PRODUCT SUBSTITUTABILITY AND THE MATCHING LAW

Gordon R. Foxall and Victoria K. James
Cardiff Business School, United Kingdom

Jorge M. Oliveira-Castro
Universidade de Brasília, Brazil

Sarah Ribier
Montpellier SupAgro, France

The applicability of matching analysis, pioneered in the context of laboratory experiments, to the investigation and interpretation of consumer choice in natural environments is explored by the examination of sequential purchases of four product categories based on information from a panel of British consumers. Over a 52-week period, participants recorded data on the brands, quantities, and prices of their purchases. Matching analysis was employed, and individual and aggregated results showed that consumers generally behave according to the predictions of the matching law when the data is averaged over 3- and 5-week periods. The matching results were assessed in light of the results of a substitutability scale, which allowed consumers to state the perceived level of substitutability, independence, and complementarity for 13 product combinations. The results are generally supportive of the predictions made based on the matching results; however, the antimatching that was expected to characterize purchase patterns for complementary products was not observed.
Key words: matching law, food choice, deviations from matching, consumer choice, replication, extension, substitutability

The possibility that matching analysis reveals the degree of substitutability, complementarity, and independence among economic commodities has a considerable history in behavioral economics (Green & Freed, 1993; Rachlin, Battalio, Kagel, & Green, 1981). Ideal matching is typically found in situations in which the reinforcers offered on multiple schedules are functionally identical; heterogeneous reinforcers are related by undermatching, overmatching, or even antimatching. These relations are assumed, even taken for granted, in a wide range of basic behavioral economic and applied behavior analytical studies (Bickel & Vuchinich, 2000; Kagel, Battalio, & Green, 1995). Much of the evidence for these assumptions

derives from experimental analyses of choice in food-deprived nonhumans, from which extrapolations are sometimes made to the natural environments of purchase and consumption in modern, affluent marketing systems (e.g., Green & Freed, 1993). In this article, we investigate the assumption that matching analysis reveals the substitutability or nonsubstitutability of reinforcers that occur in everyday settings through the analysis of panel data on consumers' purchasing behavior for food brands and products.

When presented with a choice (competing opportunities to emit response x or response y), organisms allocate their responses according to the rates of reinforcement obtained from each alternative (Herrnstein, 1997). In other words, the response rate (*B*) is proportional to the relative rate of reinforcement (*R*). The matching relation therefore takes the form:

$$\frac{B_x}{B_x + B_y} = \frac{R_x}{R_x + R_y} \quad (1)$$

where B_x and B_y represent the rate of responding on options x and y, respectively, and R_x and R_y are the rates of reinforcers obtained from each choice. Expressed in terms of ratios, this relation becomes:

$$\frac{B_x}{B_y} = \frac{R_x}{R_y} \quad (2)$$

A generalized form of the matching law states that the ratio of responses between two alternatives is a power function of the ratio of reinforcements (Baum, 1974):

$$\log \frac{B_x}{B_y} = s \log \frac{R_x}{R_y} + \log b \quad (3)$$

This can be further expressed as power function:

$$\frac{B_x}{B_y} = \left(\frac{R_x}{R_y} \right)^b \quad (4)$$

The parameter *b* represents bias and constitutes the intercept of the linear log-log formulation of the law. Deviations of this parameter from unity are interpreted as a consistent preference for one choice alternative, independent of its reinforcement rate. If *b* > 1, the participant has a bias toward responding on x; if *b* < 1, the individual exhibits a bias toward y. In the marketing context, bias may result from the positioning of alternative brands within the store, the positioning and space allocated to different brands on the shelves for the particular product category, the positioning of substitute and complementary products, stockouts, and so on (Foxall & Schrezenmaier, 2003; Schrezenmaier, 2005). The parameter *s* represents sensitivity and constitutes the slope of the linear log-log formulation. It corresponds to a deviation from strict matching, indicating that the individual favors the richer (*s* > 1) or the poorer (*s* < 1) schedule of reinforcement more than strict matching would predict. The exponent indicates overmatching if *s* > 1 and undermatching if *s* < 1. When gross complements are involved (i.e., commodities for which a decrease in the price of one is followed by an increase in the consumption of both products), the choice ratio has an inverse relationship with the reinforcement ratio, showing the opposite of what the matching law predicts. This effect, which Kagel et al. (1995, pp. 27–59) refer to as antimatching, requires that *s* < 0.

Matching is, therefore, the behavioral phenomenon whereby human and nonhuman organisms allocate responses among alternatives in proportion to the relative reinforcement obtained from each. Matching is a molar process (concerned with the relationship of the *rate* of responding to *rate* of reinforcement) that is identifiable from comparison of the relative frequencies at which responses are emitted and reinforcement obtained. Matching is found in experiments in which alternative reinforcers are direct substitutes (Davison & McCarthy, 1988). Moreover, matching requires both perfect substitutability and the exclusive control of behavior by the nominal reinforcement frequencies (Heyman, 1996). Substitutability itself inheres in the similarity of the functional attributes of economic goods and other reinforcers (Green & Freed, 1993), and this is expected by economists to manifest in the relationship between the goods in question and changes in their relative prices. Hence, in economic analysis, two commodities, x and y, are substitutes if a reduction in the price of x leads to an increase in the quantity demanded of x and a decrease in the quantity demanded of y. Strongly competing brands in the same product category can be considered substitutes. Complementarity is the converse: A reduction in the price of x leads to an increase in quantity demanded of both x and y, for example, when a fall in the price of cheese is followed by increased purchases of both cheese and crackers. Commodities are independent if a change in the price of one has no effect on the quantity demanded of the other: Shirts and suntan lotion are independent products in this sense.

The widespread view that the degree of substitutability between commodities is captured by the *s* parameter of the generalized matching law implies that $s = 1$ denotes perfect substitutability (Green & Freed, 1993; Green & Rachlin, 1991; Rachlin et al., 1981). Even though the identification of the sensitivity measure, *s*, with substitutability is not universally accepted, there is agreement even among its critics that *s* represents qualitatively different reinforcers (Baum & Nevin, 1981). Moreover, even though psychologists have, as a rule, studied nonsubstitutes and economists substitutes, both disciplines tend to underpin the expectation that sensitivity reflects substitutability. Hence, Kagel et al. (1995), as behavioral economists, argued that antimatching is firmly established within economic reasoning: The antimatching relationship between food and water is the result of the relatively inelastic demand shown by these goods, which derives in turn from their relatively low substitutability. They pointed out, however, that the psychologist Herrnstein himself "recognized that matching would not be expected with qualitatively different reinforcers or with qualitatively different response requirements" (Kagel et al., 1995, p. 59, citing Herrnstein, 1970, p. 247).

The Need for Research

There are a number of reasons for extending work on the matching law into the realm of consumer choice in natural settings. First and foremost is the need to determine whether the relationships found to exist in the laboratory between matching (and deviations from matching such as undermatching, overmatching, and antimatching) generalize to the context of everyday economic behavior of human consumers and, if so, whether the variety of such relationships is associated with predictable

patterns of brand and product substitutability, complementarity, and independence. This consideration stems from intellectual curiosity with regard to the robustness of the matching concept in extralaboratory behaviors. It is necessary to directly test the matching relationship in broader contexts rather than to use intuited correspondences and economic data, as economists and behavior analysts have frequently done. A second, related consideration is how far the findings of earlier research on matching in the context of consumer research in affluent, marketing-oriented economies can be replicated and extended. The initial research on this topic (Foxall & James, 2001, 2003) sought to establish whether matching phenomena could be identified in the case of consumers' shopping behavior and involved the analysis of data for three product categories and three consumers. This work incorporated both qualitative and quantitative phases. The results not only indicated that consumer choices at the product and brand level conformed to matching considerations but also allowed discriminations to be made between the products examined in terms of the degree of substitutability suggested by the qualitative research and the pattern of behavior/reinforcer relationships established by the matching analyses. Commodities considered to be near-perfect substitutes revealed s values that closely approximated unity on the logarithmic expression of the matching curve; by contrast, products judged to be independents on the basis of the qualitative research produced evidence of clear deviations from ideal matching, including antimatching. Similar analyses were subsequently undertaken on panel data from a larger sample of 80 consumers who purchased nine food product categories over 16 weeks (Foxall, Oliveira-Castro, & Schrezenmaier, 2004; Foxall & Schrezenmaier, 2003). Ideal matching ($s \approx 1$) was found as expected for brands within a product class that had been predicted to be near-perfect substitutes (Ehrenberg, 1988). An analysis of this same data set at the product category level by Romero, Foxall, Schrezenmaier, Oliveira-Castro, and James (2006) was based on the prediction that the substitutable products would show matching, while the independent and complementary products would show undermatching or, in the case of some complementary product combinations, antimatching. The research failed to find these systematic variations at the individual level. However, at an aggregated level (minimizing the effect that individual perceptions would have on the behavior) the results approached the expected patterns, and these effects were found exclusively when data were integrated on a weekly basis. The research reported here, based on panel data for more than 1,500 consumers, is similarly concerned with patterns of matching and matching-related phenomena for products independently judged to be substitutes, complements, or independents.

   Our conception of substitutability is behavioral, defined in terms of the actual effects of brand and product attributes, namely, their degree of functional interchangeability, ranging from practical substitutability to autonomy. Our definition is operationally related to both the matching law and behavioral economics insofar as we argue that the exponent s is a measure of the homogeneity of the competing reinforcers (Green & Freed, 2003; Kagel et al., 1995; Rachlin et al., 1981) and that economists' description of substitutability in terms of price elasticity of demand is an empirical

prediction (see Appendix 1). Green and Freed, Kagel et al., and Rachlin et al. also take this view; although they note the economists' definition, they are more concerned with the substitutability of reinforcers as related in the matching law. The results of matching make sense in terms of consumers' perceptions of substitutability based on their experience of shopping and using the goods concerned. This is an alternative index of substitutability to that of the economist, neither better nor worse but of particular interest to the applied behavioral scientist.

Finally, although studies that have applied matching analyses to data for consumers' choices of brands and products in everyday shopping situations have substantiated much of the speculative thinking that links laboratory-based matching studies and analyses of marketing phenomena (Foxall, 1999), they have also raised intriguing possibilities for further investigation. Studies to date follow a progression from the simplest demonstration of the relevance of behavioral economics to consumer research to much more complex questions with respect to the operation of consumer markets. How do the new findings elucidate theoretical and methodological issues that have emerged from previous research? The larger sample employed in the current research not only permits a more stringent test of these relationships, it also allows further investigation of ways in which matching analyses of consumer choice differ from matching analyses in the operant laboratory. One such consideration is the variation in the schedules predominantly employed in matching experiments compared with those generally encountered outside the laboratory. Classic studies of matching, as typified by Herrnstein's initial reports (Herrnstein, 1961, 1970), employ concurrent variable interval-variable interval (VI-VI) or variable interval-variable ratio (VI-VR) schedules, although as Herrnstein noted on several occasions, there are good evolutionary reasons why in natural settings concurrent VR-VR schedules are most frequent. Schedules encountered in situations of purchase and consumption typically resemble concurrent VR-VR arrangements, and that is how behavior analytic work on consumer choice has regarded price/quantity bought relationships. This has implications for the generalization of matching research beyond the laboratory, which will be further discussed in light of our results in the Discussion section.

Research has revealed how human consumers' matching behavior differs from behavior that is predictable from studies of nonhumans' patterns of choice (Foxall, Oliveira-Castro, James, & Schrezenmaier, 2007). The usual expectation of matching behavior on ratio schedules is that of exclusive choice of the more generous schedule (Davison & McCarthy, 1988), a result that is compatible with both maximization and melioration accounts of choice. For several reasons, this is not the pattern usually encountered in studies of the choices made by human consumers. Most consumers practice multibrand purchasing as a matter of course; exclusive choice of a single brand over all shopping occasions certainly occurs but is comparatively rare. Unlike nonhumans, moreover, consumers are capable of making more than one brand selection even on a single shopping trip, purchasing two or more competing brands at different prices and presumably intended for alternative consumption contexts. Further analysis of consumer choice indicates that consumers maximize as well as match, but they do this in a manner dissimilar to the matching and maximization behaviors of nonhumans. Human consumers generally make

their multibrand selections from a small subset of the totality of brands available in a product category. Their consideration sets consist of tried and trusted brands, which earlier analyses have shown reflect not only the functional reinforcement provided by these brands but also the symbolic reinforcements made available by the branding efforts of marketing companies. Consideration sets of brands thus embody both utilitarian benefits—in the case of a car, for instance, the ability to get from point A to point B—and informational benefits, such as the prestige and other forms of performance feedback provided by a Porsche or other prestigious automobile. Functional and symbolic reinforcers have proved to be amenable to identification and measurement and to exert separable effects on buying behavior (Foxall et al., 2004; Foxall, James, Chang, & Oliveira, in press; Oliveira-Castro, Foxall, & James, in press). Consumers maximize *within* the consideration sets that define the brand subset from which they choose. That is, all of the brands they consider may be premium-priced versions of the product; whichever brand is purchased on any specific shopping occasion will be purchased at a higher price than those at which functionally alternative nonpremium brands are marketed. This theme is also considered in the Discussion section in light of the present results.

Method

Sample

Participants were drawn from the ACNielsen Homescan Panel, which is composed of 15,000 randomly selected UK households representative of the UK population. Panel participants scan the barcodes printed on the packages of their purchases into a sophisticated handheld barcode reader after each shopping occasion. The information recorded for each such occasion includes brand, price paid, quantity bought (package size), number of units bought (number of packages bought), date of purchase, and name of store from which the product was purchased. Data for four fast-moving consumer product categories purchased during a period of a year (July 17, 2004, to July 15, 2005) were analyzed: fruit juice, baked beans, biscuits (cookies), and yellow fats. Yellow fats were further broken down into butter, margarine, low-fat margarine, and blended spreads. Biscuits were further broken down into sweet and savory biscuits (although the sweet biscuits subcategory was not used in this research). Table 1 shows the number of consumers who purchased each product category, the total number of purchases they made, and the average number of purchases per consumer. Only consumers who made at least three purchases in the product category were included in the analyses.

Procedure

The product combinations employed in the research were decided by the researchers, and their level of substitutability, independence, and complementarity was determined by a group of 11 consumers (staff and graduate students of a business school) who were selected based on their demographic similarity to the panel members who made up the main sample. It was not feasible to use the main sample as a source of experts because its members were anonymous and not available to contact. Members

of the group of experts were chosen based on their considerable and lifelong experience of the brands in question. Substitutability from the viewpoint of consumers was assessed by this group of experts using a variant of the Exeter Substitutability Scale (Romero et al., 2006), which indicated respondents' perceptions of whether a given product/brand was a substitute or complement of another or independent of it (see Appendix 2).

Table 1
Number of Consumers, Total Purchases, and Average Purchases Per Consumer

Product category	Number of consumers	Total number of purchases made	Average number of purchases per consumer*
Baked beans	1,639	16,203	10
Yellow fats	1,817	32,468	18
Fruit juice	1,542	23,339	15
Biscuits	1,594	75,847	48

*All figures are rounded to the nearest whole number.

The four overarching product categories studied had been selected because they and their subdivisions permitted a range of product combinations to be constructed that differed in terms of the substitutability of their components (as ranked by the expert panel). Based on these four product categories, six combinations of products were analyzed: fruit juice/yellow fats, fruit juice/baked beans, biscuits/baked beans, fruit juice/biscuits, yellow fats/baked beans, and biscuits/yellow fats. Using the further subcategorizations from the yellow fats category, six more combinations were analyzed: butter/margarine, butter/low-fat margarine, blended spreads/butter, margarine/low-fat margarine, blended spread/margarine, and blended spread/low-fat margarine. And, finally, the biscuits category could also be broken down and allowed the inclusion of the combination yellow fats/savory biscuits. This resulted in 13 product combinations being analyzed. Six of the combinations were ranked as containing substitutes by the judges, 6 as independents, and 1 as complementary. No combination was perceived as being close to the extreme complementary point of the scale. In interpreting this range of combinations, it is important to recall that the purpose of the scaling was to establish consumers' perceptions of the degree of substitutability/complementarity/independence to which our product combinations were related. This, like our investigation in general, was meant to capture the nature of consumer choice in natural settings. It was not meant, as might be the case in an experimental study, to arrange product combinations that covered the entire spectrum of the scale.

Measures and Analysis

The main technique employed was matching analysis. Within consumer research, the matching law states that the proportion of pounds and pence (dollars and cents, etc.) spent for a commodity will match the proportion of reinforcers earned (i.e., purchases made as a result of that spending; Foxall, 1999). To operationalize matching in this case two ratios were used: the response ratio and the reinforcement ratio. These ratios had been developed

for previous research and are more fully explained in earlier papers (Foxall & James, 2001, 2003; James & Foxall, 2006). The response ratio was defined as the amount spent for a product category to the amount spent for a second category:

$$\frac{Amount\ paid\ for\ product\ category\ A}{Amount\ paid\ for\ product\ category\ B} \qquad \text{(Ratio 1)}$$

Likewise, the reinforcement ratio was calculated in terms of the physical quantity bought:

$$\frac{Amount\ bought\ of\ product\ category\ A}{Amount\ bought\ of\ product\ category\ B} \qquad \text{(Ratio 2)}$$

For each combination, the product category A is the product that had been bought most over the period to which the data refer, and this is the first product named in each pairing. Amount bought was determined for liquids by the number of milliliters bought and for solids by the number of grams bought, which were then translated into units purchased, which generally followed the standard size of a purchase. For example, fruit juice was measured in 1-liter units. The product category A was determined specifically by the volume/weight purchased over the period of study. The s parameter of the generalized equation was expected to vary according to the level of substitutability of products. Following Baum's (1974, 1979) suggestions, s values between 0.90 and 1.10 were considered indicative of near-perfect matching, s values over 1.10 of overmatching, and s values between 0 and 0.90 of undermatching. In accordance with Kagel et al. (1995), in cases where the predictions of the matching law were expected to be reversed, antimatching ($s < 0$) was expected.

As in previous research, the *analogy* of ratio schedules was used, where a specified number of responses have to be performed before reinforcement is delivered. For example, a consumer may have to pay 33p (i.e., 33 responses) for a can of baked beans to be made available (i.e., reinforcement). In the experimental analysis of behavior, fixed ratio (FR) schedules keep the number of required responses equal from one reinforcement to the next, whereas variable ratio (VR) schedules allow the required number of responses to change from one reinforcement to the next. Therefore, FR schedules can be analogically compared to unchanging prices over a week and VR schedules to prices changing over a number of weeks. Thus, as in earlier research (e.g., Foxall & James, 2001; Foxall & Schrezenmaier, 2003), three analyses were performed for each product combination, but it is important to make clear that these were not based on the precisely formulated FR and VR schedules of the operant laboratory. Rather, they arose from the attempt to interpret complex human behavior in natural settings according to principles gained from the tighter context of the experimental study. The first analysis resembled an FR analysis, implying not that there was a single schedule in operation for the entire length of the study (which could not be the case as prices changed constantly) but that consumers were presented with a series of concurrent FR schedules, usually one per week. In other words, FR analysis in this context means that data were integrated on the basis of a series of weekly contingencies. This is the closest of our analyses to those employed in experimental behavior analysis, but we shall refer to it as a "weekly" (or "wkly") analysis in order to emphasize that it

emerges from *consumer* behavior analysis in the context of natural settings rather than from laboratory research. Data were also averaged over 3- and 5-week periods, using data from three concurrent FR/weekly schedules or five concurrent FR/weekly schedules in order to produce an analogy of VR schedules in the laboratory. It is important to emphasize, however, that consumers' decision making most likely takes place on the basis of price information integrated by the individual consumer on a weekly basis (Foxall & James, 2003), and that our 3- and 5-week averaging of information is a matter of our seeking patterns in the data rather than consumers' experience of quasi-VR3 and quasi-VR5 schedules. Where data were analyzed by averaging across 3-week periods, we refer to the analysis as giving a "3-wk-avg" and where values were calculated across 5-week periods as providing a "5-wk-avg."

Panel participants who had bought the two products categories in question over the 52 weeks were selected for the analysis of each product combination. Consumers who bought both products within the same week (for the weekly analysis) and within periods of 3 weeks (for the 3-wk-avg analysis) and 5 weeks (for the 5-wk-avg analysis) on at least three different occasions were identified using the software package Access. The numbers of consumers fulfilling these criteria for each combination and schedule are included in Table 2.

Table 2
Numbers of Consumers for Each Product Combination for Each Schedule

Product combination	Wkly	3-wk-avg	5-wk-avg
Fruit juice/biscuits	923	986	978
Fruit juice/yellow fats	787	934	952
Fruit juice/baked beans	473	640	669
Biscuits/yellow fats	1,331	1,301	1,431
Yellow fats/baked beans	831	1,020	1,038
Biscuits/baked beans	939	1,057	1,064
Butter/margarine	105	161	180
Butter/low-fat margarine	54	88	104
Blended spread/butter	52	193	218
Blended spread/margarine	51	102	130
Blended spread/low-fat margarine	27	62	81
Margarine/low-fat margarine	26	67	75
Yellow fats/savory biscuits	744	920	963

The number of consumers buying the two products within a particular combination generally increased for the 3- and 5-wk-avg analyses and especially for the 5-wk-avg analysis, although this was not always the case. This pattern was expected because the probability of buying the two

products over a period of 5 weeks is clearly higher than the probability of buying them within 3 weeks, which in turn is higher than the probability of buying them within the same week. The combination that produced the largest subset of consumers was yellow fats/biscuits, and the one that yielded the smallest subset was margarine/low-fat margarine. The reasons for differing volumes of consumers for each combination may vary. For some product combinations, a low volume could be due to the fact that very few consumers buy both product categories, which could be the case for margarine/low-fat margarine. Analyses at both the individual level (each consumer is treated individually) and at an aggregated level (across all consumers) were performed to replicate earlier work by Romero et al. (2006). The techniques employed in the analyses used here, including those involved in calculating data points, have been explicated in greater detail in a number of investigations published elsewhere (Foxall & James, 2001, 2003; James & Foxall, 2006; Oliveira-Castro, Foxall, & James, 2008; Oliveira-Castro, Foxall, James, Pohl, et al., 2008).

Results

Substitutability Scale

The substitutability expert panel suggested that the combination of fruit juice/yellow fats was expected to be in the middle of the continuum and the combinations of subcategories of yellow fats (such as butter/margarine) were expected to be perceived largely as substitutable products.

Six combinations were placed between 1 and 2 on the scale of substitutability, which means that they are perceived as substitutes (blended spread/low-fat margarine, butter/margarine, butter/low-fat margarine, margarine/low-fat margarine, margarine/blended spread, and butter/blended spread). Six other combinations received scores between 3.80 and 4.80 (fruit juice/biscuits, yellow fats/baked beans, biscuits/yellow fats, fruit juice/baked beans, fruit juice/yellow fats, and biscuits/baked beans); hence, these combinations are perceived as independents. Finally, the combination of yellow fats/savory biscuits obtained a mean of 5.45. As a consequence, this combination is perceived as being nearer the complementary end of the scale. Descriptive statistics of the results from the substitutability scale are included as Appendix 3. No product combination was placed at the extreme complementary pole.

Matching: The Individual Level of Analysis

Our first level of analysis was for each individual in the sample; we chose consumer number 8251060 to illustrate the results for an individual panel member. We chose this purchaser because she had purchased many of the products and product combinations. Illustrative combinations were chosen that varied in their level of substitutability. Figures 1, 2, and 3 show the results for consumer number 8251060 for all three schedules and for the combinations of substitutes (margarine/low-fat margarine; see Figure 1), independents (biscuits/baked beans; see Figure 2), and complements (yellow fats/savory biscuits; see Figure 3).
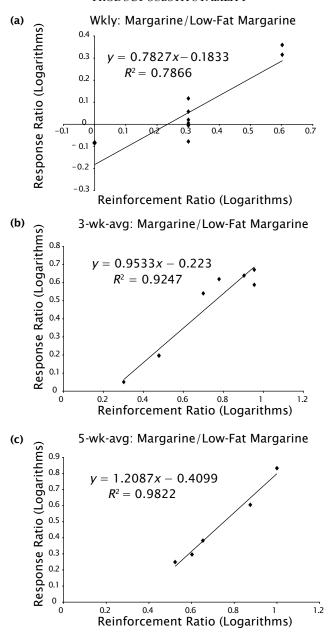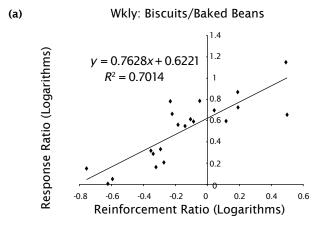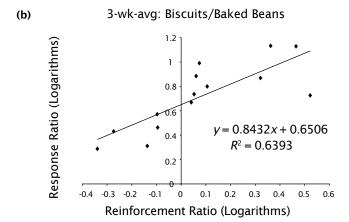
(a)

Wkly: Margarine/Low-Fat Margarine

$y = 0.7827x - 0.1833$

$R^2 = 0.7866$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

(b)

3-wk-avg: Margarine/Low-Fat Margarine

$y = 0.9533x - 0.223$

$R^2 = 0.9247$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

(c)

5-wk-avg: Margarine/Low-Fat Margarine

$y = 1.2087x - 0.4099$

$R^2 = 0.9822$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

Figure 1. Matching analysis for Consumer 8251060: substitutable products. Data for margarine/ low-fat spreads illustrate the behavior of substitute products. The response ratio is the amount of money spent on the product that was more frequently purchased over the 52-week period (measured in British pence) as a proportion of the amount spent on the alternative product similarly measured. The reinforcer ratio is the proportion of the physical amount of the product bought most frequently as a proportion of the amount purchased of the alternative product. Data are shown for each shopping occasion (usually weekly; panel a) and for spending on the two products averaged over 3-week periods (panel b) and 5-week periods (panel c).

(a)

Wkly: Biscuits/Baked Beans

$$y = 0.7628x + 0.6221$$
$$R^2 = 0.7014$$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

(b)

3-wk-avg: Biscuits/Baked Beans

$$y = 0.8432x + 0.6506$$
$$R^2 = 0.6393$$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

(c)

5-wk-avg: Biscuits/Baked Beans

$$y = 0.8664x + 0.6871$$
$$R^2 = 0.7579$$

Response Ratio (Logarithms)
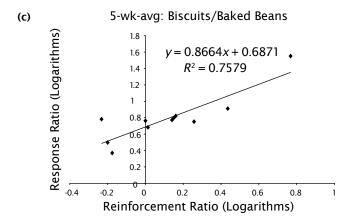
Reinforcement Ratio (Logarithms)

Figure 2. Matching analysis for Consumer 8251060: independent products. Biscuits (cookies) and baked beans illustrate independent products. Data analysis and presentational style are as described for Figure 1.
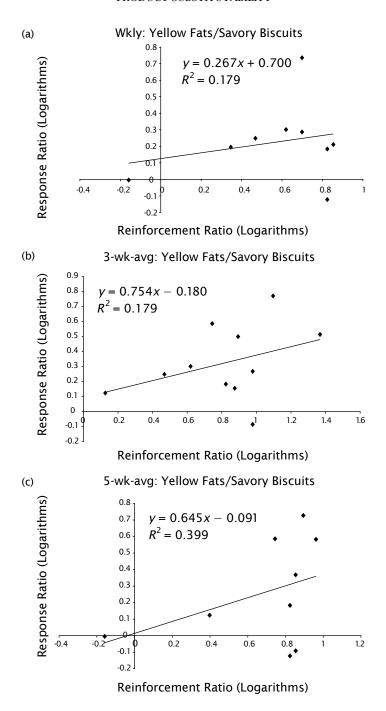
(a)

Wkly: Yellow Fats/Savory Biscuits

$y = 0.267x + 0.700$
$R^2 = 0.179$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

(b)

3-wk-avg: Yellow Fats/Savory Biscuits

$y = 0.754x - 0.180$
$R^2 = 0.179$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

(c)

5-wk-avg: Yellow Fats/Savory Biscuits

$y = 0.645x - 0.091$
$R^2 = 0.399$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

Figure 3. Matching analysis for Consumer 8251060: complementary products. Yellow fats and savory biscuits illustrate the nature of complementary products. Data analysis and presentational style are as described for Figure 1.

The choice patterns of this particular consumer are characterized largely by undermatching. For both the independent and complementary product combinations, all of the schedules show an undermatching pattern. For the substitutable products the weekly schedule shows undermatching, the 3-wk-avg analyses show matching, and the 5-wk-avg analyses show overmatching. This is an expected pattern, taking into account the predominance of matching described earlier for individual consumers.

For all three product combinations, the slopes increased between the weekly and the 3- and 5-wk-avg analyses. This is especially true for the substitutable product combinations, where the slope increased from 0.783 for the weekly analysis to 0.953 for the 3-wk-avg analysis and finally to 1.209 for the 5-wk-avg analysis. That is, the pattern of choices observed changed from undermatching to matching to overmatching as the schedule increased. The slopes were highest for the substitutable product combinations (wkly: 0.783; 3-wk-avg: 0.953; and 5-wk-avg: 1.209) compared to both the independent product combinations (wkly: 0.763; 3-wk-avg: 0.843; and 5-wk-avg: 0.866) and the complementary product combinations (wkly: 0.267; 3-wk-avg: 0.754; and 5-wk-avg: 0.645). These results met expectations regarding the degree of substitutability and the *s* parameter.

Matching: Combined Level of Analysis

The second level of analysis was for the entire sample. At the combined level of analysis, the results for each individual were summed so that the percentages of respondents whose behavior is characterized by matching, undermatching, overmatching, and antimatching can be assessed for each of the temporally based analyses.

Weekly analysis. The results for the combined matching analysis performed using weekly data are shown in Table 3. The table shows the percentage of consumers whose values of *s*, in Equation 3, were in the range of antimatching, undermatching, and overmatching for each combination or product pair. Product combinations are lined up according to their substitutability, with consumers' perceived substitute combinations on the left, those perceived as independent in the middle, and those perceived as complements to the right. As can be seen from Table 3, the predominant choice pattern observed was undermatching. For 12 of the 13 combinations, undermatching was the most frequent choice behavior found. The single exception was for the butter/margarine combination, where the percentages of undermatching and matching were identical. The highest percentage of undermatching was found for a combination of independent products (biscuits/yellow fats), with 68% undermatching. The lowest percentage of undermatching was observed for the butter/margarine combination of substitutable products (41%). The second most frequent pattern of behavior was near-perfect matching. The highest percentage of matching was for the butter/margarine (41%) and margarine/low-fat margarine (35%), which are both combinations of substitutable products. The smallest percentage of matching was found for the combination of yellow fats/savory biscuits (15%), which is the only combination of complementary products, and it is the single exception where matching was the third and not the second most common form of choice allocation. When the percentages were added, undermatching and perfect (or near-perfect) matching accounted for between

74% (fruit juice/baked beans) and 86% (blended spread/low-fat margarine) of the consumers' choices. Overmatching was the third most frequent pattern, except for the combination of complementary products (yellow fats/savory biscuits). Antimatching was infrequent (between 0% and 8%) and was the pattern least found for all combinations.

Table 3
Patterns of Matching Analysis

	Weekly				3-wk-avg				5-wk-avg			
	Percentage of consumers											
	Antimatching	Undermatching	Matching	Overmatching	Antimatching	Undermatching	Matching	Overmatching	Antimatching	Undermatching	Matching	Overmatching
Blended spread/butter	0	52	27	21	8	37	38	21	4	34	41	21
Blended spread/margarine	8	59	18	16	4	41	33	21	7	34	41	18
Margarine/low-fat margarine	0	50	35	15	0	31	48	21	5	41	33	21
Butter/low-fat margarine	0	52	28	20	1	37	43	19	3	31	40	26
Butter/margarine	1	41	41	17	2	33	42	23	3	31	43	23
Blended spread/low-fat margarine	0	56	30	14	3	49	24	24	4	42	32	22
Biscuits/baked beans	4	61	19	16	3	52	25	20	2	46	27	25
Fruit juice/yellow fats	3	54	24	19	2	48	31	19	3	46	29	22
Fruit juice/baked beans	5	51	23	21	3	44	29	24	3	43	28	26
Biscuits/yellow fats	4	68	16	12	3	57	24	16	1	52	26	21
Yellow fats/baked beans	3	59	19	19	2	51	29	18	1	47	31	21
Fruit juice/biscuits	4	60	19	17	3	49	26	22	2	47	25	26
Yellow fats/savory biscuits	7	60	15	18	4	55	20	21	5	50	22	23

3-wk-avg analysis. Although the percentage of undermatching decreased when compared to the weekly schedule, for 9 of the 13 combinations in the 3-wk-avg schedule, undermatching was the most common form of behavioral allocation (see Table 3). The percentages of undermatching varied from 31% (margarine/low-fat margarine) to 57% (biscuits/yellow fats), and overall the percentage was slightly less for substitute combinations. For 4 combinations of products at the substitutable end of the continuum (blended spread/butter, margarine/low-fat margarine, butter/low-fat margarine, and butter/margarine), the most observed pattern was matching. However, for 8 remaining combinations, matching was the second most common pattern observed, ranging from 20% (yellow fats/savory biscuits) to 33% (blended spread/margarine). Generally, the more substitutable the products were, the higher the level of matching was. For 11 combinations, overmatching was the third most common pattern of allocation. Overmatching was generally higher for the 3-wk-avg schedule

than for the weekly schedule, except for two combinations where it was the same percentage (blended spread/butter and fruit juice/yellow fats) and two other combinations where the percentage of overmatching was smaller (yellow fats/baked beans and butter/low-fat margarine). For each combination, the pattern that was found least often was antimatching, as was the case for the weekly schedule. The highest percentage of antimatching was found for blended spread/butter.

5-wk-avg analysis. Results for the 5-wk-avg analysis are similar to those for the 3-wk-avg analysis (see Table 3). Again, matching is more common than for the weekly analysis, with four of the combinations having matching as the predominant pattern (blended spread/butter, blended spread/margarine, butter/low-fat margarine, and butter/margarine), all of which are at the substitutable end of the continuum. Undermatching remained the most common pattern, with nine of the product combinations showing this as the predominant pattern. Antimatching remained infrequent, ranging from 1% for the yellow fats/baked beans and biscuits/yellow fats combinations to 7% for blended spread/margarine. Overall higher levels of matching were observed in the 3-wk-avg and 5-wk-avg analyses when compared to the weekly analysis. This was the case for all product combinations.

Matching: Aggregated Level of Analysis

The aggregated analysis treated the sample as a single consumer, thereby summarizing all purchases together. As such, it is a means of portraying the patterns contained in the data that avoids the cumbersome presentation of information for each individual separately. Regression analyses were conducted with all data points from all consumers for each product combination and each temporally based analysis (wkly, 3-wk-avg, and 5-wk-avg). The aggregated analysis was performed to minimize the effect that individual differences could have on the behavior. Table 4 summarizes the results of the aggregated analysis, reporting the slope, intercept, and adjusted *R* square for each product combination for each schedule. Table 4 also shows that the values of the *s* parameter varied between 0.446 and 0.956 for the weekly analysis. The general tendency of the slope was to decrease with the combinations' level of substitutability. Highly substitutable products such as margarine/low-fat margarine exhibited the greatest slope values, showing a near-perfect matching pattern, whereas for products that were ranked as independents, the slope tended to decrease, indicating undermatching. For complementary products, the slope clearly indicated undermatching. The results for the weekly averaged data analyses showed similar findings (that is, the slope decreased as the level of substitutability reduced). As in the individual analyses, there was also the tendency for the slope to increase as the schedule increased (wkly to 3-wk-avg to 5-wk-avg), except for butter/blended spread (where the values were the same for the 3-wk-avg and 5-wk-avg analyses) and biscuits/yellow fats.

The adjusted *R* square values varied from 0.105 to 0.806. For the weekly analysis, the adjusted *R* square for 10 of the combinations was under 0.5, indicating great dispersion and therefore low adjustment of the data to the model. For the weekly averaged analyses, the adjusted *R* square values were generally higher. Indeed, in the 3-wk-avg analysis only 2 combinations had an adjusted *R* square under 0.5 (blended spreads/low-fat margarine and

Table 4
Generalized Equation: Results at an Aggregated Level

	Product combination	Weekly			3-wk-avg			5-wk-avg		
		Slope	Intercept	Adj R^2	Slope	Intercept	Adj R^2	Slope	Intercept	Adj R^2
Substitutes	Blended spread/butter	0.749	-0.114	0.331	0.905	-0.186	0.757	0.904	-0.180	0.769
	Margarine/low-fat margarine	0.956	-0.205	0.705	0.984	-0.132	0.664	1.054	-0.129	0.806
	Blended spread/ margarine	0.704	0.115	0.378	0.824	0.1001	0.524	0.919	0.083	0.599
Independents	Butter/low-fat margarine	0.724	0.129	0.314	0.782	0.162	0.579	0.842	0.179	0.698
	Butter/margarine	0.606	0.226	0.260	0.945	0.292	0.591	0.994	0.289	0.678
	Blended spread/low-fat margarine	0.474	0.079	0.105	0.902	0.053	0.469	0.947	0.036	0.649
	Biscuits/baked beans	0.719	0.559	0.508	0.791	0.600	0.631	0.826	0.618	0.683
	Fruit juice/yellow fats	0.748	-0.349	0.493	0.821	-0.383	0.633	0.861	-0.408	0.715
	Fruit juice/baked beans	0.739	0.504	0.068	0.801	0.059	0.661	0.858	0.040	0.729
	Yellow fats/baked beans	0.696	0.402	0.388	0.791	0.430	0.515	0.849	0.440	0.592
Complements	Fruit juice/biscuits	0.752	-0.509	0.549	0.833	-0.56	0.712	0.868	-0.587	0.773
	Biscuits/yellow fats	0.699	0.155	0.415	0.899	0.124	0.605	0.898	0.180	0.693
	Yellow fats/savory biscuits	0.446	-0.056	0.158	0.749	-0.169	0.410	0.861	0.213	0.543

Note. Numbers in italics indicate undermatching. Numbers that are underlined indicate near-perfect matching.

savory biscuits/yellow fats), and in the 5-wk-avg analysis all the adjusted *R* square values were greater than 0.5, denoting a reasonable adjustment to the model. Hence, the results of the 3-wk-avg and 5-wk-avg data analyses were more adjusted to the model than were those from the weekly analysis. All the intercept values differed markedly from zero and each other, indicating that some unknown bias caused some degree of asymmetry between the options. Further analyses were performed to explore the relationship between slope values, the intercept, and substitutability. The full results of these correlation analyses are included as Table 5.
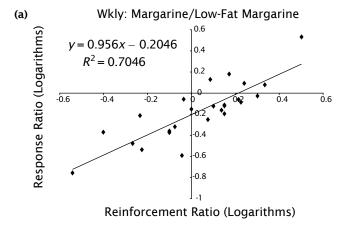
There was a significant relationship between substitutability and the slope of the aggregated weekly analysis, $r = -.232$, $p < .05$, substitutability and the slope of the aggregated 3-wk-avg analysis, $r = -.580$, $p < .05$, and substitutability and the slope of the aggregated 5-wk-avg analysis, $r = -.612$, $p < .05$. These results further confirm the observed tendency of increased slopes with increases in substitutability, supporting the connection between substitutability and the *s* parameter.
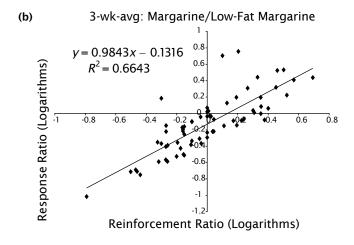
Table 5
Substitutability, Slopes, and Intercept Correlations

	Substitutability	Significance
Slope: wkly	−.232*	.380
Slope: 3-wk-avg	−.580*	.038
Slope: 5-wk-avg	−.612*	.026
Intercept: wkly	.061	.842
Intercept: 3-wk-avg	−.098	.751
Intercept: 5-wk-avg	.052	.865

* $p < .05$.

Illustrative example. Figures 4, and 5, and 6 contain graphically represented results for three selected product combinations illustrating substitute products (margarine/low-fat margarine), independent products (fruit juice/yellow fats), and complementary products (yellow fats/ savory biscuits), respectively. As with the combined analysis, these illustrative examples supported the expected pattern that as products become less substitutable the slope (the *s* parameter) reduces. As can be seen in Figures 4, 5, and 6, on the basis of the weekly analysis, the slope for the combination of substitutable products (margarine/low-fat margarine) showed near-perfect matching, for the independent products (fruit juice/yellow fats) the slope demonstrated undermatching, and for the complementary products combination (yellow fats/savory biscuits) the slope showed undermatching at a greater extent than the independent products. In the case of the weekly averaged data analyses, the differences among the substitutability continuum were also evidenced. Indeed, near-perfect matching was found for the substitutable combination and undermatching was found for the independent and complementary combinations. However, for each combination, the slopes from the 3-wk-avg analysis were greater than the slopes from the weekly analysis, and the slopes from the 5-wk-avg analysis were greater than those from the 3-wk-avg schedule.
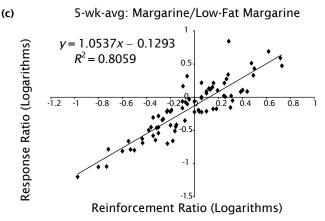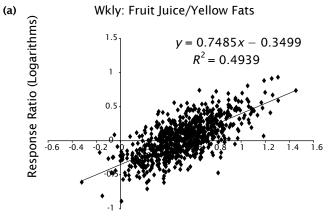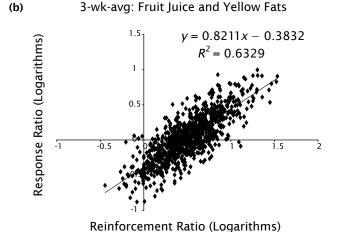
(a)

Wkly: Margarine/Low-Fat Margarine

$y = 0.956x - 0.2046$
$R^2 = 0.7046$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

(b)

3-wk-avg: Margarine/Low-Fat Margarine

$y = 0.9843x - 0.1316$
$R^2 = 0.6643$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

(c)

5-wk-avg: Margarine/Low-Fat Margarine

$y = 1.0537x - 0.1293$
$R^2 = 0.8059$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

Figure 4. Aggregated matching analysis for substitutes. Yellow fats and savory biscuits illustrate the nature of complementary products. Data analysis and presentational style are as described for Figure 1.

(a)

Wkly: Fruit Juice/Yellow Fats

$$y = 0.7485x - 0.3499$$
$$R^2 = 0.4939$$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

(b)

3-wk-avg: Fruit Juice and Yellow Fats

$$y = 0.8211x - 0.3832$$
$$R^2 = 0.6329$$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

(c)

5-wk-avg: Fruit Juice and Yellow Fats

$$y = 0.861x - 0.4075$$
$$R^2 = 0.715$$

Response Ratio (Logarithms)
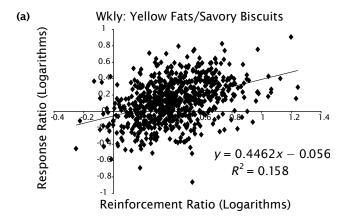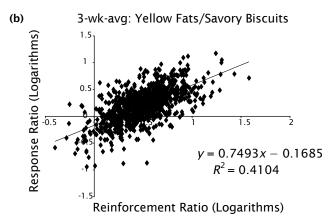
Reinforcement Ratio (Logarithms)
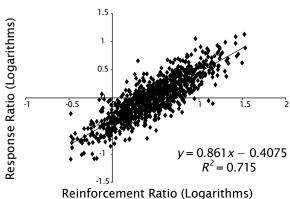
Figure 5. Aggregated matching analysis for independents. Biscuits (cookies) and baked beans illustrate independent products. Data analysis and presentational style are as described for Figure 1.
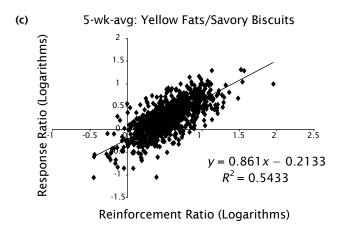
(a)

Wkly: Yellow Fats/Savory Biscuits

$y = 0.4462x - 0.056$
$R^2 = 0.158$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

(b)

3-wk-avg: Yellow Fats/Savory Biscuits

$y = 0.7493x - 0.1685$
$R^2 = 0.4104$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

(c)

5-wk-avg: Yellow Fats/Savory Biscuits

$y = 0.861x - 0.2133$
$R^2 = 0.5433$

Response Ratio (Logarithms)

Reinforcement Ratio (Logarithms)

Figure 6. Aggregated matching analysis for complements. Biscuits (cookies) and baked beans illustrate independent products. Yellow fats and savory biscuits illustrate the nature of complementary products. Data analysis and presentational style are as described for Figure 1.

Discussion

The current research supports many aspects of the earlier studies (Foxall & James, 2001, 2003; Foxall, Oliveira-Castro, & Schrezenmaier, 2004; Foxall & Schrezenmaier, 2003; Romero et al., 2006) and extends knowledge in a number of ways. Both this study and the Romero et al. (2006) study show consistency with the theoretically expected patterns for the aggregated level of analysis. The *s* parameter acted as a measure of substitutability, decreasing as the level of substitutability within each product combination decreased. The *s* of substitutable combinations was greater than the *s* of independent combinations, which in turn was greater than the *s* of complementary combinations. While the Romero et al. study suggested that this was the pattern for the weekly schedule, and the results of that study did not support the pattern for the weekly averaged data analyses, the current study presents further support for the pattern on all schedules, although the pattern was strongest for the weekly analysis of the data.

The most common pattern observed for independent products was consistent undermatching, whereas for substitutable combinations it was generally near-perfect matching for the weekly averaged data analyses. At the level of individual analysis, the results of this study showed some similarity with those generated by the Romero et al. (2006) study. As in that study, the frequency of undermatching (especially when analyses were conducted on a weekly basis) was particularly high. However, the degree of overmatching was lower in this study than in the earlier one: In the study by Romero et al., overmatching ranged between 13 and 70% for the various combinations, while in this study overmatching was found in 14 to 24% of the trials. The percentage of overmatching was especially different in the results from the weekly analysis. Overmatching in the earlier study was the second most common behavior allocation but was relegated to third place, behind matching and undermatching, in this study. This may simply be a function of differences in sample size. In common with the earlier research for those product categories perceived as substitutes, we found that there was less matching (or near-perfect matching) than expected. However, contrary to the earlier study, higher percentages of matching were evinced in the products perceived as more substitutable. This was more so for the 3-wk-avg and 5-wk-avg analyses, although it is not entirely absent from the results of the weekly analysis. The most common pattern to emerge from the 3-wk-avg schedules was matching, which was apparent for four of the six combinations that were perceived as substitutes (butter/margarine, butter/low-fat margarine, margarine/low-fat margarine, and blended spread/butter). In the case of the 5-wk-avg analysis, again four of the six more substitutable product combinations showed matching as the most common form of relationship (blended spread/butter, blended spread/margarine, butter/low-fat margarine, and butter/margarine). Hence, in the case of the weekly averaged data analyses, the most common pattern found for substitutable combinations was matching. For the combinations perceived as independent, undermatching was the most common pattern whatever the schedule. Thus, at an individual level for the weekly averaged data analyses, different degrees of substitutability produced different percentages of matching patterns, unlike in the earlier study. However, some consumers did show, for example, undermatching for substitutable products and matching

for independent commodities. All forms of behavioral allocation were observed for almost all the product combinations (some showed very little or no antimatching) and across substitutability and schedule. The variety of patterns found might be explained by the fact that different consumers could have different perceptions about the degree of the combinations' substitutability. In addition, the consumers who were asked to allocate the combinations along the substitutability continuum might have different perceptions than the shoppers who took part in the analysis.

According to Kagel et al. (1995), when participants are faced with complementary reinforcers, their behavior will exhibit antimatching, notably in the case of gross complements. However, in the case of complementary products (as defined by the group of experts), the percentage of antimatching was very low in comparison to the other forms of behavioral allocation, and our study did not confirm the antimatching effect. Indeed, undermatching was found to be the predominant pattern for even the most complementary combination (yellow fats/savory biscuits). However, this was the only combination perceived as complementary products, and this combination obtained only an average of 5.45 on the substitutability scale (see Appendix 3), which means that the combination was not perceived as complementary by all the consumers or was perceived as no more than moderately complementary by some. Hence, it is difficult to generalize this result. In order to be able to generalize results for complementary products, products that are truly complementary need to be identified (e.g., cereals and milk) and used for future research. Finally, in line with the findings of Foxall and James (2003) and Romero-Ordonez (2005), results from the current research show that parameter b (bias) in both weekly and weekly averaged analyses differed significantly from unity, evidencing possible biases, for example, in the form of availability of products, and additional, nonprogrammed response costs associated with each product.

The research we have reported contributes also to the questions of what and how consumers may be said to maximize. The issue of maximization is, as Rachlin (1995) argued, similar to that of matching: These terms relate to techniques for the analysis of behavior rather than to potentially falsifiable empirical facts. However, the kinds of panel data we have employed are integrated over time. Consumers make purchase decisions on each shopping occasion because that is when the comparative prices of the brands within each buyer's consideration set can be brought to bear on a purchase response. It is only at this point that the purchase of one or more brands, at one or more prices, can become part of the sequence of consumer choice, only at this point that the various patterns of contingencies can impinge on response outcomes. However, this does not differ fundamentally from the situation of the human or nonhuman participant in an experiment that is based on variable reinforcement schedules. The ultimate necessity of recognizing that consumer behavior must be integrated at the weekly level (or at least that of the single shopping occasion) does not preclude the integration of that behavior over longer periods if meaningful patterns of response and contingency emerge therefrom. We have argued elsewhere (e.g., Foxall & James, 2003) that consumers maximize a combination of functional or utilitarian reinforcement and symbolic or informational reinforcement, both of which were described earlier. This accounts for the tendency of many consumers to include high-priced or premium brands in their consideration

sets, either predominantly or even exclusively, and to avoid or nearly so the cheapest brand versions on the market. Consumer choice is influenced not only by the intrinsic functional benefits gained from product classes but also from the benefits provided by marketing activities whose effects inhere in the brand. The molar pattern of choice must be taken into consideration in reviewing the nature of the matching and maximization that consumer behavior can be said (by scientists making theoretical judgments) to display. This includes considerations of the structure of competition within a product market, based on the number of brands available and their classification according to the utilitarian and informational benefits they offer. Such considerations relate to the interpretation of maximization in terms of both price elasticity and the plasticity of demand (Foxall & Schrezenmaier, 2003; Foxall, Yan, James, & Oliveira-Castro, 2009). In Rachlin's (1995) terms, they enrich discussion of the ways in which transactions constitute behavioral *acts* and of how these acts fall into *patterns* of choice.

Finally, in the spirit of Rachlin's view of maximization and matching analyses as methodological perspectives rather than fixed elements of our subject matter, we draw attention to the ways in which our study differs from those normally found in behavior analysis. As in our earlier studies, we used panel data rather than experimentation and aggregate measures as well as those for single consumers. The results, nevertheless, demonstrate that the extralaboratory context of the present work yields patterns of systematic behavior related to its environmental determinants, the fundamental requirement of behavioral analysis. If there is to be an extension of the principles of behavior analysis, matching, and behavioral economics to wider spheres of human behavior and application, the methods we employed, which are not subject to the vagaries of individual histories of reinforcement and punishment to which we have no empirical access, appear to be the sole means of demonstrating that scientific analysis rather than speculative extrapolation is the way forward. Because matching usually is demonstrated in the laboratory on interval schedules, the question arises whether our studies of the allocation of behavior on apparent ratio schedules in natural settings might not be considered "matching" in the classic sense. Such criticism would have to ignore those studies that have successfully and effectively employed ratio schedules in both basic and applied research and would have to overlook Herrnstein's (1997) argument that ratio schedules are more likely to be encountered in nonlaboratory settings as a result of natural selection. The results indicate also that the range of schedules assumed here remains relevant to the analysis of consumer choice. The assumption that behavior patterns aggregated over single shopping opportunities approximate FR schedules seems reasonable given the fact that naturally occurring contingencies are unlikely ever to resemble those devised within the laboratory to the degree that they can be considered entirely homologous. Nor, as we have pointed out, is VR a perfect analogy for consumer behavior aggregated over periods of time. Reinforcement and response are not entirely independent in the case of matching on ratio schedules. Indeed, the idea of matching needs to be carefully defined in the context of buyer behavior. It is important to recognize that the matching law says nothing about consumption. All reinforcers obtained are assumed to be consumed. If matching implied simply that the proportion of buying responses for Brand A equals the proportion of reinforcers obtained

from that brand, it would be a truism. Assuming that the reinforcement value from consuming a commodity is constant and that all commodities purchased are consumed (and these seem to be reasonable assumptions), then the proportion of "purchases" would always match the proportion of "reinforcers." However, in the specific context of consumer behavior, these assumptions appear justified on the pragmatic grounds that they make possible analyses of choice in operant terms and yield results that both are predictable from the corpus of knowledge on behavioral choice and which in turn elucidate the nature of choice in general. Perhaps most important of all, they make it possible to argue that consumer behavior in natural settings differs in important respects from either human or nonhuman behavior in the operant laboratory, and this may have consequences for the general interpretation of complex choice (understood as that which is not amenable to a direct laboratory analysis).

There remain, nevertheless, imperatives for further research. To put the present article in context, this research has extended the work of Romero et al. (2006), in particular by employing a larger sample that has allowed patterns of choice to be monitored in greater detail than was the case for the earlier research, and also has produced some contradictory findings. Further research with larger samples should help to cement knowledge about the patterns of behavioral allocation in consumer choice, and it is vital that such investigation take into account a wider range of product combinations. This is especially the case with respect to complementary products. Product combinations that score consistently around the 6 and 7 points of the substitutability scale should be incorporated into future investigations. Within the aggregated analysis especially, the differences in adjusted R square require further exploration in order to ascertain whether the lower adjusted R square trend for weekly versus weekly averaged data analyses is due simply to the smaller number of consumers in the latter case or whether additional influences are involved. The study by Romero et al. highlighted the need for a qualitative dimension to research in order to explore the issue of substitutability, a technique that had proven informative in the early work of Foxall and James (2001). Do consumers themselves integrate price and product information within each week (as has been asserted) or are there dimensions of their decision making that involve integration over longer time periods? To what extent do consumers even remember products fully across weeks, and how do purchases in one week affect subsequent choices? What is the influence of interpurchase consumption on these subsequent purchase decisions? Qualitative analysis could also explore learning histories and consumers' perceived substitutability of product combinations. Answers to these questions can both supplement the interpretation of quantitative results and shape the direction and content of further investigation.

References

BAUM, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior, 22,* 231–242.

BAUM, W. M. (1979). Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior, 23,* 45–53.

BAUM, W. M., & NEVIN, J. A. (1981). Maximization theory: Some empirical problems. *Behavioral and Brain Sciences*, *4*(3), 389–390.

BICKEL, W. K., & VUCHINICH, R. E. (Eds.). (2000). *Reframing health behavior change with behavioral economics*. Mahwah, NJ: Lawrence Erlbaum Associates.

DAVISON, M., & MCCARTHY, D. (1988). *The matching law: A research review*. Hillsdale, NJ: Erlbaum.

EHRENBERG, A. S. C. (1988). *Repeat buying: Facts, theory and applications*. New York: Oxford University Press.

FOXALL, G. R. (1999). The substitutability of brands. *Managerial and Decision Economics*, *20,* 241–257.

FOXALL G. R., & JAMES, V. K. (2001). Behavior analysis of consumer brand choice: A preliminary analysis. *European Journal of Behavior Analysis*, *2,* 209–220.

FOXALL G. R., & JAMES, V. K. (2003). The behavioral ecology of brand choice: How and what do consumers maximize? *Psychology and Marketing, 20,* 811–836.

FOXALL, G. R., JAMES, V. K., CHANG, W., & OLIVEIRA-CASTRO, J. M. (in press). Substitutability and independence: Matching analyses of brands and products. *Journal of Organizational Behavior Management.*

FOXALL, G. R., OLIVEIRA-CASTRO, J. M., JAMES, V. K., & SCHREZENMAIER, T. C. (2007). *The behavioral economics of brand choice*. New York: Palgrave Macmillan.

FOXALL, G. R., OLIVEIRA-CASTRO, J. M., & SCHREZENMAIER, T. C. (2004). The behavioral economics of consumer brand choice: Patterns of reinforcement and utility maximization. *Behavioural Processes*, *66,* 235–360.

FOXALL, G. R., & SCHREZENMAIER, T. C. (2003). The behavioral economics of consumer brand choice: Establishing a methodology. *Journal of Economic Psychology*, *24,* 675–695.

FOXALL, G. R., YAN, J., JAMES, V. K., & OLIVEIRA-CASTRO, J. M. (2009). *Brand-related and situational influences on demand elasticity.* Cardiff Business School Working Paper (M2009/6), ISSN: 1753-1632.

GREEN, L., & FISHER, E. B. (2000). Economic substitutability: Some implications for health behavior. In W. K. Bickel & R. E. Vuchinich (Eds.), *Reframing health behavior change with behavioral economics* (pp. 115–144). Mahwah, NJ: Lawrence Erlbaum Associates.

GREEN, L., & FREED, D. E. (1993). The substitutability of reinforcers. *Journal of the Experimental Analysis of Behavior, 60,* 141–158.

GREEN, L., & RACHLIN, H. (1991). Economic substitutability of electrical brain stimulation, food, and water. *Journal of the Experimental Analysis of Behavior, 55,* 133–143.

HERRNSTEIN, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, *4,* 267–272.

HERRNSTEIN, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, *13,* 243–366.

HERRNSTEIN, R. J. (1997). *The matching law: Papers in psychology and economics*. Cambridge, MA: Harvard University Press.

HEYMAN, G. M. (1996). Elasticity of demand for alcohol in humans and rats. In L. Green & J. H. Kagel (Eds.), *Advances in behavioral economics* (Vol. 3, pp. 107–132). Norwood, NJ: Ablex.

HURSH, S. R. (1980). Economic concepts for the analysis of behavior. *Journal of the Experimental Analysis of Behavior, 34,* 219–238.

HURSH, S. R. (1984). Behavioral economics. *Journal of the Experimental Analysis of Behavior, 42,* 435–452.

HURSH, S. R., & BAUMAN, R. A. (1987). The behavioral analysis of demand. In L. Green & J. H. Kagel (Eds.), *Advances in behavioral economics* (Vol. 1, pp. 117–165). Norwood, NJ: Ablex.

JAMES, V. K., & FOXALL, G. R. (2006). *Matching, maximisation and consumer choice.* Cardiff Business School Working Paper (M2006/2), ISSN: 1753-1632.

KAGEL, J. H., BATTALIO, R. C., & GREEN, L. (1995). *Economic choice theory: An experimental analysis of animal behavior.* Cambridge: Cambridge University Press.

OLIVEIRA-CASTRO, J. M., FOXALL, G. R., & JAMES, V. K. (in press). Consumer brand choice: Money allocation as a function of brand reinforcing attributes. *Journal of Organizational Behavior Management.*

OLIVEIRA-CASTRO, J. M., FOXALL, G. R., & JAMES V. K. (2008). Individual differences in price responsiveness within and across food brands. *Service Industries Journal, 28,* 733–753.

OLIVEIRA-CASTRO, J. M., FOXALL, G. R., JAMES,V. K., POHL, R. H. B. F., DIAS, M. B., & CHANG, S. W. (2008). Consumer-based brand equity and brand performance. *Service Industries Journal, 28*, 445–461.

PENROSE, E. (1959). *The theory of the growth of the firm.* Oxford: Oxford University Press.

RACHLIN, H. (1995). Behavioral economics without anomalies. *Journal of the Experimental Analysis of Behavior, 64,* 397–404.

RACHLIN, H., BATTALIO, R., KAGEL, J., & GREEN, L. (1981). Maximization theory in behavioral psychology. *Behavioral and Brain Sciences, 4,* 371–417.

ROMERO, S., FOXALL, G. R., SCHREZENMAIER, T., OLIVEIRA-CASTRO, J. M., & JAMES, V. K. (2006). Deviations from matching in consumer choice. *European Journal of Behavior Analysis*, *7*(1), 15–40.

ROMERO-ORDONEZ, S. J. (2005). *Patterns of consumer choice: Product substitutability and the matching law* (MSc dissertation, University of Exeter).

SCHREZENMAIER, T. C. (2005). *A methodological exploration of the relevance of behavioral economics to the study of consumer brand choice* (Doctoral thesis, Cardiff University).

Appendix 1: Definitions of
Substitutability, Complementarity, and Independence

The source of our definitions of substitutability, complementarity, and independence is the behavior of consumers. That is, substitutes are commodities that are interchangeable in use; complements are jointly consumed; and the patterns of consumption for independents do not share common elements. Among consumer researchers, Ehrenberg (1988) strongly advocated such a behavioral portrayal of substitutability: Consumers typically confine their purchases to a small subset of the brands available within a product category, selecting among them on an apparently random basis. Ehrenberg contended that such brands are interchangeable in practice because they share identical physical formulations and are therefore functionally equivalent. Consumers switch temporarily from one to another and back again, irrespective of price variations, because they have essentially the same effects. Our examinations of multibrand choice in terms of matching show, however, that consumers are sensitive to even small fluctuations in price among the brands in their "repertoires" (Foxall et al., 2007). Similarity of physical formulation is undeniably an influence on substitutability in that it contributes to commonalities in the functional or utilitarian benefits consumers derive from the brands in a product category, as Ehrenberg suggested. In addition, however, our work shows that brands are differentiated to a greater or lesser extent by the symbolic or informational rewards created by companies' branding activities, by which consumers' status and feelings of self-esteem are enhanced. All of these considerations support our conception of substitutability, complementarity, and independence because they influence the "plasticity" as well as the elasticity of demand (Penrose, 1959).

On this basis, economists' characterizations of substitutability, complementarity, and independence can be seen as a prediction rather than as a definition of the relationships between products and brands. Behavioral economists have employed both behavioral and economic views of substitutability. In pointing out, for instance, that ideal matching ($s = 1$) is found only when identical reinforcers are offered, Kagel et al. (1995) alluded to the physical reinforcers available to nonhumans in operant matching experiments as materially identical; elsewhere, substitutability is understood in terms of relative prices and quantities demanded (e.g., Green & Fisher, 2000; Hursh, 1980, 1984; Hursh & Bauman, 1987). Given our view that substitutability, complementarity, and independence are typologies of behavior, we have used the substitutability scale as a proxy verbal measure of consumers' consumption patterns. Because $s = 1$ is characteristic of substitutability, whereas independent commodities are expected to exhibit antimatching ($s > 0$), we have assumed that undermatching ($s < 1$) and overmatching ($s > 1$) depict forms of complementarity insofar as they represent deviations from the extremes of substitutability and independence. (Hence, whereas Kagel et al. refer to commodity combinations they expect to exhibit antimatching as "gross complements," we define such commodities as "independents.") Whether these definitions of matching correspond to the economists' definitions or, as we would prefer to say, predictions, in terms of relative prices and quantities demanded remains an empirical question that is independent of our present inquiry. Our primary aim, based on our

behavioral definitions, therefore emphasized testing the nature and strength of relationships between the outcomes of matching analyses for products and the judgments of the expert group of consumers who completed the substitutability scale (see Table A1).

Table A1
Matching Criteria for the Ascription of Substitutability, Complementarity, and Independence

	Defined values of s	Criterion values of s	Relationship of compared commodities	Economists' definitions (predictions)
Matching	$s = 1$	$s = 0.9–1.1$	Substitutes	If the price of A rises, less of A is demanded but more of B.
Overmatching	$s > 1$	$0 < s < 0.9$	Complements	If the price of A falls, demand for A and B increases.
Undermatching	$0 < s < 1$	$s > 1.1$	Complements	If the price of A falls, demand for A and B increases.
Antimatching	$s < 0$	$s < 0$	Independents	Price change for A has no effect on demand for B.
References	Baum, 1974, 1979; Rachlin et al., 1981	Baum, 1974; Romero et al., 2006		Green & Fisher, 2000; Green & Freed, 1993; Green & Rachlin, 1991; Kagel et al., 1995; Hursh, 1980, 1984; Hursh & Bauman, 1987

Appendix 2: The Substitutability Scale

To what extent are the products substitutes?

> This project is part of a research program about the economic psychology of consumer choice. We are looking at combinations of products and we need to know your opinion about their substitutability. The questionnaire should take no more than 5 minutes to fill in.
>
> Thank you for your participation.

Considering **Substitutability** as the degree to which two products can serve the same purpose, please rate the degree of substitutability of the following commodities. In the following scale, **1 corresponds to complete substitutability** (for example, Coke can be replaced by Pepsi and inversely), the middle point **4 corresponds to independency** (where the two products serve two completely different purposes, for example, nails and butter), and **7 means that you see the products as complements** (one product needs the other to achieve its purpose, for example, knives and forks).

"Yellow fats" means butter, spreads, and margarine.

Product categories	Substitutes		Independents			Complements	
	1	2	3	4	5	6	7
Fruit juice and biscuits							
Fruit juice and yellow fats							
Baked beans and yellow fats							
Butter and margarine							
Fruit juice and baked beans							
Butter and blended spread							
Biscuits and baked beans							
Butter and low-fat margarine							
Yellow fats and biscuits							
Margarine and low-fat margarine							
Margarine and blended spread							
Blended spread and low-fat margarine							
Savory biscuits and yellow fats							

Appendix 3: Descriptive Statistics of the Substitutability Scale

Product combination	N	Minimum	Maximum	Mean	Standard deviation
Savory biscuits/yellow fats	11	4	7	5.45	1.37
Fruit juice/biscuits	11	4	6	4.73	0.79
Baked beans/yellow fats	11	4	7	4.64	1.03
Yellow fats/biscuits	11	4	7	4.55	1.04
Fruit juice/baked beans	11	4	6	4.36	0.67
Fruit juice/yellow fats	11	3	4	3.91	0.30
Biscuits/baked beans	11	2	4	3.82	0.60
Blended spread/low-fat margarine	11	1	6	1.82	1.47
Butter/margarine	11	1	7	1.73	1.79
Butter/low-fat margarine	11	1	3	1.64	0.67
Margarine/low-fat margarine	11	1	2	1.36	0.50
Margarine/blended spread	11	1	2	1.27	0.47
Butter/blended spread	11	1	2	1.27	0.47