first glance, then, CBDT seems to hold a promise of modeling consumer behavior in a way that, at least for some applications, will be more realistic than classical decision theory. In the present paper we study a consumer who is a case-based decision maker and whose objects of choice are products (rather than bundles). We confine ourselves to the analysis of repeated consumption choice problems. While many decisions, such as the purchase of a house, are "big" in the sense that they are made only once in (and for) a long time interval, many other decisions, such as whether to dine in or out, are "small" in that they are frequently repeated under rather similar conditions. For such decisions, the aggregate of "small" consumption choices over time implicitly defines a "choice" of a point in the bundles space.

We find that the dynamics of the decision-making process, and its dependence on memory/context/history provide some insights into consumer preferences and behavior. The theory presented here -- which we nickname "Case-Based Consumer Theory" (CBCT) -- is restricted in many ways. As a

---

1 The name should bring to mind the theory of "Case-Based Reasoning" of Riesbeck and Schank (1989). They argue that the main reasoning technique people use, especially in novel situations, is reasoning-by-analogy to previous cases, and they suggest some implementation of this idea to artificial intelligence systems.

While the term "case-based" is borrowed from Riesbeck and Schank, the basic idea is, of course, much older. Some of the more recent references in the economic literature would include Keynes (1921) and Selten (1978). However, Riesbeck and Schank appear to be the first to suggest case-based reasoning as the main reasoning technique people employ.

matter of fact, it should better be classified as a preliminary exercise rather than as a full-fledged theory. Yet we believe that it is phrased in more intuitive language, and its rationality postulates are by far less demanding than those of the neo-classical theory. It is therefore more psychologically plausible, and, within its scope of application, hopefully also more realistic than the classical theory.

A slightly more detailed outline of this paper may be helpful at this point. Section 2 provides the decision-theoretic background. We devote sub-section 2.1 to an overview of CBDT. It contains all the reader may need to know about CBDT for the understanding of the present paper. Some theoretical aspects of CBDT as applied to decision under certainty are discussed in sub-section 2.2.

In section 3 we discuss a simple example which conveys the intuition for some of our main results: we consider a consumer who solves the same choice problem over and over again, and study the pattern of consumption choices (s)he makes. We find that, if the aspiration level of the individual is low, (s)he is easily satisficed and may select a "corner" solution as an aggregate choice, which may not be "optimal" in the usual sense. On the other hand, should the aspiration level be relatively high, the consumer would keep switching among the various products. We later prove that the relative frequencies with which the various products are consumed converge to a limit, which will be an "interior" solution in this case. One way in which this result can be interpreted is that tastes, or utilities, are inherently context-dependent. An individual with a low aspiration level would tend to like a product more, the more it was consumed in the past, thus explaining habit-formation. On the other hand, a high aspiration level implies that the consumer would appear to be "bored" with products (s)he recently consumed, and would seek change. Hence, high aspiration levels seem to be related both to boredom aversion (in the small) and to convex (classical) preferences (in the large.)

Section 4 generalizes this examples and states some results. It may be viewed as studying CBDT under the special assumptions that no uncertainty is involved, and that all decision problems are equally similar to each other. For this set-up, and under the assumption of a high aspiration level, we show that the (instantaneous) utility function, which is axiomatically derived in CBDT, can also be re-interpreted (and alternatively axiomatized) by its relationship to the consumption frequencies of the various products. More specifically, the ratio of two products' utilities is the inverse of the ratio of the frequencies with which they will be consumed. As opposed to classical consumer theory, which only allows for an ordinal utility function (over bundles) to be derived from many one-shot observed preferences, we derive a cardinal utility (over products) from a single sequence of repeated choices.

While consumer choices are the main goal of this study, there is nothing in the above-mentioned sections which is specific to the consumer's problem. In particular, prices and money are not explicitly mentioned. This part of the paper may thus be viewed as a general theory of repeated choice under conditions of certainty.

We further specify the model in section 5. There we propose to study consumer's reaction to price changes by incorporating prices into the (instantaneous) utility derived from consumption. The intuition behind this model is that a boundedly-rational consumer does not solve a utility-maximization problem subject to a budget constraint, nor can (s)he be assumed to behave as if (s)he were. Rather, the consumer makes decisions on a daily basis, and is rational enough to know, say, that a higher price of a given product means that less options will be available to him/her in the future, should (s)he purchase this product. This knowledge is so internalized, that one may actually assume it is part of the utility function itself. That is, the experience of consuming the good is inseparable from the knowledge that so many dollars were sacrificed for it. A commodity which may be desirable if free may not yield enjoyment if consumed at an exorbitant price.

As shown in sub-section 5.1, this model allows one to analyze the effect of price changes on consumption, without assuming, explicitly or implicitly, that a new optimization problem is solved whenever prices vary. In sub-section 5.2 we consider the process by which consumer's choices are eventually matched to the available budget. We assume that the prices are linearly incorporated into the instantaneous utility function, and study the relationship between the "weight" of the price in this function (the "value-of-money" parameter) and the total expenditure. We implicitly assume that every so often the consumer compares the amount of money (s)he spends to the available income, and that the value-of-money is correspondingly updated.

It turns out that, in general, the consumer may end up spending more money if (other things being equal) the value-of-money is increased. However, this anomaly is restricted to consumers who are somehow exceptional. We define a consumer to be "typical" if his/her taste agrees with market prices in the following sense: if one product is more expensive than another, then, *if both were free,* the consumer would find the first also more desirable and consume it more frequently than the second. We show that for a typical consumer the expenditure function in monotonically decreasing in the value-of-money, and thus there exists a unique value for this parameter which balances the budget, and it can be found by trial-and-error.

Section 5 presupposes a given budget for the repeated choice problem. While many choices may be viewed as "small", repeated ones, few of them would have a pre-determined budget. In fact, the allocation of the budget to various classes of products -- such as "food", "entertainment" and so on -- is an integral part of the consumer's decision.

Section 6 is therefore devoted to a discussion of the generalization of the model to the case of several (repeated) decision problems which occur "simultaneously". To be precise, we describe a model in which at every stage only one decision problem is being solved, yet these problems may belong to several different categories, defined by the available choices. We omit the formal model (which is a straightforward generalization of the previous sections) and provide a simple example, which also shows that the process by which the value-of-money parameter is updated may change the budget proportion spent on the various types of products.

In section 7 we introduce the notion of the "potential" of the utility function. We provide an interpretation according to which this mathematical construct, which is derived from the primitives of our model, measures the "overall well-being" of the consumer, and is thus close in spirit to the neo-classical utility function.

Section 8 introduces similarity among acts as a way to model substitution and complementarity effects. It turns out that the similarity function between products is related to cross "derivatives" of the potential, in a way that parallels the neo-classical theory. Furthermore, we show that under an appropriate symmetry assumption one may "rotate" the bundles space in such a way that its axes are "basic commodities", among which neither substitution nor complementarity effects exist.

Finally, section 9 concludes with some comments. In particular, it attempts to outline the applicability of CBCT and to classify the cases in which it may be more successful than the (neo-)classical theory.

2. Decision-Theoretic Background

2.1. Case-Based Decision Theory -- An Overview

A full description of CBDT is certainly beyond the scope of this paper. The reader is referred to Gilboa-Schmeidler (1992) for detailed exposition, axiomatizations, variants and theoretical discussions of CBDT, as well as for comparisons of it to expected utility theory for decision under uncertainty. In this section we will only provide a very sketchy outline of CBDT, which will hopefully suffice for the understanding of the following sections.

The primitives of CBDT are:

P -- a set of decision ***problems***

A -- a set of available ***acts***

R -- a set of possible ***results*** (or outcomes)

The set of ***cases*** is defined to be

$$C = P \times A \times R$$

That is, a "case" is a triple (p,a,r), where p is the problem encountered, a is the act chosen by the decision maker, and r is the result that was obtained in this case. We will assume that at any given point in time, a decision maker is equipped with some memory M, which is simply some subset of cases, and which will be interpreted as the set of problems the decision maker can remember.

CBDT postulates two main theoretical terms -- "utility" and "similarity". As in the classical decision theory, the utility measures the desirability of the results, and is thus a function

$$u: R \rightarrow \Re$$

The notion of "similarity" is new, and corresponds in many ways to that of "subjective probability" in expected utility theory. Similarity measures the extent to which one decision problem is similar to another; that is, it is a function

$$s: P \times P \to [0,1]$$

Finally, we may describe the decision rule which is the heart of CBDT: Suppose that a decision maker, characterized by the utility u and the similarity s is faced a decision problem p, while his/her memory is $M \subseteq C$. Then every possible act $a \in A$ is evaluated by the functional

$$U(a) = \sum_{(q,a,r)\in M} s(p,q)u(r)$$

-- and the decision maker will, according to CBDT, choose a maximizer of U.

A few comments are in order. First, notice that for two distinct acts $a,b \in A$, $U(a)$ and $U(b)$ are summations over *disjoint* sets of cases. Furthermore, for some acts this summation may be over an empty set, in which case its value is defined to be zero. This value is going to play a major role in the theory: one may think of it as the decision maker's "aspiration level". To be precise, this is the "default" (utility) value the decision maker seems to be attaching to an act that was never tried in the past (i.e., for which there are no cases in memory.) If certain acts obtain higher U-value than zero, the decision maker is "satisficed", and will continue to choose among them without trying new acts and *without trying to maximize u*. Once all the acts that were tried in the past turned out to be unsatisfactory -- that is, to have negative U values -- then the decision maker will choose a new act (assuming such exists), were the choice among these will be arbitrary.

One of the main features of CBDT is that it does not require the decision maker (DM) to "engage" in hypothetical reasoning: as opposed to expected utility theory, where the very definition of an "act" involves hypothetical statements such as "If state ω occurs then I get r", in CBDT all the DM is required to "know" is the history of cases which *actually happened* and the utility he/she *actually experienced.* (The terms "engage" and "know" above are within quotation marks since one may choose a purely behavioral interpretation of the theory, according to which the DM does not have to "know" or to reason about anything.)

Without details we mention here that the decision rule of CBDT, together with the theoretical terms "utility" and "similarity" may be axiomatically derived from preferences, in a way which parallels the axiomatic derivations of "utility" and "probability", combined with the expected utility formula, in models such as Savage's (1954). (See Gilboa-Schmeidler (1992) for one such axiom system, as well as additional discussions.)

The notion of a "case" will sometimes be interpreted in a broader fashion. For instance, a case in a decision maker's memory need not necessarily have been experienced by the same DM. It may well be a "story" told by someone else. Furthermore, it need not be a real case -- it may be a hypothetical one, reflecting the DM's knowledge (or belief) about what would have occurred as a result of a possible choice.

Finally, let us briefly mention two variants of the basic CBDT model:

-- *Averaged similarity* in which one uses a functional similar to U above, with the sole difference that for each act $a \in A$, the similarity coefficients $s(p,q)$ are normalized to sum up to 1;

-- *Act Similarity* according to which acts may also be similar to each other, and the evaluation of an act a depends not only on its own performance in the past, but also on that of similar acts. Thus, on top of the utility u and the problem-similarity s_P functions, one assumes an act-similarity function s_A such that an act a is evaluated by

$$U'(a) = \sum_{(q,b,r)\in M} s_P(p,q) s_A(a,b) u(r)$$

In Gilboa-Schmeidler (1992) we axiomatize the first variant and discuss the conceptual difficulty with the axiomatization of the second.

2.2. On Certainty

Case-based decision theory does not formally distinguish between decision under certainty and under uncertainty. Yet it has a special structure in the case of certainty: for each act $a \in A$ there exists a unique $r_a \in R$ such that only cases of the form $(\cdot,a,r_a)$ may appear in the decision maker's memory with a as their second component. Thus one may identify an act with its outcome. In the case of consumer's decisions, "chicken at the restaurant", for example, is both the act chosen by the consumer and the outcome that would result from it. We will therefore omit the definition of the set of outcomes in such a model, with the understanding that it coincides with the set of acts. We will sometimes refer to the elements of this set as "actcomes", to remind the reader of their dual role.

Note that in this set-up, every act a will lead to the outcome $r_a = a$ and to the utility level $u(a)$. One may therefore ask, why isn't the decision maker maximizing this utility function, or, if (s)he tends to be satisficed -- why doesn't (s)he choose the "best" (u-maximizing) act at least in case all these utility values were found negative in the past (that is, none of the options is "satisficing".) In short -- what is the meaning of the "instantaneous utility function" u and the functional U which is maximized by the individual?

Obviously, we should not think of the function u as a neo-classical utility function. It is a theoretical construct which is derived from preferences in the context of CBDT and U-maximization. However, it can be validly interpreted in at least two ways.

According to the first interpretation, decision problems which would classically be categorized as optimization problems in which no uncertainty is present, are still viewed as problems under subjective uncertainty. Case-based decision makers (CBDM's) do not have the "knowledge" required for perfect optimization. From the viewpoint of CBDT, even a "simple" choice problem involves uncertainty about one's own utility function. Differently put, CBDM's "know" the options available to them, but they never "know" to what extent they will like each option. Thus they are always faced with subjective uncertainty, and they use their experience, to the degree it exists, in evaluating the various acts, as suggested by U-maximization.

The second interpretation allows CBDM's to "know" the "utility" function u for those acts which have been tried in the past. Yet this does not imply that the "instantaneous" utility summarizes all the relevant information. Rather, according to this interpretation, tastes, and thus decisions, are intrinsically context- or history-dependent. Hence the function u should be thought of as some derivative (with respect to the time axis) of the "real" utility U, which, in turn, is the aggregate of u values. That is, the utility function to be maximized is U, but, alas, it is constantly changing. What *is* constant over time is its "conditional derivative" u. To be more precise, consider the case of a constant similarity function $s(\cdot,\cdot)=1$. Then $u(a)$ is the rate of change of $U(a)$ in case the actcome a is chosen, while this rate is zero in case another actcome was chosen. Viewed thus, there is little wonder that u is not maximized by the decision maker.

The first interpretation highlights the subjective uncertainty which is inherent to all decision situations. It therefore emphasizes the role of memory as a source of information. The second one argues that memory (or history) directly affects the nature and "utility" of present experiences. The two are not incompatible, nor are they the only ones conceivable. Indeed, in the next two sections we will propose yet another one.

3. Short- and Long- Run Consumer Preferences: An Example

The dynamic nature of CBDT therefore suggests a new look at consumers' preferences: rather than assuming these as primitives of economic theory, they may be derived from U-maximization. More specifically, we would like to consider a consumer who makes "small" consumption decision on a "daily" basis, whose long-run (say, "annual") aggregation gives rise to the overall "preferences" as assumed by the classical theory.

Let us consider a simple example. An individual has to make a daily choice regarding his/her dinner. The available options are $\{beef, fish, chicken\}$. Let us assume that the individual has no "intrinsic" preferences over this set, so that the u-value of all actcomes is identical. Since every day the decision maker is faced with the same problem, the similarity function is constant as well. That is, the decision problems are indexed by time (days): 1,2,...,T=365 , and the similarity function is given by $s(i,j)=1$.

Now let us compare two individuals: individual 1 ("he") is easily satisficeable. Recalling that one's aspiration level is normalized to equal zero on the u-scale, we set

$$u^1(beef)=u^1(fish)=u^1(chicken)=1\ .$$

The first choice individual 1 makes may be arbitrary, but from then on he will stick to the same dish throughout the year. To the extent that one can represent the aggregate preferences of this individual in the standard theory (over the space of bundles $\Re^3$), all corner points are equally desirable, though strictly preferred to interior points. Thus the standard convexity assumption appears to be violated.

Let us now turn to individual 2 ("she") who, by contrast, is not satisficed by her meal choice. Having a high aspiration level, her utility function may be modeled as

$$u^2(beef) = u^2(fish) = u^2(chicken) = -1.$$

Let us follow individual 2's choices. The first day of the year she makes an arbitrary choice, say she chooses $beef$. On the second day, $beef$ has a negative U-value, while $fish$ and $chicken$ are still set at their default value, i.e., at zero. Say the arbitrary choice between them yielded $fish$, and on the third day only $chicken$ has a nonnegative (zero) U-level, and is chosen last. It is easy to see that U-maximization may yield the pattern $(beef, fish, chicken, beef, fish, chicken, ...)$. Furthermore, any U-maximizing sequence of choices is some concatenation of $(beef, fish, chicken)$ with itself and/or its permutations. Therefore, on the aggregate level the point $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ seems to be chosen out of the simplex of meal-mixtures.

This highly idealized example suggests a few principles which will turn out to be generalizable:

- The sequence of choices made by the consumer in consecutive periods is quite orderly. In particular, the vector of relative frequencies (of each choice) has a limit, and it therefore makes sense to discuss "the choice" made by the individual. That is, one may refer to this limit as the individual's "aggregate choice" in the 3-dimensional Euclidean space (or rather, the 2-dimensional simplex.)
- A low aspiration level -- or, equivalently, positive utility values for the products -- will correspond to the choice of an extreme point (in the 2-dimensional simplex), in spite of the fact that all products are "equally desirable". That is, the individual may choose only $beef$, only $fish$, or only $chicken$, but not some (non-trivial) convex combination of the three.

By contrast, a high aspiration level -- or negative utility values -- will correspond to the choice of an interior point in the product space. (This will turn out to be true even if the three products are not "intrinsically" equivalent.) Thus, the observed aggregate behavior predicted by our model given high aspiration levels is equivalent to that of the classical model with convex preferences. Differently put, the present model may be suggested as case-based foundations of micro-economic theory (for the case of repeated choice), where convexity of preferences in the large is derived from high aspiration level in the small.

- The aspiration level, in the context of this model, can also explain such phenomena as habit formation and routine-seeking versus boredom and change-seeking consumption choices. Should the aspiration level be low (positive utility values), the individual would seem to place little value on change for its own sake, and would not be easily bored. To the contrary, (s)he would appear to like a product more, the more (s)he has consumed it in the past. Thus habit formation may be partially explained by low aspiration levels.

On the other hand, high aspiration levels (negative utility values) may explain "boredom aversion": the very fact that a certain product was consumed more often than another will make it less desirable.

The fact that individual choice does not seem to be always the same in what appear to be the same circumstances is typically explained by postulating stochastic choice or changing preferences, or by defining utility over *sequences* of alternatives. Each of these explanations may be indispensable in some cases, but they are also allowing for a huge range of possibilities. For instance, the literature on changing preferences may well explain habit formation and boredom aversion as special cases, but in its general form the model does not allow us to infer much about the future utilities from past choices. Similarly, making sequences of alternatives, rather than single ones, the object of choice may account for these well-known phenomena, but also for a variety of less known ones. It therefore seems to be an advantage of the present model that it naturally gives rise to these behavior modes, without making unnecessary generalizations.

- Consider the case of a high aspiration level. In the example above, the utility of all products is the same, and thus simple symmetry considerations may convince us that the aggregate choice should be $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$. However, a stronger statement will turn out to be true in general: the (instantaneous) utility indices of the products will be intimately related to their limit frequencies. Specifically, for any two products, the ratio of their utility indices will turn out to be the inverse of the ratio of the limit frequencies with which they are chosen. Thus, "the utility of a product" will have a much more concrete meaning in our model than in the classical one, and it will also be cardinal for that reason. (To be precise, given the aggregate choice, the utility will be fixed up to a multiplicative scalar.) Notice, however, that these utilities are only defined for products which can be actually purchased, and not for the "bundles" which are consumed as "aggregate choices".

4. The Basic Model

We now turn to the formal model and the results which generalize the observations of section 3 above. As mentioned above, this model does not explicitly deal with budget and prices, and may be considered a general model of repeated choice. It will be further specified to deal with consumer choices in the following section.

Let the set of actcomes be $A = \{1,2,\ldots,n\}$. The set of problems is $P = \mathrm{N} = \{1,2,\ldots\}$ and the similarity function is given by $s(i,j) = 1$. Let the (instantaneous) utility function for $i \in A$ be $u(i) = u_i$.

We are interested in the (infinite) sequence of choices (elements of A) which are (stagewise) U-maximizing, as a function of the utility vector $u = (u_1,\ldots,u_n)$. Since the interesting case will be that of negative utilities, it will prove convenient to define

$$u(i) = u_i = -a_i.$$

For $x \in A^\infty$ define the number of appearances of choice $i \in A$ in the sequence x up to stage $t \geq 0$ to be

$$F(x,i,t) = \#\{1 \leq j \leq t | x(j) = i\},$$

and let $U(x,i,t)$ denote the U-value of actcome $i \in A$ at that stage. In view of the constant similarity function, this boils down to

$$U(x,i,t) = F(x,i,t)u_i.$$

For given parameter values $u=(u_1,\ldots,u_n)$ let $S(u)$ denote the set of all sequences of choices which are stagewise U-maximizing. That is,

$$S(u)=\left\{x\in A^{\infty}\middle|\, for\ all\ t\geq 1,\ x(t)\in \arg\max_{i\in A} U(x,i,t-1)\right\}.$$

Of special interest will be the relative frequencies of the actcomes, denoted

$$f(x,i,t)=\frac{F(x,i,t)}{t}$$

and their limit

$$f(x,i)\equiv \lim_{t\to\infty} f(x,i,t).$$

(We will use this notation even if the limit is not guaranteed to exist. Thus " $f(x,i)$ exists" is a meaningful statement.) The omission of the actcome index will be understood to refer to the corresponding vectors:

$$f(x,t)=(f(x,1,t),\ldots,f(x,n,t))$$

$$f(x)=(f(x,1),\ldots,f(x,n)) .$$

Finally, denote

$$Y=\{u=(u_1,\ldots,u_n)|\,\forall i\ \ u_i\neq 0\}$$

and

$$V=\{u=(u_1,\ldots,u_n)|\,\forall i\ \ u_i< 0\}.$$

We can now formulate our first result:

<u>**Proposition 1:**</u> Assume that $u \in Y$. Then:

(i) for all $x \in S(u)$, $f(x)$ exists;

(ii) There exists $x \in S(u)$ for which $f(x)$ is one of the extreme points of the $(n-1)$-dimensional simplex iff this is the case for all $x \in S(u)$, and this holds iff $u_i > 0$ for some $i \in A$;

(iii) $f(x)$ is an interior point of the $(n-1)$-dimensional simplex iff $u_i < 0$ for all $i \in A$;

(iv) if $u_i < 0$ (i.e., $a_i > 0$) for all $i \in A$, then for *all* $x \in S(u)$, $f(x)$ is given by

$$f(x,i) = \frac{\prod_{j \neq i} a_j}{\sum_{k \in A} \prod_{j \neq k} a_j} \quad ;$$

(v) for every interior point y in the $(n-1)$-dimensional simplex there exist negative utility indices $u = (u_1, \ldots, u_n)$, unique up to a multiplicative scalar, such that $f(x) = y$ for all $x \in S(u)$.

(All proofs are relegated to an appendix.)

Remark: In the case that some utility values do vanish, this result does not hold. Consider, for instance, the extreme case where $u_i = 0 \;\; \forall i \in A$. Then we get $S(u) = A^{\infty}$ and, in particular, $f(x)$ need not exist for all $x \in S(u)$. Furthermore, one may get the entire $(n-1)$-dimensional simplex as the range of $f(x)$ when it is well-defined. More generally, convergence of the limit frequencies may fail whenever there are at least two indices $i \in A$ with $u_i = 0$, and the range of $f(x)$ (when well-defined) may include some of the sub-simplices of the $(n-1)$-dimensional simplex but not others.

A few words of interpretation may be in order. First, this result shows that the insights provided by the example of section 3 generalize to this set-up, with an arbitrary (but finite) set of products and arbitrary utility function. That is, low aspiration levels (positive utility values) are related to extreme solutions of the consumer's aggregate choice problem, whereas high aspiration levels (negative u_i's) give rise to interior solutions. Similarly, high and low aspiration levels may explain boredom aversion and habit formation, respectively.

Second, Proposition 1 suggests a new interpretation of the utility function. Assume that all utility indices are negative, and consider two products $i, j \in A$. The relative frequencies with which they will be consumed (for all $x \in S(u)$) are in inverse proportion to their utility levels:

$$\frac{f(x,i)}{f(x,j)} = \frac{a_j}{a_i} = \frac{u_j}{u_i}.$$

For example, if $u(1) = -1$ and $u(2) = -3$, product 1 will be consumed 3 times as frequently as product 2. If these are the only two products, their consumption frequencies will be $3/4$ and $1/4$, respectively. Thus the utility of a product has a different meaning than in the neo-classical theory. On the one hand, the fact that product 2 has a lower utility than product 1 does not imply it will never be chosen. It will, however, be chosen less often than its substitute. On the other hand, the utility indices do more than to merely rank the products; they also provide the frequency ratios, and are therefore cardinal.

The above result also shows that *any* point in the interior of the simplex may be the result of aggregate consumption by a case-based consumer for some utility function. While it is comforting to know that the theory is general enough in this sense, one may worry about its meaningfulness. After all, even under the assumption of negative utility values, it seems to postulate only that the consumer's choices will have limit frequencies. Indeed, not all the possible sequences have such a limit, but in any finite time horizon the theory may appear close to vacuous.

The following result shows that this is not the case. It turns out that in an appropriately defined sense, very "few" sequences are compatible with case-based consumer choices.

Proposition 2: Assume that $u \in Y$ (i.e., that for all $i \in A$, $u_i \neq 0$.) Then :

(i) if $u_i > 0$ for some $i \in A$ then $S(u)$ is finite;

(ii) if $u_i < 0$ (i.e., $a_i > 0$) for all $i \in A$, then

$$|S(u)| = \begin{cases} n! & \text{if } a_i/a_j \text{ is irrational } \forall i,j \in A \\ \aleph & \text{otherwise} \end{cases};$$

(iii) denote

$$S^- = \bigcup_{u \in V} S(u)$$

and

$$S^+ = \bigcup_{u \in Y \setminus V} S(u).$$

Let $p = (p_i)_{i \in A}$ be some probability vector on A, and let λ_p be the induced product measure on A^∞ (endowed with the product σ-field.) Then S^- is contained in a λ_p-null set. Furthermore, S^+ is finite, and if p is not degenerate, S^+ is a λ_p-null set.

Thus we find that there are uncountably many sequences of choices which are U-maximizing. Even in the case $n = 2$, there are uncountably many ratios ${}^{a_1}\!/_{a_2}$, namely the irrational ones, each of which contributes two sequences to S^- , and additionally countably many ratios (the rationals), each of which contributes uncountably many sequences. Yet part (iii) of the proposition states that, overall, the set $S \equiv S^- \cup S^+$ is "small" by any of the "reasonable" measures defined above. Furthermore, if one would try to "test" the hypothesis that the consumer's choices are in accordance with our theory, with the alternative being a sequence of i.i.d. random choices, our theory, if false, will be very easy to reject. It will also be clear from the proof of Proposition 1 that finitely many observed choices may often suffice to conclude that they cannot be a prefix of any sequence in S.

The propositions presented here may be viewed as an axiomatization of the utility function u. Explicitly, for any choice sequence $x \in A^\infty$ the following two statements are equivalent: (i) there is a vector $u \in V$ such that $x \in S(u)$; (ii) $f(x)$ exists, it is strictly positive, and $x \in S(u)$ for any $u \in V$ satisfying $u_i f_i = u_j f_j$ for all $i, j \in A$.

The reader may wonder why we refer to this equivalence as an "axiomatization." Admittedly, it is rather trivial, and, which is worse, very inelegant from a mathematical viewpoint. However, one of the main roles of axioms in a theory is, according to the logical positivist view, to relate theoretical concepts to observable ones. The utility function, here as well as throughout economic theory, is a theoretical construct; for it to be meaningful, one has to know how it is measured from observable data. Thus the equivalence above does serve this purpose of an "axiomatization": it provides necessary and sufficient conditions on observations, under which the theory holds, and in this case it also allows for the derivation of the theoretical constructs in an asymptotically unique way (up to a multiplicative scalar.)

We therefore conclude that the theory of case-based consumer behavior is non-vacuous. On the contrary, it is too easily refutable. However, the usual disclaimer applies here: as almost every theory in the social sciences, this one should not be taken too literally, and should only be considered some preliminary approximation of the modeled reality. Even those who accept case-based consumer theory wholeheartedly would have to concede that in more realistic set-ups the similarity function is not constant, some uncertainty is present, and the instantaneous utility function itself changes over time.

5. Prices and the Budget Constraint

5.1 Prices

The analysis presented so far implicitly assumed that the consumer only needs to choose among several products, and that at every period exactly one of them is to be consumed. The only thing our consumer seems to care about is the utility derived from each product, with no reference to its cost or even to the availability of funds to purchase it. The model can be easily re-interpreted to include a "null product" option, i.e., a "product" whose choice implies abstention from consumption in a given period, but even if our consumer runs out of money (s)he has no incentive to choose such an option. We are therefore faced with the question, how do the prices and the budget constraint affect the consumption decisions of economic agents?

The neo-classical theory suggests that a rational consumer chooses a bundle so as to maximize his/her utility function given his/her budget constraint. This assumption is quite reasonable in certain situations, for instance if the consumer is very rational and calculated "by nature", or if (s)he is forced to be such by the circumstances. For example, a person may be aware that (s)he is facing a very tight budget constraint (say, one is to become a graduate student in economics) and (s)he therefore consciously chooses an allocation of the budget among the needs which seem indispensable.

However, for many consumers the budget constraint may not be as stringent. (Consider, for instance, the same student after finding a job.) The (neo-)classical theory maintains that they would still behave "as if" they were consciously optimizing their utility function. We do not find this account very compelling, and would like to suggest a different view, which may be more insightful for the analysis of boundedly rational consumers (especially those who can afford to be only boundedly rational.)

According to this view, the consumer does not bother to think about the optimization problem as a whole. Rather, (s)he makes "daily" consumption decisions, each of which takes into account the market prices in some "local" manner, but without looking at "the complete picture". Consider, for instance, a consumer who passes by a store and notices a product, say, a shirt. Many consumers seem to be lucky enough to think in terms of "This is a good price for this shirt" rather than "I'll be better off if I buy this shirt and give up the concert tonight." In other words, the consumption decisions, at least in the short run, are often made by direct comparison of the product desirability to its price, rather than by re-optimizing the utility function given the observed price.

Following this intuition, we propose to include a product's price in its description, and to model the utility function as a function of the pair $(product, price)$. In the context of the model presented above, let $A = \{1,\ldots,n\}$ be the set of products, each of which has an "intrinsic" utility value given by $v(i) = v_i$. These values designate the "utility" that would correspond to the choice among the products, were they free. However, when confronted with a choice problem (of which we will still think as a repeated one), the consumer also observes the prices $p = (p_1,\ldots,p_n)$, and the utility function which will be determining his/her choices is

$$u(i, p_i) = d(v_i, p_i)$$

where $d(\cdot,\cdot)$ is assumed monotonically increasing in its first argument and monotonically decreasing in the second.

We would like to interpret this utility function more or less literally: our consumer is assumed to feel quite happy for getting a certain product at a "bargain price", and suffers disutility from the mere act of paying. By this we do not necessarily mean that the consumer's self-esteem depends on his/her performance on the market (even though one may choose this line of thought as well.) Rather, we imagine the consumer as being rational enough to know that it is to his/her advantage to pay less, even without a clear idea or plan for the usage of the rest of the money.[2]

For simplicity let us consider a linear function

$$d(v_i, p_i) = v_i - \alpha p_i \qquad (*)$$

for some $\alpha > 0$. The coefficient α should be interpreted as a "value of money" parameter, which indirectly reflects the budget constraint: the richer is the individual, the smaller is α and the smaller will be the impact of changes in prices on his/her consumption decisions.

Let us consider the following example. Let $n = 2$; $\alpha = 1$;

$$v_1 = v_2 = -1;$$
$$p_1 = 1 \; ; \; p_2 = 2.$$

The consumer is choosing a sequence out of $S((-2,-3))$, and thus will be consuming product 1 60% of the time, and product 2 -- 40%. Now assume that product 2's price is being raised to $p_2 = 3$. If the consumer were to start choosing again, with an empty memory, (s)he would end up choosing a sequence in $S((-2,-4))$, that is, choosing product 1 about 67% of the time, and product 2 -- about 33%.

Comparative statistics in this model thus reaffirms the basic intuition regarding substitute products: overall (i.e., in the aggregation) the consumer will use less of a product when its price goes up (other things being the same.) However, our model allows us to do comparative dynamics as well. That is, we may ask what happens if we change the prices along the sequence of consumption choices. For instance, suppose that 10 days have past since the year started, and only then does the price of product 2 go up. At that point, our consumer has been consuming product 1 precisely 6 times, and product 2 -- 4

---

2 The introduction of prices into the utility function directly has been discussed in the literature in other contexts. We will be grateful for exact references.

times. This implies that the U -values of the two are equal. For all intents and purposes, it is as if the change has occurred with no memory.

If, however, the change in prices were to occur after, say, day 9, the first 9 consumption periods would have some residual effect on the U -values of the products. Yet the consumer's response would still be immediate -- in the sense of adding -4 to $U(2)$ whenever product 2 is consumed -- and the long-run consumption frequencies would still converge to (67%,33%).

Naturally, the limit frequencies of the consumption decisions are not guaranteed to exist if prices keep changing indefinitely. (Though the model will provide rather specific predictions in this case as well.) Should prices be fixed from a certain period on, the limit frequencies would exist, and would only depend on the latest prices.

To what extent are the parameters of (*) observable? That is, can observed consumer behavior determine the values of $v=(v_1,\ldots,v_n)$ and α? First, note that the observed relative frequencies are

$$\frac{f(x,i)}{f(x,j)}=\frac{v_j-\alpha p_j}{v_i-\alpha p_i}\ .$$

Hence one cannot hope to observe more than the ratios ${}^{v_i}\!/\!_{\alpha}$. Differently put, we may assume without loss of generality that $\alpha=1$ (by setting $v_i'={}^{v_i}\!/\!_{\alpha}$). However, under this assumption one can uniquely determine the vector v by observing the consumer behavior for two price vectors which induce different choice frequencies. To be more precise, let us repeat the definition of the "intrinsic" utility space and define the price space to be, respectively,

$$V=\left\{v=(v_1,\ldots,v_n)\middle|\forall i,\ v_i<0\right\}$$

$$P=\left\{p=(p_1,\ldots,p_n)\middle|\forall i,\ p_i\geq 0\right\}.$$

Let Δ^{n-1} denote the interior of the $(n-1)$-dimensional simplex, i.e.,

$$\Delta^{n-1}=\left\{z=(z_1,\ldots,z_n)\middle|\forall i,\ z_i>0,\ \ \textstyle\sum_{i=1}^n z_i=1\right\}$$

and define the *frequency function* $f: V \times P \to \Delta^{n-1}$ by

$$\frac{f(v,p)_i}{f(v,p)_j} = \frac{v_j - p_j}{v_i - p_i}.$$

Then we have

Proposition 3: Let there be given $v \in V$ and $p^1, p^2 \in P$, for which $f(v,p^1) \neq f(v,p^2)$. Then there does not exist $v' \in V$, $v' \neq v$, such that $f(v',p^k) = f(v,p^k)$ for $k = 1,2$.

It will be clear from the proof of this result that it can be extended to values of $v = (v_1,\ldots,v_n)$ which are not necessarily negative, as long as the vectors $v - p^k$ are negative for $k = 1,2$. More generally, postulating a particular functional form of $d(v_i,p_i)$ allows one to estimate its parameters by observing choice frequencies for various price vectors.

5.2 The Budget Constraint

Boundedly rational as our consumers may be, they may still find themselves with no financial resources left. When will this happen? How much money does a case-based consumer spend?

Assume $A = \{1,\ldots,n\}$ with parameters $v = (v_1,\ldots,v_n) \in V$, $\alpha \geq 0$, and $p = (p_1,\ldots,p_n) \in P$. (We have convinced ourselves that one may assume that $\alpha = 1$. However, it may be conceptually more convenient to think of fixed "intrinsic" value function v and adjustable "value of money" parameter α, which will eventually be used to balance the consumer's budget.)

Such a consumer is choosing the various products according to the frequency vector $f = f(v,p,\alpha) \in \Delta^{n-1}$ defined by

$$\frac{f_i}{f_j} = \frac{v_j - \alpha p_j}{v_i - \alpha p_i} \qquad \forall i,j \in A.$$

Thus the average "daily" expenditure of our consumer is

$$E = E(v,p,\alpha) \equiv f \cdot p = \sum_{i \in A} f_i p_i \ .$$

This amount should be compared to the consumer's budget constraint. Let us assume, as in classical consumer theory, that the latter is given exogenously at level $I > 0$. However, we will interpret it as an average "daily" income (say, as $\frac{1}{30}$ of a monthly salary.) Thus, if $I < E$, our consumer will face financial difficulties, will either have to borrow money, or to revise his/her consumption habits. On the other hand, if $I > E$, the consumer will have unused funds, which can either be saved or used to afford a higher frequency of more expensive products. Of course, the consumer runs a "balanced budget" if equality holds.

Since f is in the *interior* of the simplex, it is obvious that, unless all prices are equal,

$$\min_{i \in A} p_i < E < \max_{i \in A} p_i.$$

Thus the consumer can hope to balance the budget only if $I > \min_{i \in A} p_i$. However, we will assume that $\min_{i \in A} p_i = 0$. In many applications this condition holds with no loss of generality, since, as mentioned above, one may include in the model the "null product", which designates abstention from consumption and has a zero price. Under this assumption, it turns out that the obviously necessary condition for the feasibility of the consumer's problem is also sufficient for the existence of a value-of-money parameter which will balance the budget.

Proposition 4: Let there be given a value function $v \in V$, a price vector $p \in P$ with $\min_{i \in A} p_i = 0$ and an income level $I > 0$. Then:

(i) there exists $\alpha > 0$ such that $I \geq E(v, p, \alpha)$;

(ii) if $I < E(v, p, 0)$, then there exists $\alpha > 0$ such that $I = E(v, p, \alpha)$.

One way to interpret this result would be the following: assume, as mentioned above, that the consumer is characterized by the "utility" $v \in V$, and faces prices $p \in P$. At a given point of time, (s)he has a value-of-money parameter α, and (s)he spends about $E(v, p, \alpha)$ money units per day (on the average.) After a certain number of periods (days), the consumer finds out how much (s)he has spent in relation to the budget constraint. Realizing one has spent too much, one would start putting more weight on money in one's decisions. Every consumption decision is going to have a lower "utility" value, and in the model presented above this is reflected by increasing the parameter α. Conversely, the consumer may find out that (s)he has excess supply of money. The consumer may then either adjust the income (say, by saving more), or decide to consume it all. In the latter case, every price of every product will seem less important, and this will be reflected (at least from the modeler's point of view) by a lower value of α. Thus, if prices do not change over time, one could expect that in the long run α would be adjusted for a balanced budget.

The result above suggests that, as long as the consumer has the option of abstaining from consumption, a high enough level of α will guarantee that (s)he does not spend more than his/her income. Sometimes the consumer may not need a large α to obtain this goal. For instance, it may be that disregarding the cost altogether (setting $\alpha = 0$) will already result in expenditure level which does not exceed income. (This may happen either because the income is very high as compared to the prevailing prices, or simply because the consumer tends to "naturally" prefer some of the less costly products.) In this case the consumer's income exceeds his/her needs. If, however, this is not the case, that is, $I < E(v, p, 0)$, then there will be a level of α which precisely balances the budget.

It is interesting to note that even the very simple linear function $d(v_i, p_i)$ we employ in this model gives rise to a non-trivial behavior of the expenditure function. Consider the following example:

$$n = 3;$$
$$v_1 = -1;\ v_2 = -10;\ v_3 = -100;$$
$$p_1 = 1;\ p_2 = 2;\ p_3 = 0.$$

One may verify that for $\alpha = 0$

$$f(v, p, 0) = \frac{1}{1110}(1000, 100, 10) \approx (.9, .09, .009)$$

and

$$E(v, p, 0) = \frac{1200}{1110} \approx 1.08.$$

However, for $\alpha = 10$ we get

$$f(v,p,10) = \frac{1}{4430}(3000,1100,330) \approx (.68,.25,.07)$$

and

$$E(v,p,10) = \frac{5200}{4430} \approx 1.17.$$

Thus we find that the expenditure function is not necessarily monotonically decreasing in its last argument. When α increases, the weight of the prices -- as opposed to the value function -- increases. Indeed, if there is a free product, the expenditure function will have to tend to zero eventually. But on its way there it may also increase. In this example, product 2, which is the most expensive one, is consumed at a higher frequency for $\alpha = 10$ than for $\alpha = 0$. If we take a higher value for α as a proxy for a lower income level, we find here that the effect of a reduction in income on the consumption of the most expensive product may be negative.

However, one feature of the example above is a little peculiar: product 2 has a more negative "intrinsic utility" value than product 1; indeed, if prices are ignored, our consumer will use product 1 10 times more often than product 2. Yet product 2 is also more expensive. On the face of it, and without delving into the supply side and the price determination mechanism, it would appear that our consumer has a somewhat idiosyncratic taste: Given a sequence of choices between, say, free dinners in two restaurants, (s)he will choose the less expensive one more often. If these tastes were shared by all consumers, one would expect the market prices to reflect that. Consequently there is some reason to believe that our consumer is not "typical" in this respect.

Let us, therefore, define a consumer's taste -- i.e., $v = (v_1,\ldots,v_n) \in V$ -- to be *typical with respect to prices* $p = (p_1,\ldots,p_n) \in P$ if

$$v_i < v_j \qquad \text{iff} \qquad p_i < p_j \qquad \forall i,j \in A.$$

Recall that the v_i's are negative. The smaller is the value -- the higher is it in absolute value -- the less frequently will it be chosen. Hence the lower values correspond to the less desirable products, and a typical consumer taste is thus required to find more expensive products also more desirable.

We can now formulate the following result.

Proposition 5: Let there be given a value function $v \in V$ and a price vector $p \in P$, such that v is typical with respect to p. Then

(i) $E(v,p,\alpha)$ is non-increasing in α ;

(ii) if $\max_{i \in A} p_i > \min_{i \in A} p_i$ (i.e., not all prices are equal), then $E(v,p,\alpha)$ is decreasing in α; otherwise, (namely, $p_i = \hat{p}\ \forall i \in A$), $E(v,p,\alpha)$ is constant in α (and equals $\hat{p}$) ;

(iii) if $0 < I < E(v,p,0)$, and $\max_{i \in A} p_i > \min_{i \in A} p_i = 0$, then there exists a unique $\alpha > 0$ such that $I = E(v,p,\alpha)$.

Remarks:

- Note that the statement in Proposition 5 (iii) is stronger than that of Proposition 4 (ii) in that the former guarantees the uniqueness of α. This conclusion follows from two additional assumptions: the innocuous assumption that not all prices are equal, and the major one, namely that the consumer's taste is typical with respect to the prices.
- The above results guarantee, under appropriate conditions, that there will be a parameter value α for which the consumer's budget is balanced. Moreover, if $E(v,p,\alpha)$ is monotonically decreasing in α, the (unique) "balancing" value of α can be found by trial-and-error. However, these results do not shed much light on the process by which a consumer adjusts his/her α to match the budget constraint.

As mentioned above, the model suggested here implicitly views the consumer as choosing products on a "daily" basis, while stopping to think and to compare his/her expenditures to the given budget every so often, say, "monthly". This comparison may result in a new value for α. The adjustment of this parameter need not be conscious, of course -- it suffices that our consumer will be appalled by his/her lavishness, or be traumatized by having exhausted the budget well before the end of the month, to start putting more weight on prices when making consumption choices. Similarly, a consumer who discovers (s)he has been needlessly frugal will "naturally" focus on the products' "value" more than their cost, which is equivalent to lowering α.

Yet this process is not explicit in the model we suggest here. There are many ways in which it can be formalized, but the study of specific adjustment mechanisms is beyond the scope of this paper.

- The analysis carried out above assumes a utility function which is linear in the "intrinsic value" and the price. Simple and intuitive as it is, it certainly is restrictive in some ways. Consider, for example, the asymptotic frequency vector as α tends to infinity. Assuming the possibility of abstention, i.e., that there is a zero-cost "null product", we found that the "expensive" products are consumed at lower frequencies, which tend to zero. While this assumption may be reasonable enough for the choice of annual vacation, it may not be convincing for the choice of a daily meal. In such a case the model may be more realistic without the "null product", predicting a financial crisis rather than a consumer's compromise. However, the linear utility function will predict such a crisis even when it is not entirely necessary: in case $p_i > 0$ for all $i \in A$, the relative frequencies still satisfy

$$\frac{f(v,p,\alpha)_i}{f(v,p,\alpha)_j} \xrightarrow[\alpha\to\infty]{} \frac{p_j}{p_i}$$

which implies that the consumption frequencies of the more expensive products will remain bounded away from zero. This may result in violation of the budget constraint while $I > \min_{i\in A} p_i$. That is to say, such a consumer may run out of money even though (s)he could make do with less expensive products.

One may postulate other utility functions which will predict more modest consumer choice. For instance, if $p_i > 1$ for all $i \in A$, the function

$$d(v_i, p_i) = v_i - p_i^{\alpha}$$

would guarantee that

$$\frac{f(v,p,\alpha)_i}{f(v,p,\alpha)_j} \xrightarrow[\alpha\to\infty]{} 0 \qquad\qquad \text{whenever } p_i > p_j.$$

While we find that the linear function is quite reasonable at least for those cases in which a "null product" exists, it should be perceived as an example more than a proposed theory.

6. Simultaneous Problems

As mentioned above, it is a little artificial to suggest that a consumer has a given budget for one repeated choice such as annual vacation or daily meal. It seems that, as modelers, we face the following choice: either we focus on a choice which is repeated, but has no exogenously given budget, or else we assume a given income but give up the supposition that the same choice is repeated over and over again.

However, it is a straightforward generalization of our model to assume that several repeated consumption problems are solved simultaneously. That is to say, that the consumer still faces a sequence of decision problems, each of which is attached to -- or perhaps even defined by -- the set of possible choices, which is some subset of the set of actcomes A. Thus the decision problems may include as diverse problems as the choice of meal, of outfit, of means of transportation and so forth. In each problem, only the available choices are compared (in terms of their U-values), but they are all taken together in the total expenditure function. This function will naturally depend not only on utilities and prices, but also on the frequency with which various subsets of A appear as the choice set.

Let us consider a simple example. Assume that $A=\{1,2,3,4\}$ where products 1 and 2 are meal choices, and products 3 and 4 -- transportation options. A meal problem is defined by the choice set $A_1=\{1,2\}$ while a transportation problem -- by $A_2=\{3,4\}$. Let us assume that the order by which the problems are encountered by the consumer is exogenous and deterministic. For simplicity, let the odd numbers designate meal problems, and the even ones -- transportation problems.

Next assume that the consumer's taste and the market prices are given as follows:

$$v_1 = v_2 = v_3 = v_4 = -1$$

and

$$p_1 = 1 \quad p_2 = 4 \quad p_3 = 1 \quad p_4 = 2.$$

Let us first consider the case in which the consumer's "daily" income is $I=2$; that is to say, the average expenditure over all types of problems cannot exceed this income. It is easy to verify that $\alpha = 0$ balances the budget. For this value half of the meal choices will be products 1, while the other half will be product 2, and similarly products 3 and 4 will be consumed with equal frequencies. Taking into account the frequencies of the two types of problems, we conclude that the overall limit frequency vector is

$$f = \left(\tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}\right).$$

The average expenditure on meals (per choice) is 1.25, and on transportation -- .75.

Next assume that the consumer's income has dropped to

$$I = \tfrac{25}{16} = 1.5625.$$

This value, of course, was not chosen at random; rather, it is the expenditure corresponding to $\alpha = 2$. Indeed, for this value one gets

$$v_1 - \alpha p_1 = -3 \qquad v_2 - \alpha p_2 = -9$$

$$v_3 - \alpha p_3 = -3 \qquad v_4 - \alpha p_4 = -5.$$

Hence out of the meal decisions, $\tfrac{1}{4}$ of the choices will be for the (more expensive) product 2, and $\tfrac{3}{4}$ of the time product 1 will be chosen. Similarly, out of the transportation decisions, $\tfrac{5}{8}$ of the time product 3 is chosen, and $\tfrac{3}{8}$ is the share of product 4. Overall we get the frequency vector

$$f = \left(\tfrac{6}{16}, \tfrac{2}{16}, \tfrac{5}{16}, \tfrac{3}{16}\right).$$

Here the average expenditure on food per day is .875, and on transportation -- .6875, which add up to 1.5625.

A few conclusions and comments are in order:

- As can be seen in this very simple example, changes in income may endogenously change the amounts spent on various "commodities" such as "food" and "transportation". We need not assume that the consumer decides at any point how much to spend on each "category" of products (or on a "type" of problems), nor that (s)he consciously revises this decision in face of changing circumstances. The process by which the value-of-money parameter is determined also implies how is the budget to be divided.

- It is straightforward to generalize the model and results of section 5 to the case of "simultaneous" problems whose order is exogenously and deterministically given. As long as limit relative frequencies are concerned, one may extend the results to the case in which the decision problems are stochastically selected as well. However, sometimes the choice of problems is endogenous. A consumer who decides "to go shopping", for example, chooses to be exposed to decision problems which would otherwise not arise. One possible way to model these choices is to apply our model in a hierarchical manner: for instance, the decision between "go shopping" and "stay at home" is done as if the two were terminal products, and if shopping is selected, there is a sub-problem of choice among stores, within each of which there is a sub-sub-problem of choice among products and so forth.
- In the highly simplified example above, the products are partitioned into choice sets, which correspond to consumption problems. Yet there are cases in which a certain product may "solve" more than one of the problems. For instance, "dining out with friends" may solve both a food and an entertainment problem. The current version of our model does not deal with these cases since it is a model of non-durable goods. A consumption in a given day does not last to the next -- apart from what is left of it in the consumer's memory.
- One may not be satisfi(c)ed with our model, which implicitly assumes that the set of possible acts uniquely defines the problem. For instance, one may face the same possible acts for the problems "How do I get to work?" and "How do I get to the theater?", while these problems may be quite distinct, if only for the time of day at which they pose themselves. Indeed, we may insist that the problems are still defined by their possible solutions (i.e., actcomes), and that "subway at 8am" and "subway at 11pm" are different actcomes. But it seems more fruitful to relax the assumption that the similarity function is constant, and let it reflect the aspects in which various problems, whose choice sets are not disjoint, still differ.

7. The Potential

In sub-section 2.2 above we suggested the interpretation according to which the very experience of consumption of product $i \in A$ is history-dependent. Thus the utility is $U(i)$ and it changes by $u(i)$ whenever i is chosen. In this section we would like to advance this interpretation a little further, in a way that will hopefully facilitate the comparison of our theory to the neo-classical one.

We adopt all notations and definitions of section 4. We retain the assumption that $s(\cdot,\cdot)=1$. While the potential may be defined under more general conditions, this case is of particular interest as will be clear below. We note that, as the discussion in section 4, this section also deals with the general repeated choice model, whereas the budget and prices are implicit in the utility function.

For a consumption vector $x \in A^{\infty}$, let $x_t \in A^t$ be the t-prefix of x. Let $\cdot$ denote vector concatenation. Denote the set of all prefixes (all finite vectors) by $A^{*} = \bigcup_{t \geq 0} A^t$. For a function $Y: A^{*} \rightarrow \Re$, $x_t \in A^t$ and $i \in A$, define

$$\frac{\partial Y}{\partial i}(x_t) = Y(x_t \cdot (i)) - Y(x_t).$$

That is, $\frac{\partial Y}{\partial i}$ is the change in the value of Y that will result from adding i to the consumption vector x_t. Denote by $U_i: A^{*} \rightarrow \Re$ the U-value of product i. I.e., $U_i(x_t) = U(x,i,t)$. Then we have already noted that

$$\frac{\partial U_i}{\partial i}(x_t) = u(i) = u_i$$

for all $x_t \in A^{*}$. Similarly, for $j \neq i$,

$$\frac{\partial U_i}{\partial j}(x_t) = 0.$$

Following the suggestion that the utility derived from consumption is history-dependent, one may ask how can we measure the "well-being" of a consumer at some point of time. One obvious suggestion is to consider the total sum of all past experiences. Define $W: A^{*} \to \Re$ by

$$W(x_t) = \sum_{\tau=1}^{t} U_{x(\tau)}(x_{\tau-1})$$

for $x_t \in A^t$. Thus, a U-maximizing consumer may be described as maximizing his/her W function at any given point of time t. Furthermore, W is uniquely defined (up to a shift by an additive constant) by

$$\frac{\partial W}{\partial i}(x_t) = U_i(x_t)$$

for all $x_t \in A^{*}$ and $i \in A$. Hence W is a single function, such that the utility of product i, U_i, is its derivative w.r.t. i . It thus deserves the title "the potential of the utility".

In certain ways, the potential function is closer to the neo-classical utility function than either U or u. We have discussed above the distinction between u and the neo-classical utility function. However, the latter plays a rather different role from that of U as well: U is history-dependent, but is still defined for a single product; the neo-classical function, by contrast, is defined for bundles. Correspondingly, if past experiences are assumed to linger in the consumer's memory, U does not attempt to capture the "overall well-being" of the consumer, while the neo-classical utility does. On the other hand, the potential function W is also defined for "bundles" (implicit in the vector x_t), and may be interpreted as an overall measure of well-being.

Furthermore, since

$$\frac{\partial W}{\partial i} = U_i \qquad \text{and} \qquad \frac{\partial U_i}{\partial j} = \begin{cases} u_i & i = j \\ 0 & i \neq j \end{cases}$$

we get

$$\frac{\partial^2 W}{\partial j \partial i} = \begin{cases} u_i & i = j \\ 0 & i \neq j \end{cases} .$$

Hence W is concave if $u_i < 0$ (for all $i \in A$) and convex if $u_i > 0$. Indeed, we have found earlier (see Proposition 1) that negative u values induce consumer choices which are also predicted by convex preferences, that is by a concave neo-classical utility function, while positive u values correspond to a convex neo-classical utility function. Thus the potential parallels the neo-classical utility function also in terms of the relationship between concavity/convexity and interior/corner solutions. As we will see in the next section, this analogy is further accentuated when one studies substitution and complementarity between products.

Technically, the potential is a rather different creature from the neo-classical utility function. While the former is defined on the space A^* (which is topologically discrete,) the latter is defined on $\Re^n$. Conceptually, the potential is defined on consumption streams, while the neo-classical utility -- on bundles. However, under the assumption of constant similarity function (between problems), a "stream" may be identified by the relative frequencies of the products in it. (That is, all permutations of a given sequence are equivalent in terms of the behavior they induce.) Thus one may bring the potential closer to the neo-classical utility by "normalizing" it.

Formally, for $x \in A^\infty$ and $t \geq 0$, recall that $F(x,i,t)$ denotes the number of appearances of i in x up to time t. Let $T(x,i,k)$ stand for the time at which the k-th appearance of i in x occurs, i.e.,

$$T(x,i,k) = \min\{t \geq 0 | F(x,i,t) \geq k\}.$$

(This function will be taken to equal ∞ if the set on the right is empty. However, for the case $u_i < 0$ it will be finite.) Then we obtain

$$W(x_t) = \sum_{\tau=1}^{t} U(x, x(\tau), \tau - 1) =$$

$$\sum_{i=1}^{n} \sum_{k=1}^{F(x,i,t)} U(x,i,T(x,i,k) - 1) =$$

$$\sum_{i=1}^{n} \sum_{k=1}^{F(x,i,t)} (k-1)u_i =$$

$$\frac{1}{2} \sum_{i=1}^{n} u_i F(x,i,t)[F(x,i,t) - 1] \ .$$

Recall that $f(x,i,t)$ is the relative frequency of i in x up to time t. Thus,

$$\frac{W(x_t)}{t^2} = \frac{1}{2}\sum_{i=1}^{n} u_i f(x,i,t)\left[f(x,i,t) - \frac{1}{t}\right].$$

For a point $f = (f_1,\ldots,f_n)$ in the $(n-1)$-dimensional simplex, define the "normalized potential" to be

$$w(f) = \frac{1}{2}\sum_{i=1}^{n} u_i f_i^2.$$

Then, for large enough t,

$$\frac{W(x_t)}{t^2} \approx w(f(x,t)).$$

We can now redescribe the choices of our consumer: at any given point of time $t \geq 0$, (s)he has a value for the normalized potential, and behaves as if (s)he were trying to (approximately) maximize it. To be precise, the consumer is choosing a product so as to maximize $W(x_{t+1})$, or, equivalently, to maximize $W(x_{t+1})/(t+1)^2$. However, in the long run this is approximately equivalent to maximization of the normalized potential w. Considering the optimization problem

$$\begin{aligned} &MAX\ w(f) \\ &s.t. \quad \sum_{i=1}^{n} f_i = 1 \\ &\qquad\ \ f_i \geq 0, \end{aligned}$$

it is straightforward to check that, should it have an interior solution (relative to the simplex), the solution must satisfy

$$f_i u_i = const.$$

Indeed, Proposition 1 shows that, if $u_i < 0$ (for all $i \in A$), there is an interior solution satisfying the above condition. Furthermore, it shows that the "greedy", or "hill climbing" algorithm implemented by our consumer converges to this solution.

At this point the similarity between our (normalized) potential and the neo-classical utility may suggest that they are simply identical. It is worthwhile to stress the following distinctions:

1. The neo-classical utility function is assumed to be globally maximized by a one-shot decision. The consumer, knowing his/her utility or completely ignorant of it, is supposed to miraculously select an optimal point in the bundle space. By contrast, the normalized potential is locally improved upon. In case of a concave potential, we have shown that this local optimization, or "hill-climbing", "follow-the-gradient" algorithm does happen to converge to a global optimum. However, in other cases the algorithm may end up at a corner point which is not globally optimal. In particular, whenever several products have identical U values, the consumer's choice is completely arbitrary and lacks any "look ahead" property. Thus a consumer who may be satisficed need not optimize.

2. While the neo-classical utility function is a primitive of the classical model (or, at best, derived from preferences over the bundle simplex), the potential is a mathematical artifact. Our model is phrased in terms of the instantaneous utility u and the CBDT functional U, which are supposed to be cognitively significant in their own right. The potential, while potentially a useful theoretical tool, has no claim to intuitive appeal. Indeed, like the neo-classical utility, it does not seem to shed much light on the consumer's decision making process. It is our hope that the model presented above provides more insight beyond mere maximization of the potential.

3. This point is also reflected in the way the neo-classical utility and the potential are used in the consumer's choice problem: the neo-classical utility is supposed to be fixed, and defined independently of prices and income. By contrast, the potential would vary with both prices and the "value-of-money" parameter. Thus, for fixed prices and income, the potential may be used as a local objective function, whose maximization describes consumer's choices. Yet, when either of these parameters change, the "objective function" would change accordingly. In other words, the potential does not presuppose a dichotomy between preferences and the feasible set.

4. The calculations above show that the potential is a quadratic function on the bundle space. (In fact, it is also additively separable; however, this

assumption will be relaxed in the following section.) It thus appears that the neo-classical utility function is much more general, allowing for a variety of functional forms, or, equivalently, for a non-constant second derivatives. On the other hand, the theory presented here calls for generalization in a different direction, namely for the introduction of non-constant similarity function. Once we relax the assumption that $s(t_1,t_2)=1$ for any two problems $t_1,t_2 \in \mathrm{N}$, we can still use the potential W but not the normalized potential w. In other words, when the similarity between problems varies, the bundle simplex is not rich enough to describe the consumer's memory. The relative frequencies with which the products were consumed will not suffice as the *order* of their appearance will also be of importance.

Yet one may consider the potential as a way to derive (a special case of) the neo-classical utility function. Thus viewed, we may argue that the classical assumption of convex preferences (or concave neo-classical utility) is justified by (and equivalent to) the assumption that consumers have high aspiration levels.

8. Substitution and Complementarity

As mentioned in section 2, one may generalize the case-based decision model by postulating an act similarity function, such that each act's evaluation is affected by any other act's past performance to the extent that the two acts are similar.

In the context of decision under certainty, there may be some relevant notion of similarity over actcomes. Consider the example of section 3 again. One may judge *chicken* more similar to *beef* than to *fish*. Since every time *chicken* is consumed, $s_A(\mathit{fish},\mathit{chicken})u(\mathit{chicken})$ is added to $U(\mathit{fish})$, and $s_A(\mathit{beef},\mathit{chicken}) = s_A(\mathit{chicken},\mathit{beef}) = 1$ -- to $U(\mathit{beef})$, we may find, for negative utility values, that $U(\mathit{beef})$ has decreased by more than $U(\mathit{fish})$ has. For instance, consider the extreme case in which

$$s_A(\mathit{beef},\mathit{chicken}) = s_A(\mathit{chicken},\mathit{beef}) = 1$$

where the other s_A values are zero. For these similarity values, when the choice set is $\{\mathit{chicken},\mathit{fish}\}$ or $\{\mathit{beef},\mathit{fish}\}$, the relative frequencies of consumption are

$\left(\frac{1}{2},\frac{1}{2}\right)$. However, for the set $\{beef, chicken, fish\}$ they may be $\left(\frac{1}{4},\frac{1}{4},\frac{1}{2}\right)$ rather than $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$. (In fact, in this example, *beef* and *chicken* are "identical", and for any $p \in \left[0,\frac{1}{2}\right]$ the vector $\left(p,\frac{1}{2}-p,\frac{1}{2}\right)$ may be the limit frequencies of consumption. Furthermore, there will also be consumption sequences for which limit frequencies do not exist.)

One interpretation of the actcome-similarity function is that it measures substitutability: the closer are two products to be perfect substitutes, the higher is the similarity between them, and the greater will be the impact of consuming one on the boredom with another. If the similarity between two products is zero, they are "substitutes" in the sense that both can still be chosen in the same problem, but the consumption of one of them does not affect the desirability of another.

With a slightly different interpretation of the model, the actcome similarity function may also capture complementarities: suppose that the recurring choice in our model is a *purchase* decision, rather than a consumption decision. Thus at every day, or instant, the consumer chooses a certain product, but the consumer also has a supermarket cart, or kitchen shelves, which allow him/her to consume several products together. Having bought tomatoes yesterday and cucumbers today, the consumer may have a salad.

Following this interpretation, it is only natural to extend the actcome-similarity function to negative values in order to model complementarities. If the actcome-similarity between, say, *tea* and *sugar* is negative, (and so are their utilities), having just purchased the former would make one more likely to purchase the latter next.

More explicitly, assume that there are three products, $A = \{tea, coffee, sugar\}$ where $u(tea) = u(coffee) = -1$ and $u(sugar) = -2$. In the absence of actcome similarities, the relative purchase frequencies will converge to $\left(\frac{2}{5},\frac{2}{5},\frac{1}{5}\right)$. However, let us assume that

$$s_A(sugar, tea) = s_A(sugar, coffee) = -1$$

and

$$s_A(tea, sugar) = s_A(coffee, sugar) = -0.5 \ .$$

(It will later be clear why such a similarity function may be considered "symmetric".) Assuming there is zero similarity between *tea* and *coffee*, we note that after having purchased the sequence $(sugar, tea, coffee)$, the U-value of all alternatives is zero. Thus the relative frequencies will converge to $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$.

As mentioned in section 7 above, the instantaneous utility function u may be interpreted as the time-derivative of U: if an actcome i is chosen, $U(i)$ is changed by the amount $u(i)$. That is,

$$\frac{\partial U_i}{\partial i} = u_i \ .$$

Extending this interpretation to the actcome-similarity set-up, we find that, if an actcome i is chosen, $U(j)$ is changed by $s_A(j,i)u(i)$. In other words,

$$\frac{\partial U_j}{\partial i} = s_A(j,i)u_i \ .$$

Combining the above, we conclude that the similarity function may be interpreted as the ratio of derivatives of the U functions:

$$s_A(j,i) = \frac{\partial U_j / \partial i}{\partial U_i / \partial i} \ .$$

Using the potential function, we may write

$$\frac{\partial U_j}{\partial i} = \frac{\partial^2 W}{\partial j \partial i} = s_A(j,i)u_i = \frac{\partial U_j / \partial i}{\partial U_i / \partial i} \frac{\partial U_i}{\partial i}$$

or

$$s_A(j,i) = \frac{\partial^2 W / \partial j \partial i}{\partial^2 W / \partial i^2} \ .$$

The neo-classical theory defines the substitution index between two products to be the cross derivative of the utility function (w.r.t. the products quantities.) By comparison, the actcome similarity function, which measures substitution in our model, is the cross "derivative" of the potential function, normalized by the second derivative of the same function (W) w.r.t. one of the two. Furthermore, if we define the "substitutability" between two products i and j to be the impact of consuming one on the desirability (U-value) of the other, namely $s_A(j,i)u_i$, it precisely coincides with the cross derivative of the potential. In this respect we find again that the potential function in our model bears some kinship to the neo-classical utility.

It is worthy of note that in the presence of actcome similarity, the consumer (U-maximizing) choice given a sequence of past choices x_t will not depend on the order of the products in x_t, though the potential generally will. The U-value of product i is

$$U(x,i,t) = \sum_{\tau=1}^{t} s_A(i,x(\tau))u_{x(\tau)} =$$

$$\sum_{j=1}^{n} F(x,j,t)s_A(i,j)u_j$$

which obviously depends only on the number of appearances of each product in x_t. However, the value of the potential is

$$W(x_t) = \sum_{\tau=1}^{t} U(x,x(\tau),\tau-1) =$$

$$\sum_{\tau=1}^{t}\sum_{v=1}^{\tau-1} s_A(x(\tau),x(v))u_{x(v)}.$$

It is easy to see that the potential will be invariant w.r.t. permutations of x_t iff

$$s_A(j,i)u_i = s_A(i,j)u_j \qquad \forall i,j \in A$$

that is, iff

$$\frac{\partial^2 W}{\partial j \partial i} = \frac{\partial^2 W}{\partial i \partial j} \qquad \forall i,j \in A.$$

Note that this is the "appropriate" notion of symmetry in this model: the actcome similarity function itself may be symmetric without guaranteeing that the "impact" of consumption of i on desirability of j equals that of j on i. Rather, symmetry is naturally defined by $s_A(j,i)u_i = s_A(i,j)u_j$, or the equality of the cross-derivatives of the potential.

There may be some interest in models in which substitution effects are not symmetric. For instance, a consumer may be more likely to purchase *sugar* after having bought *tea*, but not vice versa. While the neo-classical model is symmetric by nature (for continuously differentiable utility function), in our model symmetry is an assumption one may choose to impose.

At any rate, only under the symmetry assumption, that is, only when the potential is invariant w.r.t. permutations can we meaningfully define the normalized potential on the bundle simplex as in section 7 above. It therefore turns out that one can use the normalized potential (or a "neo-classical utility") if and only if the potential has symmetric cross derivatives. In this case we may write

$$W(x_t) = \frac{1}{2}\sum_{i=1}^{n} u_i F(x,i,t)[F(x,i,t)-1] +$$

$$\sum_{i=2}^{n}\sum_{j=1}^{i-1} s_A(i,j)u_j F(x,i,t)F(x,j,t).$$

Under the assumption $s_A(i,i)=1$ for all $i \in A$ we get

$$\frac{W(x_t)}{t^2} = \frac{1}{2}\sum_{i,j \in A} s_A(i,j)u_j f(x,i,t)f(x,j,t) -$$

$$\frac{1}{2t}\sum_{i=1}^{n} u_i f(x,i,t).$$

Hence $\frac{W(x_t)}{t^2}$ can be approximated by the normalized potential (defined on the simplex)

$$w(f)=\frac{1}{2}fSf^T$$

where $f=(f_1,...,f_n)$ is a frequency vector and the matrix S is defined by $S_{ij}=s_A(i,j)u_j$. Since S is symmetric, it can be diagonalized by an orthonormal matrix. That is, there exists an $(n\times n)$ matrix P with $P^T=P^{-1}$ such that P^TSP is diagonal, with the eigenvalues of S along its main diagonal. Since the matrix P can be thought of as rotation in the bundle space, we may offer the following interpretation: the consumer is deriving utility from n "basic commodities", which are the eigenvectors of S, and their u-values are the corresponding eigenvalues. One such commodity may be, for instance, some combination of *tea* and *sugar*. However, the consumer can only purchase the products *tea* and *sugar* separately. By choosing the "right mix" of the products, it is as if the "basic commodity" was directly consumed.

According to this interpretation, there is zero similarity between the "basic commodities"; that is, there are no substitution or complementarity effects between them. These effects among the actual products are a result of the fact that these products are ingredients in the desired "basic commodities". (Note, however, than if some of the eigenvalues of S are multiple, there is some arbitrariness in determining the "basic commodities".)

9. Concluding Comments

9.1 Our model does not offer any explanation of how the aspiration level is determined, adjusted over time and so forth. Undoubtedly, a variety of psychological and sociological considerations are involved in these processes. For instance, one may postulate that the aspiration level of our consumer depends on the well-being and market performance of other consumers around him/her. Thus a consumer who is initially easily satisficeable may start consuming at a corner point of the products' simplex. Then, as a result of interaction with others, the consumer may become more "aggressive", as if "expecting" more of every product since (s)he has become aware of better possibilities. As this paper focuses on individual consumer's behavior, we do not address these issues here.

9.2 The notion of "aspiration level" in this model need not always be taken literally. Consider, for instance, the choice of a piece of music to listen to. We may find that our consumer listens to "Don Giovanni" three times as often as to "Turandot". In our model, the former would have a utility of -1, while the latter -- of -3. Yet it would be wrong, if not blasphemous, to suggest that these works do not achieve our consumer's "aspiration level". The fact that "Turandot" is chosen from time to time may be for the sake of change, without any implications of inadequacy regarding Mozart's work.

One may insist that the very need for change and the risk of boredom are, in a way, indications that the consumer's aspiration level has not been reached. Perhaps. In this case zero utility level may better be interpreted as a "bliss point" rather than the "aspiration level". At any rate, our theory, with negative utility values, need not conjure up sour faces in one's mind. It may also apply to perfectly content decision makers who merrily and gingerly alternate choices to keep their lives interesting.

On the other hand, a satisficed consumer need not be "happy" in any intuitive sense. For instance, consider the same consumer who, unrelated to the operas mentioned above, suffers from an acute headache. Facing this problem repeatedly, the consumer may be loyal to a certain brand of medication, thus (s)he is satisficed according to our model. Yet (s)he is by no means happy.

More generally, one may adopt a psychological distinction and attempt a tentative classification of consumption activities to "pleasure seeking" vs. "pain avoiding." While CBCT appears to apply to both types of choices, it seems that when the choice is conceived as pleasure seeking, as in the example of a concert, zero utility level is best interpreted as "bliss point" (which may still be attained.) By contrast, when the choice is conceived as designed to avoid (or stop) a negative experience, as in the case of medication, the "aspiration level" interpretation seems more plausible.[3]

At any rate, a consumer may be satisficed in the formal sense without being happy in the intuitive sense, and vice versa. While the interpretations of zero utility level as "aspiration level" or "bliss point" hopefully enhance the cognitive plausibility of the model, our definition of zero u-value is behavioral: this is the level of satisfaction beyond which the consumer does not tend to switch or to further experiment new products.

---

3 This point is due to Eva Gilboa.

9.3 As mentioned above, the assumption that the similarity function (over problems) is constant is highly unrealistic. Allowing for more interesting functions will naturally broaden the scope of phenomena CBCT may explain. One illustration is the following: assume that the similarity function depends only on time, but that it may be non-monotone. Specifically, set

$$s(i,j)=\begin{cases}1 & if \quad 10\le |i-j|\le 30\\ 0 & \quad otherwise\end{cases}.$$

Next assume that the consumer has to choose between two options, a and b, whose u-values are identical and negative. One possible U-maximizing sequence would start with 10 a's, to be followed by a sequence of 20 b's, which will be followed by 20 a's and so on. Note that such a sequence (whose cycle length is 40) is inconsistent with our results, based on a constant similarity function, which would have predicted that a and b alternate at every stage (cycle length of two.) Since longer cycles appear in applications (say, fashion changes), it seems that non-constant similarity functions are of interest.

9.4 According to our model, a consumer who attempts to avoid boredom will choose the products along a sequence in $S(u)$ where u_i is negative for all i. For instance, this sequence may be $(beef, fish, chicken, beef, fish, chicken, \ldots)$. Such a consumer will enjoy the novelty of his/her dinner every day, and presumably would minimize boredom. However, the consumer may suffer "second-order boredom": while no dish is consumed twice within any three consecutive days, the sub-sequence $(beef, fish, chicken)$ repeats itself over and over again. The heroic attempt to avoid "first-order boredom" (due to the dish itself) results in extreme "second-order boredom"; the *way* in which boredom is avoided is boring itself.

There are many ways in which one may model higher-order boredom aversion. For instance, one may define an "actcome" to be a pair (or a triple) of products, and allow the consumer to choose at each stage only among the actcomes whose first component agrees with the second component of the last actcome chosen. (Say, after having consumed $(beef, chicken)$, the three available actcomes are $(chicken, beef)$, $(chicken, fish)$ and $(chicken, chicken)$.) Alternatively, one may include in the actcome a more explicit measure of diversity.

At any rate, extreme cases of boredom aversion are bound to result in a paradox: whatever is the theory we may come up with, our consumers may be bored with following it (even in the sense of choosing the same choice probabilities) and thus may choose to violate it. (This phenomenon is sometimes referred to as "Dostoyevsky's Paradox"; see Gilboa(1991).)

9.5 The model presented above is supposed to model "small" consumption decisions which are frequently repeated. Yet it may also have some relevance to "big" decisions, where the repeated choice is to retain the status quo. For instance, a consumer who buys a house and lives in it for five years may be viewed as making 60 monthly decisions to keep consuming it. Naturally, long-run commitments and transaction costs do not allow one to change one's residence every month. But the length of stay is likely to be related to one's aspiration level in a similar way to the case of short-run decisions regarding non-durable goods.

9.6 Our primary motivation for and interpretation of our theory is descriptive. However, under some conditions it may be interpreted as a normative one as well. For instance, while low aspiration levels may lead one to be "too easily" satisficed, there is nothing patently irrational about the case-based consumer who switches among alternatives to avoid boredom. Similarly, a consumer who increases his/her value-of-money parameter, but ends up spending more money than before may be dubbed "irrational". Yet a "typical" consumer who thus reduces the total expenditure achieves his/her goal. In short, whenever the boundedly rational is not obviously silly, it may serve as a normative theory, whose main advantage is its practicality.

9.7 A general comment on cognitive plausibility may be in order. As mentioned in the introduction above, economic theory as we know it today purports to be purely behavioral. The "Official Position" of the profession is that all theoretical constructs used in the model should be tied to observable (economic) decisions, and that they derive their meaning from that relationship solely. This belief is rooted in the logical positivist tradition in the philosophy of science, coupled with the view that only economic choices are to be considered "valid data" for economic theory. Correspondingly, the Official Position holds that only such economic phenomena can be used to test a theory or judge between competing theories. In particular, introspection and evidence such as "That is not how I think" are ruled out as irrelevant and as "metaphysical nonsense".

There is little doubt that the Official Position has done a great service to economic theory in focusing debates and avoiding unnecessary ones. We certainly believe it should and will remain an important guideline for the field; indeed, throughout this paper we made an attempt to relate theoretical to observable terms, and to verify that our theory may be refuted by certain behavior patterns.

However, we do not find the Official Position tenable in its strict sense, for the following reasons:

- If all theoretical constructs, which are *in principle* observable, were also *actually* measurable, and if economic theory, using these measurements, were to provide accurate predictions, then one could refer to the theoretical constructs and the theory's axioms as a tool "that works", regardless of their intuitive appeal. Needless to say, this is not the case. As a matter of fact, some -- if not most -- of the important applications of economic theory involve qualitative reasoning rather than quantitative measurement and prediction. Indeed, one may obtain far-reaching (as well as far-fetched) conclusions merely by analyzing an economic problem *in terms of* "utility", "optimality" and the like. However, these applications are not in complete compliance with the Official Position. Since the "utility" which is used in the analysis is seldom measurable, a certain degree of faith is required of the reader in order to follow the analysis and accept or dispute its conclusions.

In other words, the state of the art is that the Official Position is partly an accepted and followed methodology, partly a myth. While we believe it may be very dangerous to discard it altogether, one also has to admit the relevance of a theory's intuitive appeal, which eventually determines the level of "faith" one scientist needs to have in order to accept another's arguments.

- The logical positivists' "Received View", which is at the heart of economic theory's Official Position, has been widely criticized in the philosophy of science and it is today largely discredited. In particular, it has been argued that the very dichotomy between "theory" and "observations" is artificial, context-dependent, quantitative or even impossible. Furthermore, it is claimed that the observations are often "theory-filtered", that is, that the theory one has in one's mind will determine the "observations" one gathers.

We find this last point of particular relevance to the social sciences, where even laboratory experiments are not completely isolated from the outside world. A subject's behavior in an experiment can seldom be interpreted without any reference to the person's conceptualization of economic environments at large. Thus, two economic theorists may very often interpret given data differently, each contending it is consistent with one's own theory. Furthermore, when the contested theories are formulated in different languages, it may be hard to obtain agreement on what it is that the two have observed.

Given this state of affairs, we believe that cognitive plausibility is a relevant criterion for the evaluation of a theory. Should one theory be more intuitive than another, we may prefer it as a way to conceptualize reality even if we cannot point to a test that undeniably falsified the latter but not the former.[4]

9.8 The above discussion should be taken as a general point regarding economic theory; it should *not* be taken to imply that we always find CBCT more intuitively appealing, or cognitively plausible than neo-classical consumer theory, nor that intuitive appeal is the only way to differentiate between them.

However, we claim that for a wide variety of applications CBCT holds a promise to provide more faithful descriptions and more accurate predictions of consumer choices than the neo-classical theory.

---

4 The argument in favor of intuitive appeal may also be reduced to an argument for simplicity, where one's intuition is taken to be part of the "observations" one has to explain, together with the "outside world" observations. Needless to say, both "intuitive appeal" and "simplicity" are inherently subjective criteria regarding which we can only hope some agreement would exist.

References

Allais, M. (1953). "Le Comportement de L'Homme Rationel devant le Risque: critique des Postulates et Axioms de l'Ecole Americaine." Econometrica **21**: 503-546.

Ellsberg, D. (1961). "Risk, Ambiguity and the Savage Axioms." Quarterly Journal of Economics (75): 643-669.

Gilboa, I. (1991). Rationality and Ascriptive Science. Northwestern University.

Gilboa, I. and D. Schmeidler (1992). Case-Based Decision Theory. Northwestern University.

Keynes, J. M. (1921). A Treatise on Probability. London, MacMillan and Co.

March, J. G. and H. A. Simon (1958). Organizations. New York, Wiley.

Riesbeck, C. K. and R. C. Schank (1989). Inside Case-Based Reasoning. Hillsdale, NJ, Lawrence Erlbaum Associates, Inc.

Savage, L. J. (1954). The Foundations of Statistics. New York, John Wiley and Sons.

Selten, R. (1978). "The Chain-Store Paradox." Theory and Decision **9**: 127-158.

Simon, H. A. (1957). Models of Man. New York, Wiley.

Appendix: Proofs and Related Analysis

Proof of Proposition 1

Let us begin with some preliminary analysis. First assume that $u_i > 0$ for some $i \in A$. Then there may be (no more than $(n-1)$) choices of actcomes with negative utility, until the first one with $u_i > 0$ is chosen. From that point on, that i is chosen for ever. Hence in this case $f(x)$ exists for all $x \in S(u)$, and it is one of the extreme points of the $(n-1)$-dimensional simplex. (Note, however, that it need not be the same for all such x's.)

Let us now consider the case of $u_i < 0$ (i.e., $a_i > 0$) for all $i \in A$. First observe that all the actcomes will be chosen infinitely often in each $x \in S(u)$: since some of them must be chosen infinitely often, the U-value of the chosen (maximal) ones is not bounded from below; hence the U-value of all of them must tend to $-\infty$.

Furthermore, at stage $t \geq 0$ actcome i is (weakly) preferred to j iff

$$U(x,i,t) = F(x,i,t)u_i \geq F(x,j,t)u_j = U(x,j,t).$$

or, equivalently,

$$F(x,i,t)a_i \leq F(x,j,t)a_j.$$

Hence for *any* stage $t \geq 0$, regardless whether i is to be chosen at it or not, we have

$$F(x,i,t) \leq \frac{a_j}{a_i} F(x,j,t) + 1.$$

(Formally, the above is proven by induction on $t \geq 0$ using the previous inequality.) Thus, for $t \geq 1$ we obtain

$$f(x,i,t) \leq \frac{a_j}{a_i} f(x,j,t) + \frac{1}{t}.$$

For $t \geq n$ none of the frequencies vanishes, and it follows that

$$\exists \lim_{t\to\infty} \frac{f(x,i,t)}{f(x,j,t)} = \frac{a_j}{a_i}.$$

However, with finitely many actcomes, this also implies that $f(x)$ exists. Furthermore, we find that it is independent of x. Finally, since

$$b_i \equiv \prod_{j\neq i} a_j$$

also satisfy ${b_i}/{b_j} = {a_j}/{a_i}$, $b = (b_1,\ldots,b_n)$ must be proportional to $f(x)$, and we conclude that

$$f(x,i) = \frac{\prod_{j\neq i} a_j}{\sum_{k\in A} \prod_{j\neq k} a_j} \quad .$$

To sum, we have proven that $f(x)$ exists in both cases, hence (i) is proven. If there are $u_i > 0$, only extreme points could be limit frequencies; on the other hand, if $u_i < 0$ for all $i \in A$, only interior points can be chosen. Thus the existence of one $x \in S(u)$ for which $f(x)$ is an extreme point implies that for some actcome i, $u_i > 0$, whence for all $x \in S(u)$ $f(x)$ is an extreme point of the simplex. This concludes the proof of (ii). Claims (iii) and (iv) were explicitly proven above. We are left with claim (v), i.e. that for every interior point y in the $(n-1)$-dimensional simplex there exist negative utility indices $(u_i)_{i\in A}$ such that $f(x) = y$ for all $x \in S(u)$, and that these are unique up to a multiplicative scalar. Let there be given such a point $y = (y_1,\ldots,y_n)$. Define

$$u_i = -a_i = -\prod_{j\neq i} y_j$$

such that for all $i,j \in A$

$${y_i}/{y_j} = {a_j}/{a_i}.$$

It follows that $f(x) = y$ for all $x \in S(u)$. Furthermore, this equation shows that the utility $(u_i)_{i\in A}$ is unique up to multiplication by a positive scalar. This completes the proof.

Remark

The proof above shows that there is some technical duality between utility and the frequency of choice. Consider the case of negative utility values and assume that the vector $a=(a_1,\dots,a_n)$ is normalized such that $\sum_{i=1}^{n} a_i = 1$. Again, let Δ^{n-1} denote the interior of the $(n-1)$-dimensional simplex, i.e.,

$$\Delta^{n-1} = \left\{ z=(z_1,\dots,z_n) \middle| \forall i,\ z_i > 0,\ \sum_{i=1}^{n} z_i = 1 \right\}$$

and define a function $g:\Delta^{n-1} \to \Delta^{n-1}$ by

$$g(z)_i = \frac{\prod_{j\neq i} z_j}{\sum_k \prod_{j\neq k} z_j}.$$

We have shown that for the utility vector $u=-a=(-a_1,\dots,-a_n)$, for all $x \in S(u)$, $f(x)=g(a)$. However, the function g is uniquely defined by

$$\frac{g(z)_i}{g(z)_j} = \frac{z_j}{z_i}.$$

Hence it satisfies $g^2 = 1$ (where 1 stands for the identity mapping.) Thus it is also true that $a = g(f(x)) = g(g(a))$. That is, the (negative of the) utility vector a can be computed from the relative frequencies f by the same function which computes f as a function of a.

It also follows that g has a unique fixed point, namely, $\left(\frac{1}{n},\dots,\frac{1}{n}\right)$. In the example given in section 3, all products had the same utility, and were consumed with the same frequency. Up to normalization of the vector $a=(a_1,\dots,a_n)$, which involves no loss of generality, the utility could therefore by *identified* with the consumption frequency. It turns out that this example was very special in this sense. However, it is generally true that the utility and frequency vectors are g-images of each other.

Proof of Proposition 2

Let us begin with (i). If $u_i > 0$ for some $i \in A$, there could be up to $(n-1)$ choices out of the actcomes whose utility is negative, and then one of those with the positive utility values will be chosen for ever. Hence there are finitely many possible sequences.

Now consider the case $u_i < 0$ for all $i \in A$. If a_i/a_j is irrational for all $i, j \in A$, after the first n stages, where each actcome is chosen once, U maximization determines the choice uniquely. Thus there are $n!$ sequences in $S(u)$. If, however, a_i/a_j is rational for some $i, j \in A$, say $a_i/a_j = l/k$ for some integers $l, k \geq 1$, after exactly l choices of actcome j and k choices of i, there will come a stage where both i and j are (among) the maximizers of U. Then the choice between them is arbitrary, and therefore there are (at least) two different continuations which are consistent with U maximization. Since for every choice made at this point there will be a similar one after $2l$ choices of actcome j and $2k$ choices of i, there are at least 4 such continuations at this stage. By similar reasoning, there are at least

$$\left|\{0,1\}^{\mathbb{N}}\right| = \aleph$$

sequences in $S(u)$. Since $\aleph$ is also an upper bound on $|S(u)|$, (ii) is established.

We need to show (iii). First consider S^-. Denote

$$S_p = \left\{ x \in A^\infty \middle| \exists f(x) = p \right\}$$

By a law of large numbers, $\lambda_p(S_p) = 1$. Hence $\lambda_p(S^- \setminus S_p) = 0$. It therefore suffices to show that $\lambda_p(S^- \cap S_p) = 0$. First consider the case where p_i/p_j is irrational for all $i, j \in A$. Then $S^- \cap S_p$ is finite and of measure zero. Next assume that $p_i/p_j = l/k$ for some integers $l, k \geq 1$. In this case $S^- \cap S_p$ is uncountable, but it is a subset of the event in which for every $m \geq 1$, the $(ml+1)$-th appearance of j occurs after the mk-th appearance of i. Since at every point there is a positive λ_p-probability of a sequence of $(mk+1)$ consecutive appearances of i, this event is of measure zero.

Finally, the set S^+ includes all sequences x such that for some $1 \leq k \leq n$

$$x(l) \neq x(j) \qquad \text{for all } l, j \leq k$$

and

$$x(l) = x(k) \qquad \text{for all } l > k.$$

Hence $|S^+| = \sum_{k=1}^{n} \frac{n!}{(n-k)!} < \infty$. Furthermore, this set is λ_p-null unless p is a degenerate probability vector. This completes the proof of the proposition.

Proof of Proposition 3

Let v, p^1, p^2 be given. First consider $i, j \in A$ such that

$$c_1 \equiv \frac{f(v,p^1)_i}{f(v,p^1)_j} \neq \frac{f(v,p^2)_i}{f(v,p^2)_j} \equiv c_2. \qquad (**)$$

From the equations

$$\frac{v_i - p_i^1}{v_j - p_j^1} = c_1 \qquad ; \qquad \frac{v_i - p_i^2}{v_j - p_j^2} = c_2$$

we derive the system

$$\begin{pmatrix} 1 & -c_1 \\ 1 & -c_2 \end{pmatrix} \begin{pmatrix} v_i \\ v_j \end{pmatrix} = \begin{pmatrix} p_i^1 - c_2 p_j^1 \\ p_i^2 - c_2 p_j^2 \end{pmatrix}$$

which has a unique solution. It only remains to note that, if $f(v,p^1) \neq f(v,p^2)$ then for every i there a j such that $(**)$ holds, hence v_i is uniquely determined.

Proof of Proposition 4

Let us study the frequency vector for large values of α. First consider $i, j \in A$ with $p_i, p_j > 0$. For these,

$$\frac{f(v,p,\alpha)_i}{f(v,p,\alpha)_j} \xrightarrow[\alpha \to \infty]{} \frac{p_j}{p_i}.$$

If, however, $p_i, p_j = 0$,

$$\frac{f(v,p,\alpha)_i}{f(v,p,\alpha)_j}=\frac{v_j}{v_i} \qquad \text{for all } \alpha \geq 0.$$

Finally, for $p_i > p_j = 0$,

$$\frac{f(v,p,\alpha)_i}{f(v,p,\alpha)_j} \xrightarrow[\alpha\to\infty]{} 0.$$

Thus for large enough α, the frequencies of the positive-price products are arbitrarily small, and consequently so is $E(v,p,\alpha)$. This proves part (i). Part (ii) of the proposition follows from the considerations above and the continuity of $E(v,p,\alpha)$ in α.

Proof of Proposition 5

Let us compare the relative frequencies of the choices for two values of the value-of-money parameter, say $\alpha_1 > \alpha_2$. Denote

$$r_{ij}(\alpha) \equiv \frac{f(v,p,\alpha)_i}{f(v,p,\alpha)_j} = \frac{v_j - \alpha p_j}{v_i - \alpha p_i}.$$

One may verify that, whenever $\alpha_1 > \alpha_2$, (and regardless whether v is typical w.r.t. [with respect to] p or not,)

$$r_{ij}(\alpha_1) > r_{ij}(\alpha_2) \qquad \text{iff} \qquad \frac{p_j}{p_i} > \frac{v_j}{v_i}.$$

(Notice that the expression on the right -- $\frac{v_j}{v_i}$ -- is the ratio of frequencies for $\alpha = 0$, while that on the left -- $\frac{p_j}{p_i}$ -- is the limit of this ratio when $\alpha \to \infty$.)

Next assume that $p_i < p_j$, or that $\frac{p_j}{p_i} > 1$. Since v is typical (with respect to p), it follows that $v_i < v_j$, that is, $\frac{v_j}{v_i} < 1$. (Recall that $v_i < 0$.) Hence, in particular, $\frac{p_j}{p_i} > \frac{v_j}{v_i}$, and $r_{ij}(\alpha_1) > r_{ij}(\alpha_2)$ follows for $\alpha_1 > \alpha_2$. Similarly, $p_i = p_j$ implies $v_i = v_j$, hence $\frac{p_j}{p_i} = \frac{v_j}{v_i}$ and $r_{ij}(\alpha_1) = r_{ij}(\alpha_2)$.

We will use the following

Lemma: Let $f=(f_1,\ldots,f_l)$ and $g=(g_1,\ldots,g_l)$ be two positive vectors such that $\sum_i f_i=\sum_i g_i$. Let $p=(p_1,\ldots,p_l)$ be a vector satisfying $p_1\le\ldots\le p_l$. Assume that for all $i<j$,

$$\frac{g_i}{g_j}\ge\frac{f_i}{f_j}.$$

Then $g\cdot p\le f\cdot p$.

If, furthermore, $p_{i+1}>p_i$ for some $i<l$, and

$$\frac{g_i}{g_{i+1}}>\frac{f_i}{f_{i+1}},$$

then $g\cdot p<f\cdot p$.

Proof of Lemma:

By induction on l. For the case $l=2$ the conclusion is straightforward. Assume, then, that the lemma holds for $k<l$ and consider $k=l$.

Observe that

$$g\cdot p=\sum_i g_i p_i=f_1p_1+(g_1-f_1)p_1+\sum_{i>1}g_ip_i;$$

Since $\frac{g_1}{f_1}\ge\frac{g_i}{f_i}$ for all $i>1$ and $\sum_i f_i=\sum_i g_i$, it follows that $g_1\ge f_1$. Hence

$$g\cdot p\le f_1p_1+(g_1-f_1)p_2+\sum_{i>1}g_ip_i.$$

Furthermore, if $p_2>p_1$ and $\frac{g_1}{g_2}>\frac{f_1}{f_2}$, $g_1>f_1$ and the above inequality is strict.

Define $\overline{f}=(\overline{f_2},\ldots,\overline{f_l})$, $\overline{p}=(\overline{p_2},\ldots,\overline{p_l})$ and $\overline{g}=(\overline{g_2},\ldots,\overline{g_l})$ by

$$\overline{f_i}=f_i \qquad \forall i\ge 2;$$
$$\overline{p_i}=p_i \qquad \forall i\ge 2;$$

$$\overline{g_2} = g_2 + (g_1 - f_1)$$
$$\overline{g_i} = g_i \qquad \forall i > 2.$$

Since for $j \geq i \geq 2$, $\frac{\overline{g_i}}{\overline{g_j}} \geq \frac{g_i}{g_j} \geq \frac{f_i}{f_j} = \frac{\overline{f_i}}{\overline{f_j}}$, and $\sum_{i>1} \overline{f_i} = \sum_{i>1} \overline{g_i}$, we may use the induction hypothesis for $k = l - 1$ and the vectors $\overline{f}, \overline{g}$ to conclude that

$$(g_1 - f_1)p_2 + \sum_{i>1} g_i p_i = \overline{g} \cdot \overline{p} \leq \overline{f} \cdot \overline{p}.$$

Finally,

$$g \cdot p \leq f_1 p_1 + (g_1 - f_1)p_2 + \sum_{i>1} g_i p_i = f_1 p_1 + \overline{g} \cdot \overline{p} \leq f_1 p_1 + \overline{f} \cdot \overline{p} = f \cdot p\ ,$$

which completes the proof of the main claim of the lemma.

Let us now turn to the "furthermore" part. Assume that for some $i < l$, $p_{i+1} > p_i$ and $\frac{g_i}{g_{i+1}} > \frac{f_i}{f_{i+1}}$. If $i = 1$, then, as noted above,

$$g \cdot p < f_1 p_1 + (g_1 - f_1)p_2 + \sum_{i>1} g_i p_i$$

and it suffices that $\overline{g} \cdot \overline{p} \leq \overline{f} \cdot \overline{p}$ to show $g \cdot p < f \cdot p$. If, however, $i > 1$, we conclude the proof by observing that $\frac{\overline{g_i}}{\overline{g_{i+1}}} > \frac{g_i}{g_{i+1}}$ and using the ("furthermore" part of the) induction hypothesis.

Let us now turn back to the proof of the proposition. Assume w.l.o.g. that $p_1 \leq \ldots \leq p_n$. Denote $f = f(v, p, \alpha_2)$ and $g = f(v, p, \alpha_1)$. For $i < j$

$$r_{ij}(\alpha_1) \geq r_{ij}(\alpha_2)$$

or

$$\frac{g_i}{g_j} \geq \frac{f_i}{f_j}.$$

Finally, $\sum_i f_i = \sum_i g_i = 1$ and the lemma may be used to derive part (i) of the proposition, i.e.,

$$E(v,p,\alpha_1) = g \cdot p \leq f \cdot p = E(v,p,\alpha_2) \qquad\qquad \text{for } \alpha_1 > \alpha_2.$$

We now turn to prove part (ii). If $\max_{i\in A} p_i > \min_{i\in A} p_i$, there exists $i < n$ with $p_{i+1} > p_i$. Then

$$r_{i,i+1}(\alpha_1) > r_{i,i+1}(\alpha_2)$$

or

$$\frac{g_i}{g_{i+1}} > \frac{f_i}{f_{i+1}}$$

and

$$E(v,p,\alpha_1) = g \cdot p < f \cdot p = E(v,p,\alpha_2)$$

follows from the "furthermore" part of the lemma. On the other hand, if $\max_{i\in A} p_i = \min_{i\in A} p_i = \hat{p}$, then $E(v,p,\alpha) = \hat{p}$ holds for all α. This completes the proof of claim (ii).

Finally, claim (iii) is a straightforward corollary of proposition 4 and claim (ii).