INTERNATIONAL
ECONOMIC
REVIEW

November 1997
Vol. 38, No. 4

CUMULATIVE UTILITY CONSUMER THEORY*

BY ITZHAK GILBOA AND DAVID SCHMEIDLER[†1]

Tel Aviv University, Israel, and Boston University, U.S.A.
Tel Aviv University, Israel, and Ohio State University, U.S.A.

This paper makes some preliminary steps towards a dynamic theory of consumer choices, restricted to the case of repeated small decisions. We assume that the consumer chooses among products, rather than among bundles, and that she bases her decision on a cumulative satisfaction index. The (instantaneous) utility of a product is closely related to the relative frequency with which it is consumed. The aggregation of choices among products implicitly defines a choice of a bundle. We propose to incorporate a product's price directly into its utility, and show that it is consistent with balancing the consumer's budget.

1. INTRODUCTION

Neoclassical economic theory depicts the consumer as a rational agent who chooses a product bundle that maximizes a utility function given the budget constraint. While this model is a powerful and insightful tool, it suffers from two major limitations this paper attempts to address. First, it is a static model that does not describe the dynamic process by which a consumer may evolve to choose an optimal bundle. Second, there are many cases in which it does not seem to be a good description, however idealized, of how consumers actually make decisions.

Should we be concerned about these limitations? We argue that we should, and that predictive power may be gained if we refine the theory, or suggest alternatives

* Manuscript received May 1994; revised June 1996.

† E-mail: *gilboa@bu.edu; schmeidler.1@osu.edu*

[1] 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 Osaka University, Japan, Tel-Aviv University, Israel, and the Universitat Autònoma de Barcelona, Spain. Special thanks are due to Enriqueta Aragones, to Simon Grant (who served as a referee), and to an anonymous referee for many specific comments on earlier versions of this paper. A preliminary version of this paper was discussed at an ISER seminar in conjunction with the Osaka University Semicentennial Symposium in November 1992.

that would be more satisfactory in these respects. Consider dynamics first. Even if it is the case that consumers invariably converge to choose optimal bundles, economic systems spend considerable amount of time away from equilibrium. Hence the dynamic patterns of consumption, and, in particular, consumers' reactions to changes in their environment are of interest.

Second, we maintain that the neoclassical model is a reliable tool for predicting behavior mostly when it is cognitively plausible, that is, when it appears to correspond to the way consumers view their decision problem. Specifically, a consumer who faces a tight budget constraint (relative to her needs), is likely to plan her consumption carefully, imagine the various bundles she can afford, and choose accordingly. For such a consumer, the neoclassical theory provides a reasonable, if idealized model. However, when we consider a consumer whose income by far exceeds the cost of her essential needs, she is less likely to reason in terms of a budget constraint. A high income induces a very large choice set, making it very hard to imagine all possible bundles. Moreover, the consumer can *afford* not to plan; even if she happens to spend more money than an optimal plan would prescribe, she would still be able to meet the budget constraint.

Economic theorists are likely to accept this point, but to find it irrelevant to behavioral predictions. Indeed, the neoclassical theory has been derived from axioms on observed choice; if we believe that consumers satisfy these axioms, we are led to conclude that they behave *as if* they were maximizing a utility function subject to a budget constraint. Moreover, the axioms appear to be unassailably reasonable. It should then follow that the theory will provide good predictions even if it does not reflect consumers' cognitive processes.

We take issue with this tenet. We argue that the underlying axioms of consumer choice appear reasonable when formulated in the language of bundles. However, when consumers choose among *products* rather than among bundles, the 'observed' preferences are theoretical constructs. 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 neoclassical theory appears to be very reasonable when formulated in its own language, but this language is not always very intuitive.

Our long-run goal is therefore to develop a theory of consumer choice that will be dynamic in nature, and that will correspond to the way 'affluent' consumers make their decisions. It is our hope that such a theory can supplement the neoclassical one in two ways; first, it may describe how consumers who end up maximizing a utility function come to do so; and second, it may expand the set of consumers whose behavior can be reliably predicted.

This paper attempts to make a preliminary step in this direction. We focus on repeated, 'small' choices, with respect to which a consumer may be viewed as 'affluent.' The choice of a meal, of a means of transportation, or of a form of entertainment are all examples where many consumers can afford to make a decision without calculating precisely how much money they will consequently be left with. Because each decision is 'small,' the consumer knows that she will be able to pay her next month's rent regardless of her current choice. Correspondingly, in these examples it is impractical to make an 'optimal' choice by imagining all possible combinations of these small choices.

We suggest the following 'cumulative utility' consumer theory (CUCT). The consumer has a cumulative 'index of satisfaction' for each possible choice i, denoted $U(i)$. At every period, a maximizer of the satisfaction index is chosen. As a result of consumption, the satisfaction index of the product chosen is updated. In the simplest version analyzed here, the update is additive and constant. That is, some number u_i is added to $U(i)$ whenever i is chosen. If it is positive, the product appears to be better the more it has been consumed. If negative, the product is less desirable as a result of consumption. The consumer is implicitly assumed to be aware of $U(i)$, but not necessarily of u_i.

Consider a consumer for whom all products appear equally desirable at the outset. If there is one for which $u_i > 0$, as soon as it is chosen, it will be chosen forever. In this case, the consumer may be said to be 'satisficed' in the sense of Simon (1957). If, however, all products have negative u_i's, the consumer will never be satisficed. We therefore refer to the value zero as the 'aspiration level.' An individual with a low aspiration level, that is, $u_i > 0$, would tend to like a product more, the more it was consumed in the past, thus explaining habit-formation. By contrast, a high aspiration level implies that the consumer would appear to be 'bored' with products she recently consumed, and would seek change.

One may consider the vector of relative frequencies with which the products are consumed as the equivalent of the neoclassical 'bundle.' Viewed thus, if the aspiration level of the individual is low, she is easily satisficed and will select a corner point as an aggregate choice, as if she had a convex neoclassical utility function over the space of bundles. This corner point may not be 'optimal' in the usual sense (of maximizing u_i). If, however, the aspiration level is relatively high, the consumer would keep switching among the various products. The relative frequencies with which the various products are consumed will converge to a limit, which will be an interior point, as if the individual were maximizing a concave neoclassical utility over the bundle space.

A discussion of related literature might help to orient the reader. From a purely mathematical viewpoint, CUCT is a special case of 'case-based decision theory' (CBDT, see Gilboa and Schmeidler 1995). CBDT is a theory of decision making under uncertainty, suggesting that, in unfamiliar decision problems, people choose acts that have performed well in similar past cases. However, the decisions we address here are quite different: there is no underlying uncertainty, and the problem is assumed to be repeated, hence familiar. Indeed, the role of 'past cases' in CUCT is also different from that in CBDT: whereas under uncertainty past cases serve mainly as a source of information, in the present study they affect preferences directly.

Our discussion of small, repeated decisions may bring to mind Herrnstein and Prelec's 'melioration theory' (see Herrnstein and Prelec 1991). Indeed, the small decisions we study are very close to what they call 'distributed choice.' Yet our focus is different from theirs. Melioration theory deals primarily with choices that change intrinsically as a result of consumption. In examples such as addictive behavior, or other problems relating to self-discipline (regarding work, savings, and the like), the 'quality' of a choice depends on how often it has been chosen in the past. By contrast, we assume that the instantaneous value of a product, u_i, is constant. Our consumer may not wish to have the same meal every day; but there is no change in the pleasure derived from it as in the case of addiction, nor are there any long-term effects as in the case of savings. Further, melioration theory compares actual behavior to an optimum, defined by the highest average quality that can be obtained if inter-period interaction is taken into account. We study more commonplace phenomena, in which no such interaction exists. Correspondingly, in our model there is no well defined 'optimum,' and, at least when high aspiration levels are involved, we do not deal with patently suboptimal choice. Finally, the decision rules in the two models differ. Whereas melioration theory uses average, CUCT employs sums.

Our model is similar to the learning model of Bush and Mosteller (1955) in that, at every period, the outcome of the most recent choice is compared to an aspiration level, and the choice is reinforced if (and only if) its outcome has exceeded the aspiration level. However, there are two main differences. First, Bush and Mosteller are motivated by learning; thus, in their model past experiences serve mostly as sources of information, rather than determinants of preferences. Second, their model is inherently random, where a decision-maker's history is summarized by a probability vector, describing the probability of each choice. By contrast, in our model the choice is deterministic. Stochastic choice is useful to summarize 'hidden variables'; however, our goal here is to expose the decision process, and we focus on the deterministic implications of our basic assumptions.

Gilboa and Pazgal (1996) extend the present model by assuming that the instantaneous payoff u_i is a random variable, whose distribution depends only on the product. They also conduct an empirical study of the random-payoff model. Finally, the basic model suggested here was applied to electoral competition in Aragones (1994). She found that change-seeking behavior on the part of the voters gives rise to 'ideological' behavior of the competing parties.

The rest of this paper is organized as follows. Section 2 presents the formal model and a few results. In Section 3 we introduce the notion of the 'potential,' which facilitates the comparison of our model to the neoclassical one. Section 4 extends the model to incorporate complementarity and substitution effects. In Section 5 we model prices and the budget constraint. Section 6 concludes.

2. THE BASIC MODEL

Let the set of products be $A = \{1, 2, \ldots, n\}$, and denote the instantaneous utility function for $i \in A$ by u_i. Since the interesting case will be that of negative utilities, it will prove convenient to define $a_i = -u_i$. For a sequence $x \in A^{\infty}$, define the number of appearances of choice $i \in A$ in 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 cumulative satisfaction index of product $i \in A$ at that stage:

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

For a vector $u = (u_1, \dots, u_n)$, let $S(u)$ denote the set of all sequences of choices that are stagewise U-maximizing. That is,

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

Of special interest will be the relative frequencies of the products, 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.) The omission of the product index will be understood to refer to the corresponding vector:

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

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

Finally, denote

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

and

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

We can now formulate our first result:

PROPOSITION 1. *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*

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

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

Proofs are in the 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.

This result shows that 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.

Proposition 1 also 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 three 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 neoclassical theory. On the one hand, the fact that product 2 has a lower utility than product 1 does not imply that it will never be chosen. It will, however, be chosen less often than its substitute. On the other hand, the utility indices do not merely rank the products; they also provide the frequency ratios, and are therefore cardinal.

Proposition 1 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 $i, j \in A$. Thus the function u can be derived from observed choice, and it is unique up to a multiplicative scalar.

We also find that *any* point in the interior of the simplex may be the result of aggregate consumption by a cumulative utility 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. Are there any choices it precludes? The following result shows that in an appropriately defined sense, very 'few' sequences are compatible with CUCT.

PROPOSITION 2. *Assume that* $u \in Y$. *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*

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

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 an uncountable subset of a* λ_p*-null set, whereas* S^+ *is finite, and if* p *is not degenerate,* S^+ *is a* λ_p*-null set.*

Thus there are uncountably many sequences of choices that are U-maximizing. 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, it is easy to verify that finitely many observed choices may often suffice to conclude that they cannot be a prefix of any sequence in S.

To sum, cumulative utility consumer theory is nonvacuous. On the contrary, it is too easily refutable. However, Gilboa and Pazgal (1996) show that, if the u_i's as well as the initial values $U(\cdot, i, 0)$ are random variables, one may choose their distributions so that every finite sequence of choices would have a positive probability.

3. THE POTENTIAL

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 by $A^* = \bigcup_{t \geq 0} A^t$. For a function Y: $A^* \to \Re$, $x_t \in A^t$ and $i \in A$, define

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

That is, $\partial Y / \partial i$ is the change in the value of Y that will result from adding i to the consumption vector x_t. Denote by U_i: $A^* \to \Re$ the U-value of product i, that is, $U_i(x_t) = U(x, i, t)$. Then u_i may be viewed as the derivative of U_i w.r.t. (with respect to) consuming product i:

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

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

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

Suppose we wish to measure the 'well-being' of a consumer at a certain time. One possibility to do so is by the 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 $x_t \in A^t$. A U-maximizing consumer may be described as maximizing her W function at any given 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 is closer to the neoclassical utility function than are either U or u. Both u and U are defined for a single product, whereas the neoclassical function is defined for bundles. Correspondingly, if past experiences are assumed to linger in the consumer's memory, neither U nor u attempt to capture the 'overall well-being' of the consumer, while the neoclassical utility does. By contrast, the potential function W is defined for 'bundles' (implicit in the vector x_t), and may be interpreted as a measure of well-being.

Furthermore, since

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

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 a pattern of consumer choices that is also predicted by a concave neoclassical utility function, while positive u values correspond to a convex one. Thus the potential parallels the neoclassical utility function also in terms of the relationship between concavity/convexity and interior/corner solutions. Technically, the potential is a rather different creature from the neoclassical utility function. While the former is defined on consumption sequences, the latter is defined on bundles. However, since all permutations of a given sequence are equivalent in terms of the behavior they induce, a sequence may be identified with the relative frequencies of the products in it. It follows that, after an appropriate normalization, we may redefine the potential on the bundle simplex.

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, that is,

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

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. We obtain

$$\begin{aligned} W(x_t) &= \sum_{\tau=1}^{t} U(x, x(\tau), \tau - 1) \\ &= \sum_{i=1}^{n} \sum_{k=1}^{F(x,i,t)} U(x, i, T(x,i,k) - 1) \\ &= \sum_{i=1}^{n} \sum_{k=1}^{F(x,i,t)} (k-1) u_i \\ &= \frac{1}{2} \sum_{i=1}^{n} u_i F(x,i,t)[F(x,i,t) - 1]. \end{aligned}$$

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,

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

At any given time $t \geq 0$, the consumer has a value for the normalized potential, and behaves as if she 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

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

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

$$f_i u_i = \text{const}.$$

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 it is tempting to identify the (normalized) potential with the neoclassical utility. However, some distinctions should be borne in mind. First, the neoclassical utility is assumed to be globally maximized by a one-shot decision. The potential is only locally improved at every stage. Should the potential be convex, for instance, our consumer may be satisficed without optimizing it. Second, the neoclassical utility is assumed to be maximized *given* the budget constraint. By contrast, the local improvement of the potential is *unconstrained.* We later introduce prices into our model in a way that retains this feature, that is, that lets the consumer follow the gradient of greatest improvement, where prices and income are internalized into the evaluation of possible small changes.

Yet, as long as prices and income are ignored, and as long as the potential is concave, our model may be viewed as a dynamic derivation of neoclassical utility maximization. Specifically, a neoclassical consumer who happens to have a concave, quadratic, and additively separable utility function over the simplex of frequencies will end up making the same aggregate choice as the corresponding unsatisficed cumulative utility consumer. CUCT thus provides an account of how such a consumer gets to maximize her utility.

Starting with a neoclassical utility, and adopting the identification of quantities with frequencies, one may suggest a hill-climbing algorithm as a reasonable dynamic model of consumer optimization, independently of CUCT. From this viewpoint, our model provides a cognitive interpretation of the gradients considered by the optimizing consumer.

4. SUBSTITUTION AND COMPLEMENTARITY

Consider a consumer who chooses a daily meal out of {*beef*, *chicken*, *fish*}. Suppose that she is indifferent among them, but seeks change. Say,

$$u_{\text{beef}} = u_{\text{chicken}} = u_{\text{fish}} = -1.$$

If the consumer judges beef as closer to chicken than fish is, after having chicken she might prefer fish to beef. To capture these effects, we introduce a similarity function between products, s_A: $A \times A \rightarrow [0,1]$, that reflects the fraction of one product's instantaneous utility that is added to another product's cumulative satisfaction index.[2]

Specifically, every time product i is consumed, $s_A(j,i)u_i$ is added to $U(j)$, for every product j, with $s_A(i,i)=1$. For instance, suppose that

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

where the other s_A values are zero. For these similarity values, when the choice set is {*chicken*, *fish*} or {*beef*, *fish*}, the relative frequencies of consumption are $(1/2, 1/2)$. However, for the set {*beef*, *chicken*, *fish*} they may be $(1/4, 1/4, 1/2)$ rather than $(1/3, 1/3, 1/3)$, as would be the case in the absence of product similarities. In fact, in this example *beef* and *chicken* are practically identical, and we can only say that the frequencies of {*beef/chicken*, *fish*} will converge to $(1/2, 1/2)$.

One interpretation of the product-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 does not affect the desirability of the other.

Next suppose that the recurring choice in our model is a purchase decision, rather than a consumption decision. In each period the consumer chooses a single product, but she also has a supermarket cart, or kitchen shelves, which allow her to eventually 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 product-similarity function to negative values in order to model complementarities. If the product-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.

Recall that the instantaneous utility function u may be interpreted as the consumption-derivative of U:

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

In the presence of product-similarity, when product i is chosen, $U(j)$ is changed by

[2] This notion corresponds to 'act similarity' in Gilboa and Schmeidler (1995). We use the subscript A to distinguish this function from the problem-similarity function in CBDT, which is here assumed constant. See also Gilboa and Schmeidler (1997) for axiomatization of CBDT with problem-act similarity function.

$s_A(j,i)u(i)$. That is,

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

Combining the above, the similarity function is 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 neoclassical theory defines the substitution index between two products as the cross derivative of the utility function (with respect to product quantities). By comparison, the product similarity function is the cross derivative of the potential function, normalized by its second derivative w.r.t. one of the two. Furthermore, if we define substitutability between two products i and j as the impact consumption of i has on the desirability of j, namely $s_A(j,i)u_i$, it precisely coincides with the cross derivative of the potential. This reinforces the analogy between the potential in CUCT and the neoclassical utility.

In the presence of product similarity, the consumer's 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

$$\begin{aligned} U(x,i,t) &= \sum_{\tau=1}^{t} s_A(i, x(\tau))u_{x(\tau)} \\ &= \sum_{j=1}^{n} F(x,j,t)s_A(i,j)u_j, \end{aligned}$$

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

$$\begin{aligned} W(x_t) &= \sum_{\tau=1}^{t} U(x, x(\tau), \tau-1) \\ &= \sum_{\tau=1}^{t}\sum_{\nu=1}^{\tau-1} s_A(x(\tau), x(\nu))u_{x(\nu)}. \end{aligned}$$

It is easy to see that the potential will be invariant with respect to 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 product 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. Hence, we define symmetry by $s_A(j,i)u_i = s_A(i,j)u_j$, that is, by the equality of the cross-derivatives of the potential. Under this assumption,

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

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

Since $s_A(i,i) = 1$ 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 $W(x_t)/t^2$ can be approximated by the normalized potential (defined on the simplex)

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

where $f = (f_1, \ldots, f_n)$ is a frequency vector and the matrix S is defined by $S_{ij} = s_A(i,j)u_j$.

Suppose that S is negative definite. In this case the hill-climbing algorithm implemented by a cumulative utility consumer will result in maximization of the normalized potential. Hence, if we start out with a quadratic and concave neoclassical utility u, it defines a matrix S for which the corresponding cumulative utility consumer behaves as if she were locally maximizing u. Furthermore, for any function u that can be locally approximated by a quadratic w, a local maximization u would result in similar choices to those of the cumulative utility consumer characterized by w.

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 $P^T S P$ 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. (These would be negative for a negative definite S.) One such commodity may be, for instance, a certain 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.[3]

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 the products are ingredients in the desired 'basic commodities.'

5. PRICES AND THE BUDGET CONSTRAINT

5.1. *Prices.* We wish to extend our model to deal with prices and income explicitly, without compromising cognitive plausibility. We also want to preserve the simple structure of repeated choice described above. Recall that we focus on a consumer who is, or can afford to be, boundedly rational in her treatment of the budget constraint. Suppose that such a consumer passes by a store and notices a product, say, a shirt. Being 'affluent,' the consumer is lucky enough to think along the lines of, 'This is a good price for this shirt' rather than of, 'I'll be better off if I buy this shirt and give up the concert tonight;' or 'This was a lot of money for this meal' rather than, 'Dining here three times is all I can afford this month if I am to pay my rent.' In other words, consumption decisions are often made by directly referring to price as a characteristic of each product, rather than by optimizing the utility function given the price vector.[4]

Assume that each product i has an 'intrinsic' utility value given by ν_i. These values designate the instantaneous utility that would correspond to the choice among the products, were they free. However, given the prices $p = (p_1, \ldots, p_n)$, the relevant utility is

$$u_i(p_i) = d(\nu_i, p_i)$$

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

We adopt a literal interpretation of these utility indices: the consumer feels happy if she gets a product at a 'bargain price,' and suffers disutility from the mere act of

[3] Note that the analogy to the additively separable case is not perfect. Specifically, the constraint that the sum of frequencies be 1 is also rotated. Hence the negative eigenvalues are not related to frequencies as directly as in Section 2.

[4] A 'rich' consumer may practically ignore prices, thinking along the lines of 'I like this shirt; I'll buy it.' This extreme case of affluence will also be captured by the model below.

paying. This need not imply that the consumer's self-esteem depends on her performance on the market. Rather, the consumer is rational enough to know that it is to her advantage to pay less. She may not have a clear plan for the usage of the rest of the money, but she has internalized the notion that money is good to have.[5]

For simplicity we consider a linear function

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

for some $\alpha > 0$. The coefficient α is a 'value of money' parameter, that 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 her consumption decisions.

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

$$\nu_1 = \nu_2 = -1;$$

$$p_1 = 1; p_2 = 2.$$

The consumer chooses a sequence out of $S((-2, -3))$, and will thus consume product 1 60 percent of the time, and product 2 -40 percent. 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, she would end up choosing a sequence in $S((-2, -4))$, that is, choosing product 1 about 67 percent of the time, and product 2 about 33 percent.

Comparative statics in this model thus reaffirm the basic intuition regarding substitute products: the consumer will use less of a product when its price goes up (other things being the same). Moreover, our model allows us also to ask what happens if we change the prices along the sequence of consumption choices. For instance, suppose that 10 days have passed 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 six times, and product 2—four 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. More generally, for any initial U-values, the long-run consumption frequencies would converge to (67%, 33%).

This model of product evaluation is purely retrospective in the sense that both desirability (ν_i) and price (p_i) are taken into account in a product's evaluation only *after* it has been consumed. One may also consider a (semi)-prospective model in which a product's desirability is combined with a posted price *prior* to making a purchase. Further, in case of price changes, one may postulate that the consumer updates her memory of past consumption to reflect the newly observed price. However, as long as prices are fixed from a certain period on, our asymptotic results would hold in these variants as well.

Retrospective evaluation is more consistent with the model presented above, whereas prospective evaluation involves more planning, and requires that the

[5] The introduction of prices into the utility function directly has been discussed in the literature by Kalman (1968) and Dusansky and Kalman (1972) as a generalization of the neoclassical theory. More generally, our model of a consumer who may not satisfy a budget constraint in the short run follows Marshall (1890). See also Friedman (1949) and Biswas (1977).

consumer be able to imagine her payoff if she were to consume each product at each posted price. Whether retrospective or prospective evaluation is more reasonable depends on several factors, such as the degree of precommitment and the cost of obtaining price information prior to making the decision. For instance, if the consumer walks into a restaurant, and then finds out what the prices are, or if she decides to use her car, and discovers that gasoline prices have gone up only at the gas station, retrospective evaluation appears plausible. By contrast, prospective evaluation seems more natural when the consumer chooses a cereal brand from the supermarket shelf. In this section we restrict attention to retrospective evaluation.[6]

To what extent are the parameters of $(*)$ observable? First, note that the observed relative frequencies are

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

Hence one cannot hope to observe more than the ratios ν_i/α. Differently put, we may assume without loss of generality that $\alpha = 1$. Under this assumption one can uniquely determine the vector ν by observing the consumer behavior for two price vectors that induce different choice frequencies. To be more precise, let us recall the definition of the 'intrinsic' utility space and define the price space to be, respectively,

$$V = \{\nu = (\nu_1, \dots, \nu_n) \mid \forall i,\ \nu_i < 0\}$$

$$P = \{p = (p_1, \dots, p_n) \mid \forall i,\ p_i \geq 0\}.$$

Let Δ^{n-1} denote the interior of the $(n-1)$-dimensional simplex, that is,

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

and define $f\colon V \times P \to \Delta^{n-1}$ by

$$\frac{f(\nu,p)_i}{f(\nu,p)_j} = \frac{\nu_j - p_j}{\nu_i - p_i}.$$

Then we have

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

It will be clear from the proof of this result that it can be extended to values of $\nu = (\nu_1, \dots, \nu_n)$ that are not necessarily negative, as long as the vectors $\nu - p^k$ are negative for $k = 1, 2$.

[6] An alternative, prospective model was suggested by Simon Grant.

5.2. *The Budget Constraint.* Affluent as our consumers may be, they may still find themselves with no financial resources left. When will this happen? How much money does a cumulative utility consumer spend?

While one may assume that $\alpha = 1$, it may be conceptually more convenient to think of a fixed 'intrinsic' value function ν and adjustable 'value of money' parameter α, that will eventually be used to balance the consumer's budget. The consumer chooses the various products according to the frequency vector $f = f(\nu, p, \alpha) \in \Delta^{n-1}$ defined by

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

Thus her average 'daily' expenditure is

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

This amount should be compared to the consumer's exogenously given daily income $I > 0$. Thus, if $I < E$, our consumer will face financial difficulties. On the other hand, if $I > E$, the consumer will have unused funds.

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 $I > \min_{i \in A} p_i$. We assume that $\min_{i \in A} p_i = 0$. In many applications this condition holds with no loss of generality, since one may include in the model a '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 that will balance the budget.

PROPOSITION 4. *Let there be given a value function* $\nu \in V$, *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* $I \geq E(\nu, p, \alpha)$;
(ii) *if* $I < E(\nu, p, 0)$, *then there exists* $\alpha > 0$ *such that* $I = E(\nu, p, \alpha)$.

The interpretation we have in mind is as follows. At a given point of time, the consumer has a value-of-money parameter α, and she spends $E(\nu, p, \alpha)$ money units per day. After a certain period, the consumer finds out how much she has spent in relation to her income. Realizing she has spent too much, she would start putting more weight on money in her decisions. In our model, this is reflected by increasing the parameter α. Conversely, the consumer may find out that she has unused income. She may then become less worried about money. The price of every product will seem less important, and this will be reflected by a lower value of α. If prices do not change over time, one could expect that in the long run α would be adjusted for a balanced budget. While there are many ways in which this process can be formalized, the study of specific adjustment mechanisms is beyond the scope of this paper.

We emphasize that the adjustment of α need not be conscious; it suffices that our consumer be appalled by her lavishness, or be traumatized by having exhausted the budget well before the end of the month, for her to start putting more weight on prices when making consumption choices. Similarly, a consumer who discovers she has been needlessly frugal will 'naturally' focus on the products' value more than on their cost, which is equivalent to lowering α.

It is interesting to note that even the very simple linear function $d(\nu_i, p_i)$ we employ in this model gives rise to nontrivial behavior of the expenditure function. Consider the following three-product example:

$$\nu_1 = -1;\ \nu_2 = -10;\ \nu_3 = -100;$$

$$p_1 = 1;\ p_2 = 2;\ p_3 = 0.$$

One may verify that $E(\nu, p, 0) \approx 1.08 < E(\nu, p, 10) \approx 1.17$. Thus the expenditure function may not be monotonically decreasing in its last argument. In this example, product 2, which is the most expensive one, is consumed at a higher frequency for $\alpha = 10$ than for $\alpha = 0$. Consequently, there are income levels for which the equation $I = E(\nu, p, \alpha)$ has multiple solutions, which induce different consumption patterns.

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 ten times more often than product 2. Yet product 2 is also more expensive. On the face of it, 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, she will choose the less expensive one more often. If these tastes were shared by all consumers, one would expect prices to reflect that. Hence there is some reason to believe that our consumer is not 'typical' in this respect.

Define a consumer's taste $\nu = (\nu_1, \dots, \nu_n) \in V$ to be *typical with respect to prices* $p = (p_1, \dots, p_n) \in P$ if

$$\nu_i < \nu_j \quad \text{iff} \quad p_i < p_j \qquad \forall i, j \in A.$$

Recall that the ν_i's are negative. The smaller is the value (the higher it is 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 $\nu \in V$ and a price vector $p \in P$, such that ν is typical with respect to p. Then*

(i) *$E(\nu, p, \alpha)$ is nonincreasing in α;*

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

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

REMARKS:

(i) 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 prices.

(ii) When $E(\nu, p, \alpha)$ is monotone in α, the unique 'balancing' value of α can be found by simple trial-and-error search procedures. Balancing the budget by adjusting one parameter is a much simpler task than finding an optimal bundle in higher-dimension spaces.

6. CONCLUDING COMMENTS

Consumers typically do not have a given budget for each type of repeated problem, such as the choice of annual vacation or of a daily meal. Our model therefore calls for a generalization, that would allow several repeated consumption problems to be solved simultaneously. Suppose that the consumer faces a sequence of decision problems, each of which is attached to a set of possible choices. The decision problems may include 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 into account in the expenditure function. It can be shown that balancing the budget for different income levels may result in different proportional expenditure on the various types of problems.

The expenditure function will naturally depend not only on utilities and prices, but also on the frequency with which various types of problems are encountered. Problems may present themselves to the consumer in a deterministic or stochastic fashion, exogenously or endogenously. That is, the consumer may decide to expose herself to decision problems (say, by going shopping), while certain problems would be determined externally (for instance, by hunger).

The aspiration level is a basic characteristic of a cumulative utility consumer. How is it determined and adjusted over time? A variety of psychological and sociological considerations are involved in these processes. For instance, one may postulate that a consumer's aspiration level depends on the perceived well-being of other consumers around 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 she has become aware of better possibilities.

Case-based decision theory stresses the similarity between a decision problem and past decisions. In the consumer theory developed here, the similarity function is implicitly assumed constant. Thus CUCT postulates that the consumer's impression of a product may be summarized by one number, namely, its U value. One may be interested in consumer choices that draw on memory of past 'cases' of consumption in more subtle ways. For instance, the weight assigned to past cases may exponentially decrease with time. (Incidentally, such a model does not require that the consumer remember more than one number per product. See Aragones 1994.) Alternatively, instances of consumption may differ in various attributes such as time of the day, recent choices, and so forth.

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. A consumer may listen 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 the consumer's 'aspiration level.' In this case zero utility level may better be interpreted as a 'bliss point.' Thus, negative utility values need not conjure up sour faces in one's mind. They may also describe perfectly content consumers 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 a consumer who often suffers from acute headaches. If she is loyal to a certain brand of medication, she is satisficed according to our model. Yet she is by no means happy.

One may adopt a psychological distinction between pleasure-seeking and pain-avoiding activities. While CUCT may apply to both types of choices, it seems that when the choice is perceived as pleasure seeking, as in the example of a concert, zero utility level is best interpreted as bliss point. By contrast, when the choice is perceived as designed to avoid (or stop) a negative experience, as in the case of medication, the aspiration level interpretation seems more plausible.[7]

APPENDIX

PROOF OF PROPOSITION 1. First assume that $u_i > 0$ for some $i \in A$. After at most $(n-1)$ choices of products with negative utility, one with $u_i > 0$ is chosen. From that point on, that product is chosen forever. 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 $i \in A$. At stage $t \geq 0$ product 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

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

[7] This point is due to Eva Gilboa.

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

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

With finitely many products, 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

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

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 product $i, u_i > 0$, 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), that is, 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 $u_i = -a_i = -\prod_{j \neq i} y_j$. 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.

Proof of Proposition 2. In light of the preceding analysis, (i) is immediate. 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 product 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 product j and k choices of i, there will come a stage where both i and j are among the maximizers of U. The choice at this stage is arbitrary, and therefore there are at least two different continuations that are consistent with U maximization. Since for every choice made at this point there will be a similar one after $2l$ choices of product j and $2k$ choices of i, there are at least four such continuations at this stage. By similar reasoning, there are at least

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

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

$$S_p = \{x \in A^\infty \mid \exists f(x) = p\}$$

By the strong 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 $(p_i/p_j) = (k/l)$ 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^+ is finite. Furthermore, it is λ_p-null unless p is a degenerate probability vector. This completes the proof of the proposition.

Proof of Proposition 3. Let ν, p^1, p^2 be given. First consider $i, j \in A$ such that

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

From the equations

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

we derive the system

$$\begin{pmatrix} 1 - c_1 \\ 1 - c_2 \end{pmatrix} \begin{pmatrix} \nu_i \\ \nu_j \end{pmatrix} = \begin{pmatrix} p_i^1 - c_2 p_j^1 \\ p_i^2 - c_2 p_j^2 \end{pmatrix},$$

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

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

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

$$\frac{f(\nu, p, \alpha)_i}{f(\nu, p, \alpha)_j} = \frac{\nu_j}{\nu_i} \qquad \text{for all } \alpha \geq 0.$$

Finally, for $p_i > p_j = 0$,

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

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

$$r_{ij}(\alpha) \equiv \frac{f(\nu, p, \alpha)_i}{f(\nu, p, \alpha)_j} = \frac{\nu_j - \alpha p_j}{\nu_i - \alpha p_i}.$$

One may verify that, whenever $\alpha_1 > \alpha_2$, (and regardless whether ν is typical w.r.t. p or not),

$$r_{ij}(\alpha_1) > r_{ij}(\alpha_2) \qquad \text{iff} \quad \frac{p_j}{p_i} > \frac{\nu_j}{\nu_i}.$$

Observe that the expression on the right ν_j/ν_i, is the ratio of frequencies for $\alpha = 0$, while that on the left, p_j/p_i, is the limit of this ratio when $\alpha \to \infty$.

Next assume that $p_i < p_j$, or that $(p_j/p_i) > 1$. Since ν is typical (with respect to p), it follows that $\nu_i < \nu_j$, that is, $(\nu_j/\nu_i) < 1$. (Recall that $\nu_i < 0$.) Hence, in particular, $(p_j/p_i) > (\nu_j/\nu_i)$, and $r_{ij}(\alpha_1) > r_{ij}(\alpha_2)$ follows for $\alpha_1 > \alpha_2$. Similarly, $p_i = p_j$ implies $\nu_i = \nu_j$, hence $(p_j/p_i) = (\nu_j/\nu_i)$ 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 $\Sigma_i f_i = \Sigma_i g_i$. Let $p = (p_1, \ldots, p_l)$ be a vector satisfying $p_1 \leq \cdots \leq p_l$. Assume that for all $i < j$,*

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

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

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

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

Since $g_1/f_1 \geq g_i/f_i$ for all $i > 1$ and $\Sigma_i f_i = \Sigma_i g_i$, it follows that $g_1 \geq f_1$. Hence

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

Furthermore, if $p_2 > p_1$ and $(g_1/g_2) > (f_1/f_2)$, $g_1 > f_1$ and the above inequality is strict.

Define $\bar{f} = (\bar{f}_2, \ldots, \bar{f}_1)$, $\bar{p} = (\bar{p}_2, \ldots, \bar{p}_1)$ and $\bar{g} = (\bar{g}_2, \ldots, \bar{g}_1)$ by

$$\bar{f}_i = f_i \qquad \forall i \geq 2;$$

$$\bar{p}_i = p_i \qquad \forall i \geq 2;$$

$$\bar{g}_2 = g_2 + (g_1 - f_1)$$

$$\bar{g}_i = g_i \qquad \forall i > 2.$$

Since for $j \geq i \geq 2$, $(\bar{g}_i/\bar{g}_j) \geq (g_i/g_j) \geq (f_i/f_j) = (\bar{f}_i/\bar{f}_j)$, and $\Sigma_{i>1} \bar{f}_i = \Sigma_{i>1} \bar{g}_i$, we may use the induction hypothesis for $k = l - 1$ and the vectors $\bar{f}, \bar{g}$ to conclude that

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

Finally,

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

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 $(g_i/g_{i+1}) > (f_i/f_{i+1})$. 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 $\bar{g} \cdot \bar{p} \leq \bar{f} \cdot \bar{p}$ to show $g \cdot p < f \cdot p$. If, however, $i > 1$, we conclude the proof by observing that $(\bar{g}_i/\bar{g}_{i+1}) > (g_i/g_{i+1})$ and using the ('furthermore' part of the) induction hypothesis. □

Let us now turn back to the proof of the proposition. Assume without loss of generality that $p_1 \leq \cdots \leq p_n$. Denote $f = f(\nu, p, \alpha_2)$ and $g = f(\nu, p, \alpha_1)$. 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, $\Sigma_i f_i = \Sigma_i g_i = 1$ and the lemma may be used to derive part (i) of the proposition, that is,

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

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

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

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 $E(\nu, p, \alpha) = \hat{p}$ holds for all α. This completes the proof of claim (ii).

Claim (iii) is a straightforward corollary of Proposition 4 and claim (ii).

REFERENCES

ARAGONES, E., "Negativity Effect and the Emergence of Ideologies," *Journal of Theoretical Politics* 9 (1997), 189–210.

BISWAS, T., "The Marshallian Consumer," *Economica* 44 (1977), 47–56.

BUSH, R.R., AND F. MOSTELLER, *Stochastic Models for Learning* (New York: John Wiley and Sons 1955).

DUSANSKY, R. AND P.J. KALMAN, "The Real Balance Effect and the Traditional Theory of Consumer Behavior: A Reconciliation," *Journal of Economic Theory* 5 (1972), 336–347.

FRIEDMAN, M., "The Marshallian Demand Curve," *Journal of Political Economy* 57 (1949), 463–495.

GILBOA, I. AND D. SCHMEIDLER, "Act Similarity in Case-Based Decision Theory," *Economic Theory* 9 (1997), 47–61.

——— AND ———, "Case-Based Decision Theory," *The Quarterly Journal of Economics* 110 (1995), 605–639.

——— AND A. PAZGAL, "History-Dependent Brand Switching: Theory and Evidence," mimeo, 1996.

HERRNSTEIN, R.J. AND D. PRELEC, "Melioration: A Theory of Distributed Choice," *Journal of Economic Perspectives*, 5 (1991), 137–156.

KALMAN, P.J., "Theory of Consumer Behavior When Prices Enter the Utility Function," *Econometrica* 36 (1968), 497–510.

MARSHALL, A., *Principles of Economics* (London: Macmillan 1890).

SIMON, H.A., *Models of Man* (New York: John Wiley and Sons 1957).