A rational analysis of marketing strategies

Nisheeth Srivastava (nsrivast@cse.iitk.ac.in)
Department of CSE, IIT Kanpur
Kalyanpur, UP 208016 India

Edward Vul (evul@ucsd.edu)
Department of Psychology, 9500 Gilman Drive
La Jolla, CA 92093 USA

Abstract
Rational accounts of decision-making are incompatible with the prevalence and success of ubiquitous marketing strategies. In this paper, we demonstrate, using computational experiments, how an ideal Bayesian observer model of preference learning is compatible with the manipulation of purchasing decisions via a number of well-known marketing techniques. The ability of this model to predict the effects of both familiar and novel marketing interventions suggests it as a plausible candidate theory of consumer marketing. Simultaneously, by clarifying the logic underneath the interplay between environmental exposure and preference distortions seen in economic decisions, this model rationalizes the seemingly irrational susceptibility of consumers to marketing.

Keywords: decision-making; preference learning; advertising; marketing; rational analysis

Introduction
Marketing constitutes a genre of economic activity that is mysterious to existing formal accounts of consumers’ decision process. While such formal theories require consumers to be economically rational, doing so would make them impervious to marketing techniques. In fact, the very existence of marketing as a viable genre of activity violates the predictions of current formal accounts of consumer behavior. What possible new information can the 43rd viewing of an insurance company’s ad give a consumer? Clearly, consumers receive a lot more information about products than just their ‘utility’ through such repeated interactions. Such associative influences have been difficult to document and incorporate into formal theorizing - hence have historically been ignored in marketing research - save as unspecified exogenous influences to be parametrized in econometric analyses. Now that online activity can be meticulously logged across content platforms, we argue that the sort of side information that was treated as noise in earlier generations of marketing theories can be incorporated to construct computational models that can make testable predictions about the efficacy of marketing interventions. This is the goal we pursue in this paper.

We do so by developing a psychological model of preference formation that can quantitatively relate manipulations of marketing variables to consumer demand. Our approach diverges from existing accounts of consumer/buyer behavior in several key aspects. First, unlike classic (Belk, 1975) and modern (Malhotra, 1988) integrated models of consumer behavior, our model can offer constrained quantitative predictions by virtue of relying on only observable variables (such as price distributions, exposure frequency, and transaction history), rather than relying on unobservable and immeasurable consumer valuations and beliefs. Second, unlike current quantitative models of consumer psychology behavior that consider choice mechanisms in very narrow settings (e.g., reference price models; (Winer, 1986)), we attempt to provide a general account that can capture the effects of many marketing interventions. Finally, unlike classical approaches designed for brick-and-mortar retail that have relied on population-level market and consumption variables, our approach considers frequency distributions of individual observers’ transactions, which are increasingly more measurable and relevant in internet commerce.

Existing models of consumer behavior
In the absence of formal theory, existing quantitative models of marketing are primarily econometric - they regress multiple available variables against outcomes of interest, use focus groups or deductive arguments to suggest that such variables can be changed by particular marketing interventions, then extrapolate these changes to the consumer base to predict how much the underlying outcomes will change.

Econometric models of marketing interventions are fundamentally data analytic models that impose microeconomic constraints on estimated parameters. Thus while they are good at retrospectively estimating the effects of marketing interventions on demand curves, they can only make predictions about such effects by extrapolating parameters. Perhaps the most rigorous models of consumer behavior are in the domain of pricing. Price sensitivity has been shown to follow a Weber law, such that consumers are sensitive to proportional price changes (Monroe, 1973). Moreover, consumers seem to evaluate prices relative to a “reference” price range that varies across products and categories (Kalyanaram & Winer, 1995), appears to be learned from transaction history (Emery, 1969), and may be influenced by brand strength (Biswas, 1992). These models are typically used to explain and motivate narrow experimental manipulations, and while they hold promise for predicting changes in aggregate demand curves from transaction history, they have not been applied in this way; perhaps largely due to the fact that they do not integrate the effects of long-term marketing strategies. Although these models can capture the effects of long-term marketing strategies on demand curves via free parameters to account for changes in reference price with branding, advertising, etc., they do not offer a predictive account of how marketing actions will influence the reference price, and thus can only retrospectively describe their effects on demand curves.

In contrast, theories of consumer behavior that aim to explain the psychological mechanisms of a broad range of marketing interventions rely on qualitative, verbal accounts of psychological processes and invoke unobservable, and immeasurable, latent traits and beliefs of consumers (e.g., (Bettman, 1979)). While these theories offer pithy qual-
ative summaries of marketing researchers’ intuitions about the psychology of consumers, they are neither designed to, nor capable of, offering quantitative predictions.

The account we present in this paper aims to capitalize on the strengths of these different approaches: First, by constructing a model at the level of individual choices, we capture intuitions about the psychology underlying consumer behavior. Second, by basing the individual choice on historically observed price distributions and transactions, we accommodate known relative price-range effects. Finally, by relying on only externally observable quantities as the inputs to the individual choice model, we make our theory empirically identifiable in the same manner as standard econometric models.

A cognitive model of consumer psychology

Ultimately, population demand curves are created from aggregating individuals’ buy/don’t buy decisions, therefore any formal analysis of the efficacy of marketing techniques must model how they influence individual purchasing choices. Framed this way, the question such an analysis must ask is ‘how do prices and other market signals influence purchase decision?’

The standard way of addressing this question is to treat choices as the outcomes of utility maximization. On this view, whatever choices an observer makes can be attributed to some underlying hedonic calculation which shows a higher evaluation for the chosen option. While this is a mathematically elegant way of describing the choice outcome, it is a very poor description of the process underpinning these choices. Prior research has demonstrated that consumers’ price estimates of products tend to be drastically altered by presentation formats (Tversky & Kahneman, 1981), the set of available options (Huber & Puto, 1983), as well as a variety of seemingly irrelevant psychological primes (Ariely, Loewenstein, & Prelec, 2006).

The success of various marketing strategies in increasing consumer preference for the same underlying product (Kirmani & Rao, 2000) lends credence to a less optimistic view of consumer preferences: choices are based on dynamic, context-dependent comparisons between options, rather than reliable hedonic value judgments (Ariely et al., 2006). Consumers are likely to make any particular decision by drawing upon past experiences with choices among similar options (Gilboa & Schmeidler, 1995). Given variability in experiences, variability in recall, and variability in the comparison process used to generate preferences, the resulting preferences will be considerably uncertain. Our theory is that marketing strategies capitalize on this uncertainty by manipulating the information available to observers at intermediate steps of the preference-construction process to influence preferences.

The principal contribution of our work is demonstrating how Bayes-optimal combination of prior choice-relevant observations yields an interpretable, simple, testable, and parsimonious account of marketing psychology. In particular this account predicts the efficacy of a number of interesting marketing strategies on several important consumer choice outcomes by virtue of their influence on a small, factored representation of consumer price history and knowledge.

Consumer representation

What are the observable building blocks of a theory of marketing psychology? An intuitive simplification of a typical economic transaction is that a buyer decides that the price for a particular product is fair in a particular context. Thus the observable units of individual transactions are consumers, prices, products, choices, and auxiliary contextual information (e.g., physical location, web portal, company brand, etc). Of these units marketers cannot directly influence consumers’ choices (b), but they can affect prices (m), products (x), contexts (c), and critically, the frequency and co-occurrence statistics with which consumers encounter each.

Although the full set of experiences of an observer can be described as a joint distribution of \( p(b,c,x,m) \), there are several reasons to consider the agent’s representation not as this complete joint distribution, but instead a factored set of several conditional distributions. First, it seems implausible for humans to keep track of the full joint distribution given the extreme sparsity of observations therein. Second, an argument from introspection suggests that not all conditional probabilities are equally easy to access as we would expect if they were all calculated from the same joint distribution: e.g., \( p(m|x,c) \) (how much does yogurt cost at Safeway?) seems intuitive while \( p(x|m,c) \) (what costs $5 at Safeway?) seems to require an awkward explicit search. Third, by factoring the joint distribution, a consumer can learn about the distribution of goods and prices from observing the transactions of others independently of tracking her own choices. Finally, a fourth, practical, reason to factor the joint distribution in a consumer choice model is that it makes it usable for predicting consumption behavior; whereas a model based on the full joint distribution would be inestimable to marketers who do not have access to the full set of experiences of a particular consumer.

Thus, to retain psychological plausibility, and practical usability, we assume that individuals represent the important elements of only some conditional and marginal probabilities from the joint distribution of purchasing decisions, products, prices, and contexts. Specifically, we assume consumers learn the following distributions from observations of the world around them:

- \( p(c) \) - what contexts populate a consumer’s daily life?
- \( p(x|c) \) - what products are available in this context?
- \( p(m|x,c) \) - how much does this product cost in this context?

And from their own experience, they keep track of:

- \( p(b|m,x,c) \) - how often do I purchase a particular good in a particular context, at a particular price?
Consumer choices

These tracked conditional distributions can be combined via the rules of probability to estimate the joint distribution over consumption choices, products, prices and contexts; and thus any conditional distribution of interest. Of particular interest in our case are the conditional distributions that observers must use to make consumer decisions:

- $p(b|x)$ - do I want to buy product $x$? (preference)
- $p(m|b,x)$ - what price am I willing to pay to buy $x$? (valuation)
- $p(b|m,x)$ - how does willingness to buy change with price? (demand curve)
- $p(b|c)$ - will I make a purchase in a given context?
- $p(c|b,x)$ - if I am going to buy $x$, in which context will I do so? (brand/retailer selection)

Each of these distributions capturing key aspects of consumer behavior can be predicted by marginalizing and conditioning the joint distribution obtained via

$$p(b,m,x,c) = p(b|m,x,c)p(m|x,c)p(x|c)p(c).$$

On our account, consumers determine their propensity for buying particular goods using accumulated evidence of previous purchases:

$$p(b|\tilde{x}) = \frac{\sum_{c,m}p(b|m,\tilde{x},c)p(m|\tilde{x},c)p(\tilde{x}|c)p(c)}{\sum_{c,m}p(m|\tilde{x},c)p(\tilde{x}|c)p(c)}.$$  \hspace{1cm} (1)

What is more interesting to a firm, though, is finding the greatest price a consumer would be willing to pay to purchase a product. Prior research has suggested that people typically generate a range of prices that they would be willing to pay for a product (Mazumdar, Raj, & Sinha, 2005). We formalize this intuition by casting this as a distribution over possible prices,

$$p(m|b=1,\tilde{x}) = \frac{\sum_{c}p(b=1|m,\tilde{x},c)p(m|\tilde{x},c)p(\tilde{x}|c)p(c)}{\sum_{c}p(b=1|m,\tilde{x},c)p(\tilde{x}|c)p(c)}.$$  \hspace{1cm} (2)

which directly gives us the distribution of prices at which consumers are willing to purchase a good.

With only a slight reformulation, this yields the relationship needed to obtain classical demand curves: purchase propensity as a function of price:

$$p(b|m=m_a,\tilde{x}) = \frac{\sum_{c}p(b=m_a|m_a,\tilde{x},c)p(m=m_a|\tilde{x},c)p(\tilde{x}|c)p(c)}{\sum_{c}p(m=m_a|\tilde{x},c)p(\tilde{x}|c)p(c)}.$$  \hspace{1cm} (3)

Of particular interest to a retailer, is the propensity of consumers to purchase while in their store,

$$p(b|\hat{c}) = \frac{\sum_{c,m}p(b|m,x,\hat{c})p(m|x,\hat{c})p(\hat{x}|c)p(c)}{\sum_{c,m}p(m|x,\hat{c})p(x|\hat{c})p(c)}.$$  \hspace{1cm} (4)

Finally, brands and retailers alike are interested in the likelihood that a consumer will choose their store or brand when making a purchase of a particular product:

$$p(c|b=1,\hat{x}) = \frac{\sum_{m}p(b=1|m,\hat{x},c)p(m|\hat{x},c)p(\hat{x}|c)p(c)}{\sum_{c,m}p(b=1|m,\hat{x},c)p(\hat{x}|c)p(c)}.$$  \hspace{1cm} (5)

Critically, each of these key facets of consumer choice and behavior will change in predictable ways under various marketing interventions designed to alter the conditional distributions that consumers keep track of. Thus, this formal setup, while sparse, allows us to test the influence of manipulating prices and context information on consumer demand curves.

Model predictions

To substantiate our intuitions about marketing-based distortions of consumer preferences, we simulated a small test market, containing three purchase contexts, two goods, and five price labels where a consumer’s purchases were generated via the following generative model. A purchase context was sampled from a random seed distribution $p(c)$, a product was sampled from a discrete random seed probability $p(x|c)$ for this context, a price label was sampled from a random seed probability $p(m|x,c)$ for the already sampled tuple $\{x,c\}$. Finally, this observation was flagged as a purchase decision with a small probability ($p = 0.2$), and within the samples thus flagged, purchase decisions were randomly generated while maintaining an inverse relationship with price.

Using this generative procedure, we sampled 10000 events to obtain baseline empirical estimates for each of the conditional distributions implicated in our account. The experimental results we report in succeeding sections were constructed by appending this baseline event history with manipulated event sequences corresponding to various marketing interventions.

Rationalizing product-brand associations

The most obvious form of marketing is advertising by displaying the product and its associated brand. This form of advertising could be rationalized as providing information to potential consumers. It is harder to make a similar argument for event sponsorships and brand awareness campaigns, wherein companies advertise only brands, not products. What rational purpose is served by simply presenting the company’s logo to a consumer, disconnected from product information? Also, why belabor people with redundant and uninformative visuals over and over again? Surely once or a couple of occasions would be enough to convey any information? Why are “tip of tongue” (Mowen & Gaeth, 1992) and brand recognition metrics (Munoz & Kumar, 2004) so popular, influential, and desirable? The answer, of course, is that firms aim to increase the rate at which consumers think of their brand. But why would increasing the ease with which consumers think about the brand change consumer purchasing decisions?
On our account, changing brand recall and recognition amounts to changing the context probability \( p(c) \) for that brand (Figure 1 left). The immediate effect of increasing brand recognition and recall is an increase in \( p(c|b,x) \): given that a consumer has decided to buy a product, which brand will she choose? So long as the brand is associated strongly with products a particular product \( x \) \( (p(x|c)) \) an increase in \( p(c) \) will yield an increase in \( p(c|b,x) \); in other words consumers will be more likely to choose brand \( c \) when asking themselves “I want to buy an \( x \), which brand/retailer should I choose?”.

However, our model also predicts that simply increasing \( p(c) \) will have no effect on the consumers’ eagerness to buy its specific products \( p(b|x) \) or increase their eagerness to buy the brand \( p(b|c) \). Our account suggests one immediate strategy for increasing consumers’ eagerness to buy the brand: selectively increasing \( p(x|c) \) for \( x \) with high \( p(b|x) \) – in other words, strategically associating the brand with desirable goods. If the advertising that increases \( p(c) \) also strategically increases \( p(x|c) \) in this manner, then not only are consumers more likely to choose brand \( c \) when making a purchase \( (p(c|b)) \), but they will overall be more likely to purchase the brand \( (p(b|c)) \). Moreover, this increase in propensity to buy the brand yields a uniform increase to the demand curve for the brand \( (p(b|m,c)) \); Figure 1 right), showing just how effective a carefully selected increase in brand-product association can be.

Another interesting theoretical prediction from our model concerns the overuse of promotions presenting that brand without an associated product; this may be counterproductive as it might result in product-brand delinking. This could occur if, for instance, a company overemphasizes event sponsorships over product ads, such that the linking probability \( p(x|c) \) is diluted by frequent observations of brand \( c \) without associated products \( x \). Since such dilution will be accompanied by \( p(c) \) gains, this will be a risk primarily for already familiar brands, for which \( p(c) \) improvements are showing diminishing returns. In such situations consumers will show high brand awareness \( p(c) \), but this will not translate into changes in consumption behavior \( p(b|c) \).

This account also reaffirms other important elements of brand competition. In particular, it emphasizes product differentiation (Dickson & Ginter, 1987), frequently cited as one of the major causes of ad campaign failures. If the product \( x \) that a brand is associated with is considered to be a unique entity (e.g., “a Diet Coke”) rather than a generic category (e.g., “a diet cola beverage”), then the gains of increased brand recognition will translate directly to increased demand for that brand’s product. However, when a market is over-crowded, product differentiation becomes harder and costlier, thus gains in \( p(c) \) will be lost because \( p(x|c) \) does not adequately pick out the product of that particular brand, thereby reducing the potential gains from a higher \( p(c) \). Furthermore, this account emphasizes the arms race nature of branding campaigns – the advantage is determined by relative frequency, rather than absolute frequency of brand exposure, which naturally imposes barriers to entry in existing competitive markets, as suggested previously by (Schmalensee, 1982) using empirical data.

**Rationalizing loss-leader strategies**

Classically, the economic tension between the retailer and consumers’ incentives maintains a price equilibrium. One potential advantage for the retailer is the relatively high costs of searching for low prices for every product, which motivates consumers to generalize about price (dis)advantages of retailers in aggregate, rather than for isolated products. Thus, insofar as consumers use aggregate price advantages to predict
prices for new products, there is an opportunity to offer some carefully chosen products at a discount, and thus distort the aggregate inference: \( p(m|c) = \sum c p(m|x,c) p(x|c) \).

Specifically, by choosing to offer price reductions on salient products (high \( p(x|c) \)), promoters can skew the consumers’ estimates of the overall priciness of brand \( c \). Figure 2A-B shows this strategy in action. Reducing prices on highly observed items reduces observers’ estimates of how pricey a particular retail outlet might be \( p(m|c) \) which informs their propensity to make purchases in such contexts \( p(b|c) \).

Although exploiting salient products to distort consumers’ estimates of overall price tendencies may seem exotic, it is not original to this paper. Amazon adopts this very policy by undercutting competitors on their most popular products, while increasing prices on less salient goods (Del Rey, 2015). This strategy makes perfect sense under our account: with a keen enough understanding of \( p(x|c) \) (which we expect Amazon’s Big Data provides) it may even be possible to increase overall prices while simultaneously decreasing consumers’ estimates of the prices offered by the brand.

### Rationalizing money-losing brand extensions

The principal way in which manufacturers can benefit from brand ‘equity’ is by extending the repertoire of products associated with it. The problem lies in the possibility of dilution of the brand’s association with individual products by virtue of exposure alongside multiple products. In the simple account of brand extensions, manufacturers bring new products to market to increase profits at the expense of brand equity. Our analysis, however, reveals the possibility of an inversion of this basic process - a manufacturer could potentially improve brand equity by bringing a low-priced new product to market - trading off profits (or even losses) for brand equity.

In this situation, a company would manufacture a new product that sells for low prices at high volumes, and is indelibly associated with the company’s brand. Such a scenario would most likely play out for companies whose primary products are big-ticket, low volume items, e.g. cars, vacations etc, and that are looking to improve their visibility. Availability of the product at sufficiently low prices will raise \( p(b|x,m,c) \), which will in turn increase not just \( p(b|x) \) for this low-price and likely low margin product, but also \( p(b|c) \) and \( p(c|b) \), thus increasing brand equity at fairly low cost.

To test this possibility, we added exposure to a new good specific to a particular context to the baseline event history in our simulation, available at the bottom two price labels in a ‘cheap’ condition and at the top two price labels in a ‘pricey’ condition. We measured gain in brand equity as relative change in \( p(r|c) \) from that measured in the baseline condition for this context. Figure 2C, which plots the relative gain in \( p(r|c) \) for 100 model simulations from different initializations, shows how brand equity improves through adding a loss-leader, and drops through adding a relatively expensive product to the product line. The latter is more profitable, so this simulation demonstrates the existence of a competitive tension between brand equity and capital - companies could potentially trade one off against the other sequentially, modulo diminishing returns from product-line overcrowding.

### Conclusion

Beginning with the intuition that marketing strategies influence consumers’ preference formation processes via associative influences within the preference construction process, we have created a theory of consumer preference formation that is grounded strongly in observable correlates for marketing variables. With a series of computational experiments, we have substantiated various predictions that this model makes about the impact of both existing and novel marketing strategies, thus rationalizing several lines of consumer research findings via a simple inductive explanation of how consumption preferences are formed. The model opens up a large space of possible experiments testing the effect of each of the variables we have defined on consumer behavior. Table 1 suggests a number of directional hypotheses derived within our framework. We expect the strong observability of our model, in combination with its novel hypotheses, will benefit both theory and practice of marketing and consumer research, particularly in online retail settings, where the conditional distributions implicated in our account are easy to access.
Figure 2: Predictions for advanced and speculative marketing strategies (Left) Flooding retail displays with cheap or discounted goods reduces observers’ internal estimates of the price distribution \( p(m|c) \), (middle) which promotes their propensity to make purchases in the retailer’s chosen context. (Right) Similarly, the introduction of a cheap brand extension to the market can result in an increase in \( p(b|c) \) – a measure of brand equity. All changes are measured from baselines estimated on the initial event history. Histograms show results for 100 simulations each.

References


