Distribution Agreement

Signature:

______________________________ ______________

Chen Lin Date

MODELING COMPLEX DECISIONS UNDER MULTIPLE CONSUMPTION SCENARIOS

By

Chen Lin

Doctor of Philosophy

Business

Douglas Bowman, Ph.D.
Advisor

Sandy D. Jap, Ph.D.
Co-Advisor

Michael Lewis, Ph.D.
Committee Member

Manish Tripathi, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

MODELING COMPLEX DECISIONS UNDER MULTIPLE CONSUMPTION SCENARIOS

Chen Lin

B.COMPUTING (Hon.), National University of Singapore, 2007

Advisors: Douglas Bowman, Ph.D.

Sandy D. Jap, Ph.D.

An abstract of

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Business

2012

ABSTRACT

MODELING COMPLEX DECISIONS UNDER MULTIPLE CONSUMPTION SCENARIOS

The objective of this dissertation is to understand complex consumer decisions under multiple consumption scenarios accounting for the interdependencies across product categories within the consumption portfolio in order to capture a full model of consumer behavior. The context includes multi-category/multi-channel media choice (Essay 1), and broader applications of consumption in seemingly disparate product categories (Essay 2).

Specifically, Essay 1 focuses on predicting media choice and time allocation, taking into account interdependencies between traditional and new media under a utility-maximizing framework. Using a rich database of individual-specific media activity diaries, it suggests that accounting for media interdependencies is extremely important and generates unique insights on consumer-level media switching, media-multiplexing, potential sources of substitutability/complementarity resulting from media attention and penetration, and consumer heterogeneity often ignored in aggregate data.

Essay 2 takes a first step in modeling difficult-to-observe psychological processes that govern consumer decision making by examining consumption across seemingly disparate categories. This research proposes a hierarchical multinomial processing tree model to empirically examine the driver, which is defined as the "latent trait", which governs consumer choices across five seemingly disparate product categories: media consumption, automobile purchases, financial investments, soft drinks and cell phone plans through a dataset consisting of 5,000 consumers in the United States. Essay 2 further investigates how consumer behavior systematically varies from one category to another and suggests new approaches to segment and profile consumers based on behavior across multiple categories. Finally, by comparing the latent trait approach with the latent class approach, it contrasts discrete and continuous representations of consumer heterogeneity and discusses related substantive and empirical issues.

MODELING COMPLEX DECISIONS UNDER MULTIPLE CONSUMPTION SCENARIOS

By

Chen Lin

B.COMP (Hon.), National University of Singapore, 2007

Advisors: Douglas Bowman, Ph.D.

Sandy D. Jap, Ph.D.

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Business

2012

ACKNOWLEDGEMENT

I would like to express my heartfelt gratitude to my advisors, Dr. Douglas Bowman and Dr. Sandy D. Jap for their invaluable guidance, patience, support, and most importantly, their mentorship throughout my graduate study at Emory University. I would also like to gratefully and sincerely thank my doctoral committee members, Dr. Michael Lewis and Dr. Manish Tripathi, for their kind sharing of knowledge, expertise and advice in helping me with my doctoral work. Essay 1 has benefited from early interactions with Dr. Sriram Venkataraman. Dr. Jagdish N. Sheth and Dr. Ryan Hamilton also provided useful feedback and suggestions on Essay 2. I also appreciate the generous support and great learning experience provided by the Marketing Department and Goizueta Business School. Finally, nothing can be achieved without the everlasting love from my dearest family and beloved ones, Zhijian Lin and Shuping Dong. Thank you.

Table of Contents: Chapters

Table of Contents: Tables

Table of Contents: Figures

Table of Contents: Appendices and Supplementary Materials

1 INTRODUCTION

"To look at a leopard through a tube, you can only see one spot."

-From Ancient Chinese Idiom (422 AD)

The premise of marketing research is built upon investigating closely related variables and decisions. Marketing strategies are planned and delivered with an integrative mix of the 4P's (Price, Product, Promotion and Place), each of which may form a hierarchy of interrelated components. Consumer decision making is a compound function of his/her past experiences, firm communications, social influences, individual differences and other contextual factors. Understanding interactions and interdependencies among these complex relationships often yields richer and more complete insights on the value creation process than examining a single piece or a subset of information independently.

The objective of this dissertation is to understand complex consumer decisions under multiple consumption scenarios accounting for the interdependencies across product categories within the consumption portfolio. The context includes multi-category/multi-channel media choice (Essay 1), and broader applications of consumption in seemingly disparate product categories (Essay 2). While both essays aim to examine a full model of consumer behavior, Essay 1 focuses on predicting media choice and time allocation taking into account interdependencies between traditional and new media under a utility-maximizing framework, whereas Essay 2 undertakes a more general approach to investigate the underlying latent processes that govern consumer decision making across seemingly unrelated product categories through a multinomial processing tree model.

Specifically, Essay 1 examines “media multiplexing” behavior, which refers to consumers serially consuming small “chunks” of multiple media types (television, radio, Internet and print) within a short time period, and notes that a key challenge for integrated marketing communications (IMC) media planners is predicting which media or combination their target audience is likely to consume at any given time. I propose a structural forecasting model that incorporates media-multiplexing behavior of both traditional and new media, their interdependencies, time effects and consumer heterogeneity and calibrate the model using a rich database of individual-specific media activity diaries. Results suggest that accounting for media interdependencies is extremely important even if media-multiplexing is not critical to the media planner or is not commonly occurring in the data. My individual-level analyses generate unique insights on consumer-level media switching, media-multiplexing, potential sources of substitutability/complementarity resulting from media attention and penetration, and individual heterogeneity often ignored in aggregate data. I demonstrate another application of my model by predicting consumers' media choice after an exogenous shock in the marketplace, i.e. the exit of print media.

Essay 2 takes a first step in modeling latent processes that govern consumer decision making by examining consumption across seemingly disparate categories. Marketing activities today are coordinated in a variety of categories and in a variety of formats, and consumers naturally shop around a globe of unrelated product categories that are beyond the traditionally defined “shopping basket”. This research proposes a hierarchical multinomial processing tree model to empirically examine the driver, which is defined as the "latent trait", which governs consumer choices across five seemingly disparate product categories: media consumption, automobile purchases, financial investments, soft drinks and cell phone plans through a dataset consisting of 5,014 consumers in the United States. I further investigate how consumer behavior systematically varies from one category to another and finally suggest new approaches to segment and profile consumers based on latent traits across multiple categories. In doing so, this paper contributes to the consumer decision literature in three ways: 1) theoretically, the latent-trait approach provides rich support in examining the underlying psychological processes; 2) methodologically, the relative merits of models with continuous versus discrete representations of consumer heterogeneity are discussed; and, 3) substantively, new insights on targeting and profiling based on latent processes rather than observed behavior are presented with respect to managing across seemingly unrelated product categories.

To summarize, this dissertation takes an integrative view in examining the full picture of consumer behavior and understanding interdependencies in multiple consumption scenarios through structural and psychometric models. New insights on targeting, media planning, and profiling are discussed.

2 ESSAY 1: MEDIA MULTIPLEXING BEHAVIOR: IMPLICATIONS FOR TARGETING AND MEDIA PLANNING

2.1 INTRODUCTION

Innovations in networking products, mobile technologies, and adoption of high-speed Internet have significantly (and permanently) altered the manner in which consumers' utilize media today. From 1999-2004, consumption of the Internet with alternative media – namely television (TV) – has grown. Figure 1 indicates that their total combined utilization has increased from 174 minutes per week in 2001 to 300 minutes per week in 2004. This difference represents a 72 percent increase over a 4-year period. According to Nielsen's latest Three Screen report[1], Americans now spend 35 percent more time using the Internet and TV at the same time than they did in 2008. That translates to about 3.5 hours of overlapping TV/Internet time per month for the average American consumer. Nearly 59 percent of consumers reported doing this at least once a month as well.

Integrated marketing communication (IMC) specialists have already begun responding to these changes by coordinating their advertising across multiple media taking into account the cost advantages and the improved targeting ability of new media relative to traditional broadcast media. In 2008, of the $141.7 billion spent on advertising in the United States, Internet advertising grew by 7.3 percent, while traditional print media advertising declined by 19.7 percent over 2009[2].

These shifts in the way we consume media coupled with the increasing trend in coordinated marketing activities across multiple media channels underscore the need for a detailed understanding of the consumer's choice of new and traditional media channels

[1] http://blog.nielsen.com/nielsenwire/wp-content/uploads/2010/03/3Screens_4Q09_US_rpt.pdf
[2] Kantar Media, http://www.kantarmediana.com/news/03172010.htm

as well as media interdependencies (Schultz, 2002). This is because the success of any IMC program today hinges on the media planners' ability to harness potential synergies[3] across multiple media (Naik and Raman, 2003). Despite the significance of this need, there is scant research in this area (Mantrala, 2002; Danaher and Rossiter, 2011).

Armed with such a predictive model, IMC planners can develop a strategy for targeting consumers across different media channels. We provide the foundational beginnings for this task. Specifically, we investigate the phenomenon of media multiplexing - whereby people serially (and potentially simultaneously) consume small, most likely incomplete "chunks" of media within a short period of time and may then switch to the consumption of programming/content on another media channel (Simmons Research Report 2010; Pilotta et al. 2004; Pilotta and Schultz, 2005; Smith et al 2006). One implication of this behavior is that content creators may need to produce content in smaller chunks so that media multiplexing consumers consume a complete segment of program content and perhaps continue for the next segment[4]. Advertisers and media planners who previously focused on predicting consumers' single media choice in isolation of other media now face the problem of tracking *where* a target audience may be at any given point in time and predicting which media or combination of media their target audience is likely to consume given potential inter-media synergies.

We introduce a consumer-level demand model and calibrate it using a proprietary dataset of individual-specific media consumption diaries (recorded for one week for each

[3] Synergy is defined as occurring when "the combined effect of multiple activities exceeds the sum of their individual effects" (e.g.. Belch and Belch 1998, p. 11). Naik and Raman (2003) show the retail-level sales gains from running the same advertising campaign simultaneously across both television and print media instead of just one media.
[4] Fox's American Idol series is a classic example of how key announcements are made after advertising breaks so as to reduce the likelihood of consumers switching to competing channels.

individual in half-hour increments). From a methodological standpoint, in order to account for media multiplexing, we relax the discrete-choice assumption of commonly employed choice models. The flexible-demand specification we introduce also allows us to account for richer sources of inter-media dependencies, including: (1) media switching and co-consumption, (2) potential complementarity and substitutability of the channels, (3) observed and unobserved consumer heterogeneity and process variation (e.g., satiation), and (4) day, time, and media specific variables such as technology penetration and attention span. These factors in concert rationalize consumers' media choices, be they single media or multiplexing. I use the proposed demand model to address the following questions:

1) <u>What is the improvement in predicting consumers' media choices if we account for inter-media synergies and multiplexing?</u>

   Accounting for media synergies within a single utility specification significantly improves model forecasts. For a single media choice, the overall hit rate of my model is approximately 37 percent better than a corner-solution-only discrete continuous model (i.e., a single category Hanemann model). This suggests that accounting for multiplexing tendencies, as is done in my model, is extremely important. When consumers multiplex, the relative performance of my main model is even better. My main model generated perfect predictions for five of the 11 media multiplexing configurations observed in my data.

2) <u>For a given target audience, which media are complements and which are substitutes?</u>

Using a data-driven approach, I examine controllable media specific factors, media attention and penetration, that impact substitution patterns across media alternatives. *Ceteris paribus*, holding the same attention level across media alternatives, the baseline utility for television is the highest relative to computer, radio, print or any possible combination of these alternatives alone. However, the cross-media effect of attention on a media-specific baseline utility suggests interesting asymmetries across the media alternatives. For example, consumers with high attention for the computer are unlikely to multiplex traditional forms of media (radio, print, and television). In contrast, consumers with high attention for these traditional media options have higher utility for multiplexing all forms of media (including new and traditional).

We are also able to identify the consumer's differential utility for each media and, hence, which media they are willing to give up. For example, consumers who own personal computers at home are unlikely to give up print media, whereas consumers who use computers at their workplace are more likely to do so. Additionally, cable television subscribers tend to have lower valuations for computer media. I relate my findings to marketplace trends and discuss the impact on print media in a subsequent counterfactual study.

3) <u>Which media will consumers migrate to in the event of an exogenous shock to a media alternative (e.g., the exit of print media)?</u>

We consider the consequences of the foregoing questions via a counterfactual analysis for two large markets, namely Los Angeles and New York City. The issue is particularly timely, as a growing number of regional newspapers struggle to remain profitable. We find that in Los Angeles, 54 percent of the previously print media consuming consumers' would switch to other social activities (that do not involve any other media) and 33 percent would switch to computer, while 10 percent would switch to TV and three percent to radio. Similar results are found for New York City. 54 percent switch to other social activities and 33 percent switch to computer, just as in LA. However, only seven percent would switch to TV while five percent would switch to radio. These findings provide an exploratory understanding of local-market-level media migration.

My study serves as a critical *first step* toward a more detailed understanding of consumers' media choices taking into account media channel interdependencies and multiplexing behavior. Ultimately, I hope future research in this area will facilitate a better understanding of how firms can achieve higher levels of efficiency, synergies, and effectiveness in their advertising resource-allocation efforts both within and across media. For example, equipped with the proposed model, consumers' media choices and consumer-directed advertising schedules, one can extend the models proposed by Rust and Alpert (1984), Danaher and Rust (1996) and Wilbur (2008) to a multi-media channels setting[5].

The remainder of this paper is organized as follows. In 2.2 I briefly review related streams of literature and my general approach to modeling multiplexing behavior. This is followed by the empirical model in 2.3, a description of the data (2.4), and estimation and results in 2.5. I then consider the complementary and substitutionary roles of media in conjunction with multiplexing behavior in 2.6, a counterfactual analysis of the

[5] Although this data do not allow us to examine issues such as return-on-investment directly, they allow me to provide novel insights that would move us toward that goal.

elimination of print media in 2.7, and conclude with a discussion of key findings, implications for management, limitations and directions for future research in 2.8.

2.2 LITERATURE REVIEW

The IMC literature emphasizes effective interaction and integration of various forms of marketing communications. It is built on the premise that the appropriate deployment of multiple media strategies - for example, general advertising, direct response, sales promotion and public relations - can clarify and make communications consistent. However, the empirical IMC literature to date has focused either on understanding consumers' media firms' individual media choice such as television (Vakratsas and Ambler, 1999), radio, or Internet (Manchanda et al. 2006). Within the focal media, these studies either measure the effectiveness of these individual channels for consumer-directed advertising, or they model the consumers' single media choice problem such as the selection of television program channels (Rust and Alpert 1984, Shachar and Emerson 2000), thereby generating insights for the creation and scheduling of TV content. Despite widespread reporting in industry trade publications, to the best of my knowledge, consumers' choice of multiple media (be it sequential or simultaneous) has not been addressed in the current literature. Pilotta and Schultz (2005) claim that the experience of media multiplexing is a shift in the logic of cultural perception and attunement from successive experience to simultaneity and synesthesia of media that, in turn, restructures attention. To this end, research that investigates consumers' consideration of media channels across varying usage situations has found that the media channel's perceived benefits and perceived synergies affect consumers' media channel consideration (Wendel and Dellaert 2005).

There is also a related stream of research examining firm-side media choice within an IMC setting. Naik and Raman (2003) demonstrate the firm-side sales gains from advertising across two media (i.e., television and print). Their study is the first to demonstrate the retail-level sales gains from honing media synergies[6]. However, their study does not examine consumers' media-choice decision and therefore their ability to predict consumer sensitivity to media exposure is limited. However, the focus of this line of research on synergy is the consequential impact of media consumption, whereas the focus of my research is in the prediction and nature of consumption and their migration away from these channels in response to an exogenous shock.

In summary, although there is a growing interest in the use and management of multiple media and their resulting synergies, there exists a small amount of literature that helps marketing managers plan communications strategies across multiple media using individual consumer preferences. An exception to this is Smith et al., (2006) who use individual-call-level data across print, direct mail, event sponsorships, radio, computer, and retail sources as inputs to the firm's lead generation process. Their descriptive model accommodates serial correlation across the media sources, carryover effects, and concave response functions. However, this approach does not give insight into the nature of potential interdependencies across media such as whether the media channels act as substitutes or complements to each other. In contrast, my approach provides these insights as well as insights into the migration patterns of consumer choice across channels and the effects of exogenous shocks, such as the demise of a media channel.

[6] This study complements studies that have investigated the synergies in marketing-mix elements. Examples include Naik, Raman and Winer (2005) where synergies in price promotions and advertising are examined.

I use a *multivariate discrete-continuous choice* model that is utility-theory consistent and harnesses consumption interdependencies across multiple channels so as to rationalize individual-specific media multiplexing. I do this by drawing on recent advancements in the multi-category choice-modeling literature in marketing and economics. Specifically, I introduce a new consumer-level media forecasting model that directly accommodates media multiplexing of both traditional (television, radio and print) and new (Internet) media, along with unobserved heterogeneity and satiation.

In marketing, there has been an emerging trend toward understanding the cross-category effects of marketing-mix activities. Much of this research occurs in the grocery retail context (e.g., how consumer purchases in one category are related to their purchase in another). The goal of much of this literature is to help firms better coordinate their marketing activities across categories and products within each category so as to maximize profits. Multi-category demand specifications were developed to accommodate consumers' simultaneous purchase incidence in multiple categories/products. For example, Song and Chintagunta (2006) study cross-category price effects using aggregate store-level data, Manchanda et al. 1999 and Chib et al. 2002 focus on recovering category interdependencies using household-level multi-category purchasing decisions, and whether consumer preferences and sensitivity to marketing activities are household- and/or category-specific is presented by Ainslie and Rossi 1998, Seetharaman et al. 1999, Mehta, 2007.

Here, cross-category (or product) interdependencies are accommodated in a variety of ways, ranging from non-structural specifications such as Manchanda et al. (1999) and Edwards and Allenby (2003), to a more semi-structural one like Dubé (2004), to a completely structural approach as undertaken in Kim et al. (2002) and Bhat (2005). My approach is akin to Song and Chintagunta (2007) and Bhat (2005) in that multi-category interdependencies (i.e., media channels, in my case) are accommodated via a behaviorally consistent utility-maximization framework, while accounting for a mixture of discrete (e.g., which media or media channels to consume?) and continuous (e.g., how much to consume?) choices.

I introduce an econometric demand model that can account for media-multiplexing behavior, and in doing so builds on the seminal work in economics by Wales and Woodland (1983) and Hendel (1999). As in Kim et al. (2002) and Bhat (2005), consumers in my setting can choose one or more alternatives among imperfect media substitutes subject to a budget constraint (i.e., time, in my case). This is in contrast to Hendel (1999) and Dubé (2004), where multiple-discreteness is accommodated using a stream of future consumption occasions and a discrete-choice model among perfectly substitutable goods for each consumption occasion.

To summarize, I contribute to and advance the IMC literature as follows:

(i) I complement the Naik and Raman (2003) study where they utilize aggregate advertising data to speak to issues at the firm or market level, while I rely on individual-choice data across media and focus on predicting media choice/s at the *individual* consumer level. I calibrate the proposed demand model using proprietary individual-level diary panel data that tracks the specific media choices (newspapers, TV, radio or Internet) of 1,775 individuals. The panelists' media choices are known for each half-hour slot, every day, for seven consecutive days. I <u>do not</u> focus on the implications of consumers' media choices on retail sales/brand choices or advertising directed to the consumer through these media, as the data used in this essay only contain information on the panelist's media choice[7].

(ii) Methodologically, my utility-theory-consistent demand model incorporates the simultaneous consumption of different media (both traditional and new), time effects, media technology penetration and other exogenous factors[8]. Instead of treating media choices as independent of each other (which has been the dominant approach of the empirical IMC literature to date), I explicitly model the determinants of a consumer's media choice and interdependencies between media by drawing on the emerging multi-category choice-modeling literature (see Chintagunta 2002; Song and Chintagunta 2006). In doing so, I have strived to specify a richer theory-driven model of consumer behavior that explicates the determinants of consumer choice of one or more media versus other alternatives.

(iii) In order to accommodate media multiplexing, I relax the discrete-choice assumption of commonly employed choice models. Furthermore, multiplexing is accommodated within a single utility specification.

(iv) Finally, the ability to harness synergies across multiple media of the proposed approach allows media planners to consider how changes in media-market structure and reshaping forces impact media consumption and migration (c.f., Ansari et al. 2008). This issue of migration across multimedia channels is completely unexplored. By using the individual consumer as the unit of analysis, I

[7] This data vendor does not track the panelist-specific retail-level brand choice data. More details on the data can be found in section §4.

[8] The demand model I introduce can be applied to other settings where consumers choose multiple alternatives and are face some constraint that forces then to make appropriate tradeoffs in the consumption of these alternatives.

am also able to examine how changes in media-market structure - e.g., the elimination of print media (Featured in *TIME Magazine, Feb 16th,* 2009) - impact subsequent media consumption and media migration patterns.

2.3 THE EMPIRICAL MODEL

Consumer choice in our setting is characterized by the multiplexing of K media alternatives such as Internet, television, radio and print media within a one hour period[9]. The objective of this essay is to predict individual i's media consumption choice k (or multiple choices when the consumer engages in multiplexing), which is the discrete component in the proposed model, and as continuous media time allocations t_k (consumption quantities) while accounting for potential interdependencies across media alternatives. Therefore, a simple and parsimonious multiple discrete-continuous extreme-value model (Bhat 2005) or a discrete-continuous demand system akin to Kim et al. (2002) is particularly well suited for this context. In this research, I propose a stochastic utility specification variant of Bhat (2005) while incorporating important media planning decisions: timing effect (including time-of-the-day and day-of-the-week), observed and unobserved consumer heterogeneity, attention span and media penetration. I will compare the proposed model with a single discrete-continuous model (Hanemann 1984) in later sections to show the improvement in model fit and prediction in accounting for media multiplexing.

---

[9] . The choice of a one-hour period is largely a function of the data structure, which is collected via self-report diaries over half hour increments. I develop the model along one hour increments, but note that there is no theoretical reason why the increment should be one hour or more or less. In the estimation section I also demonstrate the model's flexibility to one hour, half hour, or even two hour increments and show evidence as to the robustness of our results.

I first specify the utility derived by any representative consumer in each hour block. The functional form of the proposed utility specification is a generalized variant of the translated consumer-expenditure-system (CES) utility function, and given by:

$$U_i(\mathbf{t}) = \sum_{k=1}^{K} \frac{\gamma_k}{\alpha_k} \psi_{ik} \left\{ \left(\frac{t_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\} \tag{1}$$

where $U(\mathbf{t})$ is a quasi-concave, increasing, and continuously differentiable function with respect to the dependent variable that consists of continuous consumption quantity (Kx1)-vector $\mathbf{t}$ ($t_k \geq 0$ for all k), which in my case is the time allocated to each media alternative. The model also specifies the first alternative to be the outside good, i.e. not consuming any media alternatives, but instead engaging in other activities. Multiplexing is accommodated within a single utility specification by summating across K media alternatives, with ψ_k , γ_k and α_k being parameters associated with good k to capture interdependencies across media channels.

Specifically, ψ_k reflects the baseline marginal utility, which controls whether or not a media option is selected for positive consumption (i.e. the extensive margin of choice). The role of γ_k is to enable corner solutions for which only one media option is chosen, though it also governs the level of satiation. Satiation in media is likely to differ across traditional and new media, and each individual may have a different capacity in consuming media over time. Higher values of γ_k imply a stronger preference (or lower satiation) for media k, i.e. the consumer is willing to give up more of other media alternatives to obtain one unit of alternative k. For example, the "couch potatoes" will have a much higher γ value for television. I will utilize these two elements of the proposed model to assess the extent of competition (or substitution) between different media. The purpose of α_k is solely to allow satiation. When α_k = 1 for all *k*, this represents the case of absence of satiation effects or, equivalently, the case of constant marginal utility. The utility function in such a situation collapses to $\sum_k \psi_k x_k$, which represents the perfect substitutes case as applied in Hanemann (1984). Intuitively, when there is no satiation, the consumer will invest all expenditures on the single good with the highest baseline (and constant) marginal utility (i.e., the highest value). This is the case of single discreteness. As α_k moves downward from the value of 1, the satiation effect for good *k* increases. Thus, for a given extensive margin of choice of media *k*, γ_k and α_k influence the quantity of media *k* consumed (or the intensive margin of choice) through their impact on satiation effects[10].

Furthermore, ψ_k , γ_k and α_k can all be further parameterized to reflect individual differences in evaluating or consuming media options. Specifically, a multiplicative random element is introduced to the baseline marginal utility of each media option as follows:

$$\psi_{ik} = \psi(z_{ik}, \varepsilon_{ik}) = \psi(z_{ik}) \cdot e^{\varepsilon_{ik}} \quad (2)$$

where z_{ik} is a set of attributes characterizing alternative *k* and the decision maker including baseline constants, time of the day, day of the week, demographics, attention levels and media penetration, and ε_{ik} captures idiosyncratic (unobserved) characteristics that impact the baseline utility for media *k* for individual *i*. $\psi(z_{ik})$ is further

[10] However, γ_k and α_k parameters for each media option cannot be identified separately. There are several ways to configure these parameters and researchers need to choose the one that best fits this data.

parameterized as $\exp(\beta' z_{ik})$, which then leads to the following form for the baseline random utility associated with media *k*:

$$\psi(z_{ik}, \varepsilon_{ik}) = \exp(\beta' z_{ik} + \varepsilon_{ik}) \tag{3}$$

The z_{ik} vector in the above equation includes a constant term. The overall random utility function of Equation 1 then takes the following form, with the first good being the outside good:

$$U_i(\mathbf{t}) = \sum_k \frac{\gamma_k}{\alpha_k} \left[\exp(\beta' z_{ik} + \varepsilon_{ik})\right] \cdot \left\{ \left(\frac{t_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\} \tag{4}$$

From the econometrician's perspective, the individual is maximizing random utility subject to the binding linear budget constraint that $\sum_{k=1}^{K} t_k = T$, where *T* is a total budget of 60 minutes, in this case.

For identification purposes, one set of γ_k and α_k has to be fixed. In cases where α_k values are estimated, these values need to be bounded from above at the value of 1, whereas in cases where γ_k values are estimated, these values need to be greater than zero. To enforce these conditions, α_k can be parameterized as

$\alpha_k = 1 - \exp(-\theta_k' y_k)$ or $1 - \dfrac{1}{1 + \exp(-\theta_k' y_k)}$ depending on user specifications, and

$\gamma_k = \exp(\varphi_k' w_k)$; $z_k, y_k,$ and w_k are individual-specific variable vectors and $\beta, \theta,$ and φ are corresponding coefficient vectors. That is, the satiation parameters (*i.e.,* α and γ values) are specified as functions of observed and unobserved individual attributes.

In my identification setting, separate α values (rather, the corresponding functional forms) are estimated for all goods including the outside good. The γ values for all goods (except the outside good) are constrained to be equal to 1 (*i.e.,* the corresponding φ values are constrained to be equal to 0).

2.3.1 OPTIMAL TIME ALLOCATIONS

Optimal time allocation occurs when individual consumer *i* chooses the media channel and consumption quantities that maximize his or her utility for under the one-hour time constraint. This is achieved by forming the Lagrangian and applying the Kuhn-Tucker (KT) conditions. I derive the individual-specific consumption vector for the random utility-specification subject to the linear time constraint. Purely for convenience sake, I drop the subscript *i* for now. The resulting Lagrangian is:

$$\mathrm{L} = \sum_k \frac{\gamma_k}{\alpha_k}\left[\exp(\beta' z_k + \varepsilon_k)\right] \left\{\left(\frac{t_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right\} - \lambda\left[\sum_{k=1}^{K} t_k - T\right] \tag{5}$$

where λ is the Lagrangian multiplier associated with the time constraint (that is, it can be viewed as the marginal utility of the total time)[11]. Recall, the KT approach immediately satisfies all the restrictions of utility theory, and the resulting first-order conditions provide the basis for deriving the probabilities for each possible combination of corner solutions (zero consumption) for some goods and interior solutions (strictly positive consumption) for other goods. The singularity imposed by the "adding-up" constraint in the KT approach and differencing the indirect utilities with respect to one of the goods

[11] This modeling framework can also accommodate financial budget constraints to avail of these media choices. However, this data do not contain price information. Therefore, I restrict this analysis and specification to be subject only to time constraints. This will be elaborated on in more detail at a later point.

(the outside good in this case), generates (K-1) interdependent stochastic first-order conditions.

The first-order conditions for the optimal time allocations (the t_k^* values) are given by:

$$\left[\exp(\beta' z_k + \varepsilon_k)\right] \left(\frac{t_k^*}{\gamma_k} + 1\right)^{\alpha_k - 1} - \lambda = 0, \text{ if } t_k^* > 0, \; k = 1, 2, \ldots, K$$

$$\left[\exp(\beta' z_k + \varepsilon_k)\right] \left(\frac{t_k^*}{\gamma_k} + 1\right)^{\alpha_k - 1} - \lambda < 0, \text{ if } t_k^* = 0, \; k = 1, 2, \ldots, K \qquad (6)$$

The optimal demand satisfies the conditions in Equation 6 plus the budget constraint $\sum_{k=1}^{K} t_k^* = T$.

I specify an extreme value distribution for ε_k and assume that ε_k is independent of z_k (k = 1, 2, …, K). The ε_k's are also assumed to be independently distributed across media alternatives with a scale parameter of σ normalized to 1. Let V_k be defined as follows:

$$V_k = \beta' z_k + (\alpha_k - 1) \ln\left(t_k^* + 1\right) \; (k = 1, 2, 3, \ldots, K) \qquad (7)$$

For the first good, the KT condition may then be written as:

$$\lambda = \exp(\beta' z_1 + \varepsilon_1) \alpha_1 \left(t_1^* + 1\right)^{\alpha_1 - 1} \qquad (8)$$

Substituting for λ from above into Equation 6 for the other activity purposes (k = 2,…, K), and taking logarithms, we can rewrite the KT conditions as:

$$V_k + \varepsilon_k = V_1 + \varepsilon_1 \text{ if } t_k^* > 0 \; (k = 2, 3, \ldots, K)$$

$$V_k + \varepsilon_k < V_1 + \varepsilon_1 \text{ if } t_k^* = 0 \;\; (k = 2, 3, \ldots, K) \tag{9}$$

To complete the model structure, I specify a standard extreme value distribution for ε_j and assume that ε_j is independent of x_j (j = 1, 2, …, K) . The ε_j's are also assumed to be independently distributed across alternatives. From Equation 9, the probability that the individual participates in M of the K activity purposes ($M \geq 2$), given ε_1, is:

$$\begin{aligned}
&P(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0) = | J | \int_{\varepsilon_1=-\infty}^{+\infty} \int_{\varepsilon_{M+1}=-\infty}^{V_1 - V_{M+1} + \varepsilon_1} \int_{\varepsilon_{M+2}=-\infty}^{V_1 - V_{M+2} + \varepsilon_1} \cdots \int_{\varepsilon_{K-1}=-\infty}^{V_1 - V_{K-1} + \varepsilon_1} \int_{\varepsilon_K=-\infty}^{V_1 - V_K + \varepsilon_1} \\
&f(\varepsilon_1, V_1 - V_2 + \varepsilon_1, V_1 - V_3 + \varepsilon_1, ..., V_1 - V_M + \varepsilon_1, \varepsilon_{M+1}, \varepsilon_{M+2}, ..., \varepsilon_{K-1}, \varepsilon_K) \\
&d\varepsilon_K d\varepsilon_{K-1} ... d\varepsilon_{M+2} d\varepsilon_{M+1} d\varepsilon_1 \,,
\end{aligned} \tag{10}$$

where J is the Jacobian whose elements are given by (see Bhat 2005):

$$J_{ih} = \frac{\partial [V_1 - V_{i+1} + \varepsilon_1]}{\partial e_{h+1}^*} = \frac{\partial [V_1 - V_{i+1}]}{\partial e_{h+1}^*}; \; i, h = 1, 2, \ldots, M - 1. \tag{11}$$

Substituting the extreme value density and distribution functions one can uncondition out ε1 from Equation 10 to obtain the following unconditional probability expression: The probability that the individual allocates some time to the first M of the K goods ($M \geq 1$) is given by the following closed-form expression:

$$\begin{aligned}
&P\left(t_2^*, t_3^*, ..., t_M^*, 0, 0, ..., 0\right) \\
&= \left[\prod_{i=1}^{M} c_i\right] \left[\sum_{i=1}^{M} \frac{1}{c_i}\right] \left[\frac{\prod_{i=1}^{M} e^{V_i}}{\left(\sum_{k=1}^{K} e^{V_k}\right)^M}\right] (M-1)!
\end{aligned} \tag{12}$$

where $f_i = \left(\frac{1-\alpha_i}{t_i^* + \gamma_i}\right)$. In the case when M = 1 (*i.e.*, only one alternative is chosen), there are no satiation effects (α_k =1 for all k) and the Jacobian term drops out (that is, the continuous component drops out, because all time is allocated to media option 1). Then, the model in Equation 8 collapses to the standard MNL model. Thus, the main model is a multiple discrete-continuous extension of the standard MNL model[12].

In addition, when an outside good is present, for identification purposes let $\psi(t_1, \varepsilon_1) = e^{\varepsilon_1}$. Then, the utility functional form needs to be modified as follows:

$$U(\mathbf{t}) = \frac{1}{\alpha_1}\exp(\varepsilon_1)t_1^{\alpha_1} + \sum_{k=2}^{K}\frac{\gamma_k}{\alpha_k}\exp(\beta' z_k + \varepsilon_k)\left\{\left(\frac{t_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right\} \qquad (13)$$

Note that there is no translation parameter γ_1 for the first good, because the first good is always consumed[13]. The identification considerations discussed for the “no-outside-good” case carries over readily to the case with an outside good.

2.3.2 UNOBSERVED HETEROENEITY

Examining the impact of consumer heterogeneity on a consumer’s media choice is useful in that it gives insight into the degree of competition between media and the extent to which media substitutability might vary across consumers. First of all, heterogeneity in consumer preferences and differing sensitivities to marketing actions provide the basis for segmentation and targeted communication programs (Allenby and

---

[12] Note that when α_k = 1 for all k, $V_k = \beta' z_k - \ln p_k$. Even if M = 1, when Equation 9 collapses to the MNL form, the scale σ is estimable as long as the utility takes the functional form $V_k = \beta' z_k - \ln p_k$ and there is price variation across goods. This is because the scale is the inverse of the coefficient of the ln p_k term (see Hanemann, 1984).

[13] This added feature of this model allows us to readily accommodate media consumption along with other non-media activities.

Rossi 1999). In the grocery context, it has been shown that failure to account for demographic variables such as income and household size (Gupta and Chintagunta 1994) leads to biased estimation and characterization for segment membership. For instance, they find that low-income households tend to be more sensitive to price and promotion, whereas larger households prefer larger brand sizes.

In practice, armed with market research, brand managers provide media planners with a set of consumer characteristics that best describe their target consumers. Media planners then purchase media slots based on media-specific audience-metrics for the target group. Advertising creatives are then developed taking into account consumer characteristics, advertising responsiveness and programming available to consumers in the chosen time-slots in the respective media. Consistent with practice, I take the perspective of the media planner. I accommodate observed consumer heterogeneity via consumer demographics on age, gender, household income, household size, and location (urban versus rural area), and unobserved heterogeneity via the heteroscedasticity and correlations in the error structure.

Specifically, I investigate unobserved heterogeneity in baseline utility as well as in the satiation parameters. I include random coefficients for each baseline preference constants, and introduce one common error component among all media options to generate heteroscedasticity and covariance in unobserved factors across activity types[14]. This means that consumers may follow a certain common mechanism in evaluating media options relative to the outside good. The error term ε_j may be partitioned into two independent components: ζ_j and $\eta' w_j$. The first component, ζ_j, is assumed to be

[14] I consider other error component specifications such as a common error component for all traditional media choices, or alternative combinations of the media choices, etc. However, the results are not significantly different.

independently and identically standard Gumbel distributed across alternatives. The second component, $\eta' w_j$, allows the estimation of distinct scale (variance) parameters for the error terms across alternatives. w_j is a column vector of dimension K with each row representing an alternative. The vector η (of dimension K) is specified to have independent, normally distributed and mean-zero elements, each element having a variance of ω_j[152]. Let ω be a vector of true parameters characterizing the variance-covariance matrix of the multivariate normal distribution of η. The result of this specification is a covariance of among alternatives in group h. Therefore, equation 12 can then be rewritten as:

$$P\left(t_2^*, t_3^*, \ldots, t_M^*, 0, 0, \ldots, 0\right) \\ = \int_{\eta} \left[\prod_{i=1}^{M} c_i\right]\left[\sum_{i=1}^{M} \frac{1}{c_i}\right]\left[\frac{\prod_{i=1}^{M} e^{(V_i + \eta' w_i)}}{\left(\sum_{k=1}^{K} e^{(V_i + \eta' w_i)}\right)^M}\right](M-1)!\, dF(\eta \mid \omega) \qquad (14)$$

where F is the multivariate cumulative normal distribution. I extend the model in Equation 14 in a conceptually straightforward manner to also include random coefficients on the independent variables z_k, and random coefficients in the α_k satiation parameters.

2.4 DATA

In this section, I describe the dataset and the current industry practice for targeting. I illustrate the extent to which media multiplexing occurs (and related switching, time effects and consumer heterogeneity) and discuss the practical implications of the proposed model.

---

[15] Other distributions may also be used for η. Note that the distribution of η can arise from an error components structure or a random coefficients structure or a combination of the two, similar to the case of the usual mixed logit model

2.4.1 MEDIA CONSUMPTION DATA

The data is from Universal McCann's Media in Mind Diary and consists of a panel of 1,775 individuals who were asked to report their media activities -- i.e., computer (including Internet), television, radio, print (newspapers and magazines), or other social activities such as having dinner, exercising, sleeping, etc., which I refer to as the "outside" option[16] in 2006. This media diary is conducted annually with a randomly-selected, nationwide sample in the United States, and is considered the largest survey on consumer media consumption conducted by any media agency.

The timing intervals in the survey are defined by half-hour time slots. Thus, at any given time a panelist could consume one or a combination of these alternatives. I use a randomly chosen set of 1,500 individuals for estimation and a hold-out sample of the remaining 275 respondents. For each respondent, I also have select demographic information, including age, gender, household income, household size, and location information such as whether the respondent is from an urban or rural area. Furthermore, adoption of media technologies is examined for each respondent. Specifically, the variables that I am interested in are home computer ownership (83.7 percent), workplace computer availability (44.7 percent), and cable television subscription (64.9 percent). Adoption for television sets and radio devices are virtually 100 percent and thus not included in the analysis.

Respondents report their activities and attention levels for each media channel every half hour for seven consecutive days, except for the time periods from 1AM - 3AM and 3AM-5AM, which are each recorded as two individual observations. I aggregate the

[16] This data do not contain information on media content. Furthermore, the geographic location of the respondents and date of survey administration is not recorded by this data provider. This prevents me from fusing programming schedule data from other sources such as Nielsen EDI with this media diary.

data up to the one-hour level to allow for more flexible modeling of continuous components rather than artificially imposed half-hour intervals. Thus, over the course of the week, there are 22*7=154 time slots. The dependent variables are defined as the time spent on each media channel at each time slot, the total of which equals the time constraint[17]. Therefore, the complete dataset contains 1,500*154=231,000 observations. Media consumption occurs 40.7 percent of the time with single media consumption occurring 32.0 percent. Thus, media multiplexing accounts for 8.7 percent of the observations (n=19,994). A large share (40.2 percent) of this multiplexing occurs during the 7AM-9AM and 7PM-9PM time slots, which are prime advertising spots for television and radio. This underscores the importance and relevance of accommodating media-multiplexing in audience-targeting, media-choice, and resource-allocation issues. Figure 2 displays the summary statistics for both single- and multi-media consumption. This figure suggests that while most of the sample engages in single-media consumption – i.e., computer 27.9 percent and radio 21.5 percent -- with television being the highest (44.6 percent), and nearly one out of five consumers in the sample multiplex media. Moreover, some of the multiplexing options are more popular than single media options. For example, co-consumption for computer and television (2.1 percent) and for computer and radio (2.1 percent) are higher than consumption for print media alone (1.9 percent). Table 1 of the Appendix displays additional detail on individual and joint media consumption by day of the week.

[17] In the case of multiplexing, the total time will be treated as t =60 minutes for each alternative chosen, as I do not observe the exact amount of time allocation to each media within the half-hour slot that is reported.

2.4.2 MEDIA CHANNEL SWITCHING

Table 1 displays the switching matrix for the average time movement across media. Each cell denotes the number and percentage share of occurrences of switching from the media channel in a row to the media channel in the column in any consecutive time slot. For example, among those who used the computer in the past period, 18 percent remained with the computer channel in the next time period, while eight percent switched to television, print, or an outside option and seven percent switched to radio. In general, I observe that consumers are most likely to switch to television from radio (14 percent), print (16 percent), or an outside option (14 percent). The diagonal elements of these tables indicate the extent to which state-dependence occurs; specifically, television exhibits the highest degree of state-dependence (22 percent), followed by radio (21 percent), computer (18 percent) and print media (nine percent).

I also find that on weekends (Saturday and Sunday), 87 percent of the consumers in the sample are more likely to switch away from media altogether to an outside option, whereas during the week (Monday through Friday), only 56 percent of consumers will switch away entirely from media. While I do not observe noticeable differences in state-dependence and inter-media switching between males and females, I do observe that younger audiences (ages 18-34) have more state-dependence for computers (20 percent) than older audiences (ages 65-75, 12 percent), and are more likely to switch *to* this medium from television, radio, print, or an outside option (9.25 percent on average) than older audiences (5.25 percent on average).

2.4.3 TIME-VARYING PROPERTY

The current industry practice is to rely on aggregate-time-series summaries by media or by consumer segments to identify prime advertising slots. As an example, in this context, the number of consumers who consume a single (or more) media are often aggregated across households and perhaps even across time slots as in Figure 1 of the Appendix. As shown in Figure 1, while the patterns of media multiplexing follow a saddle shape distribution with 9AM and 9PM being the peaks, the compositions of media activities also vary with time. If consumers are single-media consumers, a frequency-based approach to targeting is both attractive and practical. However, with this approach, the firm cannot determine whether the observed media-multiplexing in this Figure is emanating from one or multiple individuals. Moreover, this approach potentially creates exposure inefficiencies. For example, if a target consumer is consuming television and the Internet at the same time, then she is likely to receive advertising messages for the same product across two channels[18]. If the two media channels are synergistic, this approach may yield higher returns to advertising spend than if the same consumer were to receive the ad across different media at different times or were simply limited to one media channel alone. However, if advertising repetition results in lowering consumer interest (Craig et al., 1976), then such an approach may reduce ROI. Obviously, this problem can potentially worsen or improve as the proportion of multiplexers in the exposed audience increases.

[18] This is often the case, since advertising agencies are organized around media verticals, i.e. print, television, radio and interactive/digital. Each division is tasked with generating media planning and purchasing for the same client, with little to no coordination with other media teams when it comes to media planning and scheduling.

2.4.4 CONSUMER HETEROGENEITY

To a limited extent, the media switching matrix and time-varying consumption patterns provide some insights on state dependence and consumer heterogeneity. Since Dube (2004) notes that state dependence and consumer heterogeneity cannot be identified - although Dube et al. (2009) suggest some conditions (i.e., long time frame such as multiple months and considerable price variation) under which it might be possible - these conditions are not applicable in this setting. Therefore, I focus the discussion on consumer heterogeneity, from both observed demographics and unobserved random components, to sharpen the estimates for the media targeting exercise. By disaggregating at the household level, I am able to gain a more granular (and informative) view of multiplexing behavior. As an example, Figure 2 of the Appendix plots the average share of time that each household spends across media over the course of a 24-hour day and evidences considerable systematic variation that needs to be accounted for in media-allocation decisions.

I model consumer heterogeneity via the addition of four parameter estimates that capture the random effects (or more specifically, random coefficients) and error components in the baseline and satiation parameters. I include observable consumer demographics as well as self-reported variation in terms of technology availability (i.e., media penetration) and individual media attention levels (e.g., an exogenous “attractiveness” measure). These measures are consistent with the recognition that consumers may access and evaluate emerging media technologies in systematically different ways and these factors are critical to consider in a media consumption context.

To summarize, the empirical dataset in this essay suggests that media multiplexing is prevalent, and a model that examines media switching, time-varying properties, consumer heterogeneity and media specific factors is necessary for understanding this phenomenon.

2.5 RESULTS

In this section, I describe the model covariates and contrast goodness-of-fit measures in the proposed main model and its out-of-sample predictive ability.

2.5.1 MODEL COVARIATES AND CONSTANTS

I included several variables in order to obtain a rich understanding of consumers' baseline marginal utility with respect to their media-choice decision as well as their time-usage behavior. Throughout the analyses, I used a model specification with media-specific constants and select covariates in the baseline marginal utility and for the media-specific satiation parameters[19]. Covariates include contextual variables such as: (1) time-of-day dummies (12AM is the base time slot), (2) day-of-week dummies (Saturday is the base day-of-week), (3) socio-demographics (household size, age, income, gender, location of residence, etc.) and (4) stated media attention level (respondents indicate as high, medium or low for each media that is being consumed, and I further convert this to a 3-point scale with 1=low, 2=medium and 3=high attention) and (5) stated availability of media technologies: personal computer at home, at work and cable television service.

I now discuss the estimation results of the proposed demand model. Then I contrast the results with an alternative demand specification, i.e. a single discrete-continuous demand model (the base specification). The benchmark model is a Haneman

[19] An exciting direction for future research would be to add the media content being consumed in each channel so as to decompose the content into underlying attributes and attribute levels to include in the baseline utilities.

(1984) model, which models a single category discrete continuous demand function. This model has been further developed by Chiang (1999) and Mehta et al. (2010). Specifically, I contrast results from the demand models using the parameterized form of the baseline marginal utility parameters, as this allows me to show the equivalence between models in a straightforward manner.

2.5.2 MODEL SELECTION AND GOODNESS-OF-FIT MEASURES

For model selection, I examine the performance of the Hanemann model relative to the proposed multicategory-choice model along both in-sample and out-of-sample statistical criteria.

The pseudo log-likelihood for the main model with the aforementioned covariates in the baseline utility and the satiation/translation parameters as media-specific constants is -4,564.68. The log-likelihood value at convergence of the Hanemann model with the main model's baseline utility covariates as explanatory variables is -24,573.97. I rely on Bayesian Information Criterion (BIC) for model selection. The BIC for the main model is 11,204.19, with 168 parameters. The BICs for the Hanemann model is 51,291.69 with the same number of parameters. Thus, the proposed model outperforms the Hanemann model for single category discrete-continuous demand, underscoring the importance of accounting for multiplexing behavior in media choice.

In-sample assessment. Since the research objective is forecasting consumers' media choice, I contrast the in-sample hit rates for both models. These results are presented in Table 2. Interestingly, the proposed model outperforms the benchmark model even when single media are consumed. The benchmark model predicts 50.8 percent for computer, 31.7 percent for television, 29.5 percent for radio, and 18.7 percent for print. In contrast, single-media predictions from the proposed model are 99.1 percent for computer, 99.2 percent for television, 97.6 percent for radio, and 86.7 percent for print, underscoring the superiority of the proposed demand-model from a goodness-of-fit standpoint. Furthermore, the main model produces excellent predictive ability for multiplexing choices that would not be captured by simple single discrete-continuous demand framework. To summarize, the proposed model leads to a hit rate of 97.0 percent as compared to 60.3 percent from a Hanemann model. This hit rate is higher than those in previous research (e.g. around 80 percent in Hansen et al. (2006), 89 percent in Mehta (2007), 78 percent in Anslie and Rossi (1998), 77 percent in Chung and Rao (2003), 84-94 percent Schweidel et al. (2011)), and more closely approximates that the rate observed in Manchanda et al. (1999) (i.e., 99 percent). It is worth noting that Manchanda et al. (1999) took a model fitting approach, whereas my structural approach provides some theoretical insight while robustly achieving similar hit rates in modeling multi-category choices. As a robustness check, I further select three subsamples consisting of 5000 random observations each. Table 2C reports the hit rates, which are are stable and high across all three subsamples (around 95.0%-97.8%).

Out-of-sample assessment. Two-hundred and seventy-five individuals that were held out from the estimation sample are used for model validation. I report the out-of-sample hit rates for the proposed and benchmark models in Table 4A. Again, the main model (95.9 percent) outperforms the benchmark models (60.1 percent) overall and for every possible single media choice. Table 4B displays the proposed model performance for all possible media interrelationships.

Given the superiority of the proposed model using in-sample and out-of-sample fit criteria, I now consider the baseline-preference constants and satiation-parameter estimates. In the interest of space, I do not examine or contrast the parameter estimates for the benchmark model directly (these are provided in Table 2 of the Appendix). In a subsequent section, I will assess the managerial significance of these differences (including the complementary and substitutionary roles for the media) and implications for targeting and media planning via a counterfactual experiment.

2.5.3 BASELINE-PREFERENCE CONSTANTS

The estimated baseline-preference constants and the corresponding t-statistics are presented in Table 3 for time of day and day of week, attention levels, media penetration, demographics and satiation. I treat any activity that does not involve consumption of any integrated marketing media as the base alternative; these essentially capture generic tendencies to participate in each media channel relative to the outside good. Since the outside alternative is considered more often than the inside alternatives, all the baseline-preference constants are negative. The baseline-preference constant for the print media option is more negative than other media, indicating the lower participation level of households in this media activity relative to the outside good than other media activities.

2.5.4 TIME-OF-DAY AND DAY-OF-WEEK EFFECTS

The estimates reveal significant baseline-utility differences across the time of day and day of week for each media alternative. Ceteris paribus, relative to the outside alternative, the baseline utility for print media is highest on Monday (.341, $p<.01$), suggesting that the media planner is likely to gain greater audience via print than other media alternatives on Monday. The proposed model could easily accommodate the details of which magazines and newspapers were consumed by the panelists and could generate recommendations for one-to-one targeting (i.e. for a specific household, which media within a given time slot; and within that media, which sub-media).

Let us now consider the day-of-week and time-of-day effects. Within a single media, computer for example - ceteris paribus the baseline utility is highest on Monday (.411) than any other day of the week (the range is from -.012 to .305). Similarly, the baseline utility for computer peaks during the morning (7AM-8AM) and the evening (8PM-10PM), and similar patterns are found for radio (7AM-9AM and 4PM-7PM) which are more likely to be consumed when people commute to and from work. Television peaks during prime time and this finding lends prima facie support to the industry practice of pricing prime-time advertising slots much higher than other time spots.

2.5.5 EFFECTS OF HOUSEHOLD AND INDIVIDUAL SOCIO-DEMOGRAPHICS ON BASELINE UTILITY

To my surprise, within the household socio-demographic variables, there was very little difference in baseline utilities for the computer across age groups (.01 for those less than 35 and over 65). Instead, I observe that the baseline utility increases with age for television (-.103 for those less than 35 and .117 for those above 65) and print (-.106 and 0.305, respectively). Low-income individuals derive the lowest baseline utility from reading newspapers and magazines (-.010 relative to any other media or social activity). Smaller households have lower utilities for all media channels. Relative to men, women experience a smaller baseline utility from television and print consumption (-.071 and -.059, respectively) and higher utility from either computer (.031) or radio (-.056) consumption. Relative to consumers in rural markets, urban dwellers derive lower utility from computer (-.093) and television (-.020) and higher utility from radio (.049) and print (.101). While this may seem counterintuitive, it can be rationalized as follows. If one were to surmise that urban dwellers are likely to derive higher utility from other social activities than rural dwellers[20], the baseline parameters for individual media, which are all relative to the outside good, can be higher (and more positive) for rural dwellers than urban dwellers.

2.5.6 EFFECTS OF HOUSEHOLD ATTENTION LEVEL AND MEDIA PENETRATION ON BASELINE UTILITY

As expected, higher household attention for a focal media is associated with a higher level of utility for the same media. Ceteris paribus, holding the same attention level across media alternatives, the baseline utility for computer (7.740) is the highest relative to other media alternatives (2.089 for television, 2.212 for radio, and 3.599 for print). Historically, these attention levels would imply a 100 percent share of that media channel. However, by estimating the cross-media effects of attention on a media-specific baseline utility, I observe that these strong attention levels can enhance or weaken the utilities for alternative media. For example, while computer has the highest own attention-level effect, the attention level for computer has a negative impact on the baseline utility for other media, e.g., -.335 for television, -.421 for radio, and -.239 for print, suggesting that consumers who focus on computer media may be less likely to multiplex. In other words, for these consumers, non-computer media options might act as substitutes. In a subsequent section, I will examine the marginal rates of substitution to examine this possibility more rigorously. In contrast, consumers who have high levels of

[20] In part this is due to the sheer difference in options to which they have access to, such as operas, theaters, shopping centers, parks, etc.

attention for television (2.089) are likely to multiplex print (.104), but unlikely to multiplex computer (-.05) and radio (-.095) media. Thus, for these consumers, utility for print and computer media are likely also increased.

Higher attention levels for all media alternatives negatively impacts the baseline utility for computer, with the highest degradation from attention levels for print (-1.410). However, this substitution pattern is not found for the other three media alternatives. In fact, high attention levels for print media benefit the baseline utility for television (.469) and radio (.651). It is not uncommon that I observe people read newspapers when watching television or listening to radio. The allocation of attention suggests that the consumer may intrinsically be attracted to multiple media channels, and the joint consumption is not necessarily due to satiation only.

I consider the role of media availability by examining media penetration in the sample[21]. Respondents who have subscribed to cable television service derive more negative baseline utilities for computer consumption (-.063), but positive utilities for television (.078), radio (.004) and print media (.07). In contrast, consumers who own a personal computer have higher utilities for computer media (.043) and print media consumption (.178), but negative utilities for television (-.084) and radio (.295). In other words, consumers with a computer at home are more likely to have higher utility for print media. In addition, if respondents use computers at work, then they experience more negative baseline utilities for all other media (e.g., -.10 for television, -.00 for radio, and -.05 for print).

[21] I acknowledge that this is a data-driven approach, and better data would be preferable. HoIver, this approach is not inconsistent with the grocery choice literature (Manchanda et al. 1999; Kim and Allenby 2002), which also does not examine shelf availability directly.

2.5.7 SATIATION ESTIMATES

Recall that the role of α_k is to reduce the marginal utility with the increasing consumption of media k, hence its interpretation as a satiation parameter. High α_k values indicate low satiation effect. These satiation parameters were introduced as a constant in our model specification – i.e., one constant for each media alternative. From the satiation-parameter estimates and the corresponding t-statistics displayed in the last row of Table 3, it can be observed that significant satiation effects exist in the time-investment patterns for each media. Further, it can be observed that the satiation level is lowest for computer/Internet (.999), and relatively the same for other media channels (.447-.479), indicating that consumers appear to be satiated more quickly by traditional media than by new media.

2.5.8 UNOBSERVED HETEROGENEITY

I investigate unobserved heterogeneity in the baseline utility as well as in the satiation parameters by including random coefficients for a baseline preference constant and introducing one common error component among all media options to generate heteroscedasticity and covariance in unobserved factors across activity types. The high standard deviations of the random coefficients for computer (.209 for the baseline utility and .298 for the satiation parameter) and print media (.665 and .208, respectively) indicate higher variance due to unobserved factors for computer and print, compared to radio and television. However, the common error components among media options do not seem significant.

2.5.9 ROBUSTNESS OF THE TIME PERIOD SPECIFICATION

As noted earlier, I used a one-hour time period specification, but this choice is only a pragmatic starting point. The data was collected in half-hour slots and could just as easily be aggregated to two hour time periods (which is what I did between the hours of 1AM to 5AM). To assess the robustness of the proposed specification, I estimated the proposed model using half-hour time slots and observe that the parameter estimates are robust, both in direction and significance. These estimates are displayed in Table 3 of the Appendix. Together, this suggests that the proposed model is able to process and predict the continuous components of choice in a fairly flexible manner than what is dictated by the dataset structure.

2.6 MEDIA AS COMPLEMENTS AND SUBSTITUTES

Having observed that attention levels and media penetration (i.e., availability) can systematically impact the consumer's utilities across media, I now examine potential complementarity and substitution interdependencies among the channels. My goal in this section is to empirically describe some (in)consistencies with marketplace wisdom; I do not claim to provide an explanation (i.e., the "why") for observed patterns. This I leave to future research.

Recall from the discussions in section 2.3 that ψ_k is the marginal utility at the point of zero consumption. It can be inferred by computing the marginal utility of consumption with respect to media *k*, which is:

$$\frac{\partial U(t)}{\partial t_k} = \psi_k \left(\frac{t_k}{\gamma_k} + 1 \right)^{\alpha_k - 1} \tag{15}$$

Alternatively, the marginal rate of substitution between any two media k and l at the point of zero consumption of both media is given by $\frac{\psi_k}{\psi_l}$. For two alternatives i and j, a higher baseline marginal utility for media i relative to media j implies that an individual will increase overall utility more by consuming media i rather than j at the point of no consumption of any media option. In other words, the consumer would be willing to give up media j in exchange for consuming more of media i. Thus, a higher baseline ψ_k implies a reduced likelihood of a corner solution for media k.

Using the recovered parameter estimates, I can compute the own and cross-marginal effects to examine individual media substitution patterns across the choice alternatives. These are presented in Table 5, which uses only the media-specific intercepts in the baseline utilities of each media alternative; this represents consumers' intrinsic preference for switching from a media option in the column to an alternative on the row from the point of zero consumption. Hence, on average, consumers would gain more utility from the television than from computer (4.289), radio (3.126) or print (16.268). In contrast, print media is easily substituted for all other media (.061, .264). These rates of substitution can be further generalized to a specific day of the week and time of the day, or a specific demographic. For example, for a certain group of media enthusiasts, the substitution patterns between joint consumption and single consumption would show the incremental benefits of multiplexing over consuming media separately and thereby revealing potential synergies across media channels. It is worth noting that this is the conventional approach to understanding such interdependencies (recall Figure 1).

I am fortunate to also obtain media attention measures and penetration. Hence, I decompose the aggregate marginal rates of substitution by computing the individual effect of a consumer's attention level and media availability on the complements and substitute roles, which I believe to be not only more nuanced measures of media usage but also contain more face validity than a solely demographics perspective.

I calculate the marginal rate of substitution for each individual media and all possible combinations at the point of zero consumption first for the baseline constants, and then based on attention for media (i.e., computer, television, radio, and print) and media penetration (i.e., cable television ownership, pc ownership, and workplace pc use). In the interest of space, the full tables of estimates for each of these aspects are displayed in the Appendix, Tables 4B-4H.

Substitutes versus complements. When Table 5 is expanded to consider all possible media combinations (Table 4A in the Appendix), the results suggest that television is the "stickiest" media; i.e., consumers gain more utility from television than from computer, radio, print, or any possible combination of these. In this vein, I observe radio to be the next most preferred, followed by the computer. These baseline constants on marginal rates of substitution also suggest that single media consumption is generally more favorable than multiplexing options.

However, if I examine these substitution patterns in light of controllable media-specific factors, such as consumer attention (Tables 4B-4E in the Appendix) and technology availability (Tables 4F-4H in the Appendix), I observe some intriguing asymmetries across media alternatives and their combinations which I will subsequently describe. In these tables, I highlight the combinations of individual and joint media options (e.g., media A and media B) that may serve as strong substitutes for another. For example, in Table 4B, I observe that for consumers with high attention to computer, the computer media option dominates (i.e., consumers are less willing to give up the computer) for all forms of alternative single and joint media combination. Furthermore, because of the high attention to the computer, other media channels are heavily impacted in that consumers are more willing to give up the computer media alternatives for any consumption choices that involve the computer option. For example, for these same consumers, television is substitutable for computer, print, and the combinations of: computer and television, computer and radio, computer and print, computer and television and radio, computer and television and print, computer and radio and print, and computer and television and radio and print.

In the tables I also italicize the instances in which media alternatives might be viewed as complements. Specifically, a high marginal rate of substitution of a multiplexing media option (for example, media AB) with a single media option (A, which is a subset of the joint consumption) suggests that consumers are willing to give up more single consumption of media A for joint consumption with B. In that case, I conclude that media A and B are being viewed as complements. Tables 4B-4H in the Appendix highlight the instances where I pin down such interdependencies and how these individual media are used together. For example, I observe that while computer, radio, and print may act as substitutes, so also do all possible combinations of these alternative media (i.e., computer and radio; computer and print; radio and print; and computer, radio and print) alone. However, when coupled with television, these media act as complements (i.e., computer and television; television and radio; television and print;

computer, television and radio; computer, television and print; television, radio and print; and computer, television, radio and print). In this manner, one can assess the extent to which all media and combinations thereof might serve as substitutes or complements to a focal media choice. I now consider some broad results and insights.

Attention for media. In general, I find that while it is true that consumers with high attention for a particular media are less likely to give up that media for substitutes, I do find that the various media are differentially prone to multiplexing with alternative media forms. For example, consumers with high attention for computer are unlikely to multiplex traditional forms of media (see the lower diagonal of Table 4B in the Appendix). However, the reverse is not true; those who have high attention for traditional media (i.e., television, radio, and print; tables 4C-E) are more prone to multiplex not only the consumption of new media such as the computer, but also consumption of all other forms of media, i.e., television, radio, and print. This is consistent with marketplace findings, published in early 2009, which suggest that television has strong spillover effects with computer (i.e., digital media) consumption (http://adage.com/print/134790).

Media availability. I incorporate the data on media penetration to speak to multiplexing behavior in light of media availability. For example, I find consumers who own PCs at home are less likely to give up print media (see the print column of Table 4G in the Appendix). The individuals in the sample who consumed computer media tended to be females over 35, with higher household incomes than television viewers or radio listeners. The individuals in the data who consume print media also tend to trend higher in income than consumers of all other forms of media. This income differential is consistent with marketplace reports, which indicates that computer media consumers tend to have higher incomes (The Pew Internet Project, 2010)[22]; this could be partially due to the fact that there is a higher penetration of broadband connections among these households than lower income households[23].

In contrast, consumers who have cable television at home would easily give up computer media consumption relative to all other forms of media (see the computer column of Table 4G in the Appendix). This is consistent with the demographics of the sample, which indicate that television watchers tend to be individuals who are older (i.e., over 65) males, with household incomes less than $35,000. According to recent statistics published by the U.S. Census and Nielsen, this is also the demographic for television viewers in the U.S. population; consumers over 65 watch the most television per month (198.3 hours) and 30.4 percent of television viewers earn less than $30,000 annually (TV Dimensions Study, 2010)[24].

The data also suggest intriguing location differences with respect to media consumption. For example, it is well known that thepresence of a PC at work facilitates employees' personal or non-work activity. In fact, 2008 estimates suggest that workers spend as much as 25 percent of their work time engaged in personal activity (Cheng, 2008)[25]. To this end, I observe that consumers who have workplace PCs are also more likely to give up print media (see the print column of Table 4H in the Appendix) and less likely to give up computer media. This would be consistent with the hypothesis that

---

22 http://www.marketingcharts.com/television/Ialthier-use-internet-differently-15178
23 http://www.marketingcharts.com/direct/higher-income-equals-higher-tech-usage-rates-15160
24 http://www.sterlingsatellite.com/info/television-vieIrship-numbers.html
25 http://arstechnica.com/old/content/2008/09/report-workers-spend-25-of-work-time-goofing-around-online.ars

workers tend to consume news at work instead of home. This is in contrast to individuals who own PCs at home (Table 4G in the Appendix), yet are unlikely to give up print.

Together, these results suggest the possibility that print media may be in the most tenuous position in terms of media consumption. In other words, as new forms of media and increased multiplexing emerge, it is not clear how print media consumption will be impacted. There is speculation that print media may play a reduced role in consumers' media consumption portfolio. In the past year alone, in part because of reduced subscription levels, many newspapers in the U.S. have exited the market (CNN News, March 19th, 2009)[26]. Because of these emerging realities, I now turn to a counterfactual analysis that assesses the impact of a print media exit on consumers' individual and multiplexing media consumption.

2.7 COUNTERFACTUAL ANALYSIS

Since the demand model is derived from first principles, i.e. utility-theory-consistent, I can treat the resulting parameter estimates as preference parameters, which enables me to conduct "what-if" analyses. Since consumers in the empirical context are consuming print media, I can treat the entry/exit of print as being exogenous and simulate the choices that result from shutting off the newspaper option. Thus, the counterfactual experiment can predict short-run media choices that will stem from a temporary shut-down of print media such as newspapers. This is because large structural changes in the marketplace can potentially result in changes to the recovered demand parameter pre-change, as per the Lucas critique (Lucas 1976)[27].

---

[26] This is particularly true in rural or small markets. However, since I am unable to identify such respondents in the dataset, I contrast an identifiable, albeit extremely large market, namely New York City.
[27] This counterfactual can also be viewed as consumers' response to a temporary shock like the non-delivery of newspapers.

I make the print option unavailable for respondents in Los Angeles. This is achieved by setting the time allocated to print to 0 for respondents in these two markets, predicting co-consumption of other media alternatives in conjunction with the outside good. Since the proposed model allows for state-dependency in media choices, redirecting choice for the current period impacts utility derived in the current consumption period as well as beyond.

I compare the results of current period impact of the pre- and post-policy change for Los Angeles in Table 6. I observe the impact of the policy change in Los Angeles using parameter estimates from the main model. Column 2 of the table states the total number of observations that select a specific media (displayed in the row classifier) pre-policy change. Of 679 observations in the sample, the corresponding columns that are associated with the same row denote the number of observations that now switch to the new media (column choice) - i.e. switch away from the previously chosen media. For example, 234 observations have print media alone, pre-policy change. Of these, 126 (53.8 percent) switch to an outside good, 77 (32.9 percent) switch to computer, 23 (9.8 percent) to television and eight (3.4 percent) to radio. By age groups, I observe that seniors (56.8 percent) switch to outside activities more than young adults (41.2 percent). Within the media choice set, young adults (ages 18-35) are more likely to switch to computer (44.4 percent), as compared to the population average (32.9 percent).

Based on these results, it seems that the majority of print media audiences will be lost to outside activities, but for those who remain in the media consumption, most people will switch to computer when print media is unavailable to them. Furthermore, a traditional single-category Hanemann type model would not reflect the 445 previously multiplexing consumers (which are almost twice as large an audience as single print media consumers), or would falsely treat them as separate observations for single media. In contrast, the model enables media planners to tap into these consumer segments and predict the traffic stream to each and every media outlet and combination. While the qualitative substitution patterns remain very similar across the two markets, the magnitude of switching and, hence, the resulting switching elasticities reveal some degree of heterogeneity in the effects of such a media landscape change.

In the New York City market based on 999 observations (Table 5 of the Appendix), I observe similar results, although slightly more traffic to the computer and less to television and radio. A shutdown in print media will result in 53.7 percent of the single print media consumption switching to an outside activity, while 33.4 percent switch to computer, 7.4 percent switch to television and 5.5 percent switch to radio. Senior people are more likely to opt for outside activities (66 percent), whereas young adults migrate more to computer (37.5 percent) and radio (14.6 percent).

These changes in predicted substitution patterns within each market are important as they have direct bearing on targeting and media-buying strategies. While I cannot assess the economic significance of these differences with this dataset, consideration of the in-sample and out-of-sample predictions collectively suggest that the statistical properties of the main model and its forecasting ability substantially out-perform other commonly employed media-choice models in the IMC domain.

In summary, the model selection criteria and wealth of novel insights generated by the main model reinforce the gains from the model relative to other models currently being employed by media planners. Taken together, these results suggest that accounting for media synergies - as done in the model - is extremely important even if media-multiplexing is not critical to the media planner or is not commonly occurring in the data.

2.8 DISCUSSION

This research is a first step in investigating the critical issue of predicting consumers' media choice (or choices) and its implications for media planning and targeting. Rather than focusing on the synergistic effects within elements of a firm's marketing mix which has received considerable attention in the extant literature, I examine synergies across media channels at the individual consumer-level. Doing so allows me to complement the extant literature and address a managerially relevant question that is mission critical to the success of any IMC initiative. The empirical approach involves a proprietary database of daily diaries of media consumption and product preferences for over 1,775 panelists who report their media consumption habits over one-hour increments throughout a given day and for a whole week. The proposed model estimates suggest that consumers spend less time consuming computer or radio contents on the weekends and more time watching television and reading print media instead. Additionally, I reveal interesting (and often asymmetric) interdependencies across media channels by examining marginal rates of substitution. Attention spans, media adoption and individual differences also play a role. I also show that failure to account for media multiplexing can significantly reduce the reliability of the media planner's audience predictions.

My access to the proprietary dataset in this essay enables me to estimate a model that better predicts consumers' media choice while accounting for media synergies and switching. I demonstrate the gains from the proposed model relative to the current media choice models. A novel contribution of the proposed model also is the ability to identify the relevant alternatives to single media consumption and migration patterns in the event of an exogenous shock in the marketplace, i.e. the exit of print media.

2.8.1 LIMITATIONS

In this study I take the position of the media planner and my central research objective is trying to predict consumers' media choice. Given their focal audience's demographics, media planners can use the model-generated predictions to have more targeted media plans. However, the dataset used in this essay prohibits further discussion of advertising effectiveness and related media resource allocation decisions. Another limitation is that this dataset does not afford the ability to model switching within less than a half-hour time slot. This is not that problematic as most, if not all, traditional media contracts are negotiated at the half-hour level. Lastly, similar to the empirical approach in the variety-seeking literature, the interior solutions in our model are a result of satiation; future research is needed to model complementarity across alternatives in a more flexible manner.

2.8.2 IMPLICATIONS FOR MANAGEMENT

This research underscores the need for media planners to move away from single media focused choice models to the use of demand models that better incorporate and account for media synergies and media multiplexing. Furthermore, this study underscores the need for media planners to avail themselves of new individual-level media-choice data instead of market-level data. Commonly used market-level data mask underlying consumer- and time-varying specific media choices and limit the ability to uncover valuable sources of media interdependencies. The proposed structural demand-model can also be used by policy makers and media planners to predict the extent of switching to an outside activity or alternative media (single- or media-combination) as a result of shutting down any particular media option.

2.8.3 DIRECTIONS FOR FUTURE RESEARCH

This research provides a first step in understanding and modeling customer multiplexing behavior in multi-media environments. I encourage future research that examines interdependencies not only across more channels or platforms (for example, social media, online search, mobile activities, etc.). With the emergence of the new interactive media landscape and technologies, consumers' media preferences may evolve over time. Despite the recent literature that examines migration patterns across multiple retail channels (Ansari, Mela and Neslin, 2008), the issue of consumer-level media migration over time is completely unexplored. How does the complementarity or substitutability between online and offline channels vary over time? What factors contribute to this change? An intriguing area to investigate is the dynamics of media choice, thereby adding a fourth dimension (time) to the media-planning framework.

Finally, I encourage future research on developing models that predict the ROI on media spending after accounting for media multiplexing and the implications of these shifts for media planners' resource-allocation decisions. Together, these research questions ultimately generate insights on how to better measure, understand and manage customer engagement in a continuously changing media environment.

3 ESSAY 2: MODELING CONSUMER DECISIONS ACROSS SEEMINLY DISPARATE CATEGORIES A LATENT-TRAIT APPROACH

3.1 INTRODUCTION

"To look at a leopard through a tube, you can only see one spot."

-From Ancient Chinese Idiom (422 AD)

The premise of marketing research is built upon investigating closely related variables and decisions. Marketing strategies are planned and delivered with an integrative mix of the 4P's (Price, Product, Promotion and Place), each of which may form a hierarchy of interrelated components. Consumer decision making is a compound function of past experiences, firm communications, social influences, individual differences and other contextual factors. Understanding interactions and interdependencies among these complex relationships often yields richer and more complete insights on the value-creation process than examining a single piece or a subset of information independently.

The task for marketing managers today is increasingly complex and customer-oriented. Traditional practice involves brand managers planning and organizing marketing activities around individual brands, then shifting towards category managers who coordinate purchasing, merchandising and prices of a set of brands within a category (Zenor 1994) and occasionally across categories within the "market basket" (Bell and Lattin 1998; Seetharaman et al. 2005). Most recently, as marketing practice embraces customer orientation and customer management, managers note that consumer purchases are never just limited to a few brands, or grocery shopping basket. In fact, consumers naturally shop around a globe of disparate product categories that are more complex and diverse than the traditionally defined market basket in retailing research. Here, the term "disparate" is similar to "non-comparable" (Johnson 1984), which describes the degree to which choice alternatives can be represented by the same attributes, but offers a broader and more generalized description of categories that are utterly dissimilar and difficult to compare with each other than merely a function of the number of common and distinctive features associated with alternatives as in comparability (Tversky 1977). For example, consumers drive certain cars and listen to certain radio channels; they prefer certain soft drinks and behave in certain ways when it comes to financial investments. These categories have typically been studied in isolation, but they collectively reflect a more complete and realistic picture of consumer demand rather than steady snapshots for consumer behavior as in previous research. This research aims to examine consumer choices across seemingly disparate product categories in order to specify a fuller model of the consumer demand problem.

Insights from understating behavior across seemingly disparate categories would be increasingly relevant in today's retail context for customer valuation, targeting, cross-sell and resource allocation (Reinartz and Kumar 2003; Shah and Kumar 2008). Marketing activities are coordinated in a variety of categories and in a variety of formats. Supermarkets such as Wal-Mart make assortment decisions for product categories that are not closely related, including consumer electronics, furniture, apparel, grocery and many others. The rewards from loyalty programs such as Air Miles can be accumulated or redeemed in many outlets, ranging from gasoline services and package holidays to supermarket shopping. Brand extension efforts make Virgin Group a conglomerate that builds presence across different business areas. Moreover, due to the growing ability to track consumer purchase patterns cross categories using CRM and web-based tools, Internet retailers (such as Amazon and Groupon) and platform providers (such as Facebook) are proactively managing across a wide assortment of categories and having access to a rich database of consumer behavior that was not able to be tracked traditionally. Managers are urged to embrace the challenge of creating a broader and richer description of customer behavior and understand the deeper underlying process of consumer decision making.

Identifying and assembling purchase patterns from individual categories can assist in segmentation and targeting. To date, behavioral-based segmentation focuses primarily on "what consumers did" rather than "why they did it". The objective of this research is to help managers get at the "why" question by studying and inferring the latent processes from observed behavioral data that are accessible to today's firms (rather than incurring the additional cost of augmenting with experiments, survey data or brain scans). The genesis is that consumers are alike because they share similar thought processes, not because they display similar observed behavior as assumed in traditional segmentation approaches.

For example, if a consumer bought a BMW coupe, traditional models may suggest selling him tires (as complements), or financial services (categorizing him as a buyer with high income), or another BMW car (based on his/her purchase history), or in the worst case, some unrelated products (from coincidences and random errors). In fact, better targeting occurs when retailers find out the true underlying decision process that the consumer employs and advertise the product that fits into his/her value: for example, a pair of Nike shoes for those who appreciate the sporty spirit, or a piece of artwork for those who value the design, or a life insurance plan for those who value control and security. Though these categories are disparate to the BMW coupe, they provide a deeper understanding of consumers' underlying decision making processes that traditional models fail to address (Manchanda et al. 1999; Ma and Seetharaman 2004). Nevertheless, one reason that previous research on cross-category behavior often restricts to related categories is due to data availability. It is difficult to get access to customer purchase data across a wide range of seemingly disparate categories and infer consumers' underlying processes based on observed behavioral patterns. The objective of this research is thus to build a theory-driven model that helps managers to understand and measure the impact of the underlying processes that explain systematic co-variations across seemingly disparate categories based on behavioral data.

In fact, the process of aligning decisions across seemingly unrelated categories occurs naturally and bilaterally. Consumers constantly make choices for every aspect of their lives, from complex decisions such as which car to purchase, in which stock to invest, and to which cell phone plan to subscribe, to more routine ones such as which soda to drink and which television channel to watch. There could be many types of underlying processes that explain co-variations across categories. One famous example in the marketplace, originally to illustrate the power of data mining (*Financial Times of London*, Feb 7th, 1997), is of "Beer and Diapers". It is observed that beer and diapers, two categories which appear to be unrelated, tend to be purchased together simultaneously by male customers. Traditional models on multi-category choice behavior would only capture this phenomenon through demographics and random errors, and fail to recognize the deeper rationale that male customers seek convenience when making shopping trips. Another example is that we may observe certain consumers tend to be "innovators" of many categories as they always prefer the latest new products or services, ranging from apparel and cell phones to automotives. We may also observe that certain consumers are more inclined to purchase or hold multiple types of products, either because of the need for variety-seeking, or because of a limited capability to reach a single decision (Dowling and Uncles 1999). In this case, people are alike not only because they coincidentally display similar observed behavior but also because they share similar latent decision processes. While traditional segmentation research attempts to group people of similar observed outcomes together and explain their behavior with a same-response coefficient, assuming that "birds of a feather flock together" (Desarbo et al. 2004; 2006; Heilman and Bowman 2002), this research provides a first step in categorizing customers as a set of value-based process parameters for theory-driven segmentation and profiling exercises.

This research contributes to the literature in three ways. First, it takes a first step in modeling a complete picture of consumer decision problems by examining consumption across seemingly disparate product categories. Second, it investigates the latent processes that govern consumer decision making across decision stages and across categories to advance our understanding in both dimensions of customer behavior: the breadth of their consumption portfolio, and the depth of their latent decision processes. Third, it provides richer insights on targeting and profiling based on continuous latent processes rather than discrete observed behavior. Specifically, I propose a hierarchical multinomial processing tree model to empirically examine the underlying processes, which are defined as the "latent traits" that govern consumer choices across five seemingly disparate product categories[28]: media consumption, automobile purchases, financial investments, soft drinks and cell phone plans through an asyndicated dataset consisting of 5,014 randomly selected consumers in the United States.

The model is estimated using Bayesian methods with weakly informative hyperprior distribution and a Gibbs sampler based on two steps of data augmentation. While the latent process structure remains the same across these categories, I further investigate how consumer behavior systematically varies from one category to another and finally suggest new approaches to segment and profile consumers based on collection of continuous latent traits (rather than discrete observed behavior) across multiple categories. Lastly, I compare the latent-trait approach with the latent-class approach and identify conditions under which they may yield in similar or dissimilar results from a data-driven perspective.

Latent trait models have a long history in psychometric studies of psychological constructs such as verbal and quantitative ability (see, e.g., Lord and Novick 1968; Langeheine and Rost 1988) but have not received much attention in the marketing field. Essentially, any person-level difficult-to-observe continuous parameters, whether well-defined or undefined, goal-oriented or heuristic-based, can all be considered as latent traits. It can take place at many levels of decision making. For example, at the product category level, need for convenience is the latent trait that explains the phenomenon of beer and diapers. It is highly likely that male consumers would exhibit the same trait when choosing brands and products, such as choosing the most accessible diaper brand on the shelf, or choosing the beer that they are most familiar with. Furthermore, decision

[28] I will explain the selection of categories in the data section.

processes can often be casted into a tree model in a natural and principled manner, and latent traits can be best viewed as the branches that lead to decision nodes at each stage. Depending on the firm's interest in key decision variables and availability of data, the tree structure can be adapted in a specific setting. For example, if managers are interested in the impact of "need for convenience" on store and assortment choices, then the tree will start from a consumer who chooses between the more "convenient" stores (i.e., stores within a certain distance) and less convenient stores, then chooses between more "convenient" assortments (e.g., shelf allocation in the case of beer and diapers). At each stage, the latent trait of "need for convenience" determines the consumer's paths in taking upper or lower decision branches. If a firm's interest lies in capturing the latent trait of "innovativeness" in category and brand management, then the decision tree will start from a consumer choosing between the newer (more innovative) and more established product category, followed by decisions in brands, and finally in products.

As noted earlier, many types of latent traits may affect consumer decision making and this research is at best offering a *process* for studying the impact of latent traits. For exposition and without loss of generality, I examine one specific type of latent trait, which is defined as "polygamy". Polygamous loyalty has been documented in the literature to describe the behavior of "divided loyalty" among a number of brands (Dowling and Uncles 1997; Bowman 2004). Polygamy is the tendency of individuals to seek *multiple* types of products, services, or brands, as opposed to holding to a *single* one. It is an idiosyncratic *trait* that a consumer has and, when manifested, it can lead to interior solutions where their constrained utility is maximized on the budget constraint with strictly positive quantities of two goods (i.e., multiple goods are chosen from the alternative set). It is noteworthy to distinguish polygamy from *variety-seeking* behavior, which can be viewed as a subset or outcome of polygamy that describes the *switching* behavior among brand/product/service alternatives, as opposed to *loyalty* (Khan et al. 1986). While consumers engage in variety-seeking activities merely as a result of satiation (Kim and Allenby 2002), they may seek polygamy for various reasons such as sensation, diversification, convenience, security, complementarity and/or inability to reach a single decision. Polygamy may take place at many levels of the decision process. For example, at the product level, investors may hold different stocks as a portfolio; at the brand level, diners may order different brands of wines at one occasion; at the product-type level, consumers may want both a laptop and a desktop; and finally, at the product-category level, consumers almost always hold multiple categories. In addition, depending on the product category, consumers are likely to experience a satiation effect or "heavy-user" effect when moving across layers of decision processes. For example, if consumers purchase multiple types of automobiles, they may be less likely to purchase multiple brands within each type. Nevertheless, for media consumers who enjoy the large variety of website choices that Internet offers, it is more likely that they will also subscribe to multiple television channels at a time. Such variations across levels of decision processes and product categories allow better identification when estimating the parameters and enrich potential insights that latent trait can generate.

In summary, by testing one specific latent trait of "polygamy", this research takes the first step to empirically investigate the *continuous* latent processes that govern consumer *behavior* across seemingly broad and disparate product categories and across different decision-making stages to advance understanding in both dimensions of customer behavior: the breadth of their consumption portfolio and the depth of their latent decision processes. Specifically, this research addresses: 1) whether latent trait has an impact on consumer decision making and the magnitude of such impact, if any; 2) how a latent trait is manifested across different levels of decision making; and 3) how the effect of a latent trait varies across seemingly disparate categories. In doing so, this research contributes to the consumer decision literature in three ways: 1) theoretically, the latent-trait approach provides rich support in examining the high level processes; 2) methodologically, the relative merits of models with continuous versus discrete representations of consumer heterogeneity are discussed; and 3) substantively, by providing new insights on targeting and profiling with respect to managing across seemingly unrelated product categories.

The remainder of this essay is organized as follows. Section 3.2 reviews related streams of literature and my general approach to modeling consumption across seemingly disparate categories. This is followed by the empirical model in Section 3.3, a description of the data (Section 3.4), and estimation and results in Section 3.5. I then conduct a latent-class segmentation analysis ex-post based on the first stage latent trait parameters in Section 3.6, and conclude with a discussion of key findings, implications for management, limitations and directions for future research in Section 3.7.

3.2 LITERATURE REVIEW

Marketing research on consumer choice across seemingly disparate categories and latent traits is scarce. Nevertheless, related literature on cross-category models of consumer choice and decision making processes has been popular. Consistent with the shift in practice, marketing research has progressed gradually towards examining the full picture of decision problems. As shown in Figure 1, the literature on cross-category behavior evolves from standard single category choice models with homogenous demand specifications and independent category decisions (McFadden 1980; Guadagni and Little 1983; Bucklin and Gupta 1992; Berry 1994) to models addressing correlations between two or three related, by and large complementary product categories (Erdem 1998; Manchanda et al. 1999; Heilman and Bowman 2002; Chung and Rao 2003), and most recently to multi-category choice models (aka market basket models) that describe purchase behavior in typically eight to ten categories within grocery shopping trips (Ainslie and Rossi 1998; Bell and Lattin 1998; Seetharaman et al. 2005; Mehta, 2007). In doing so, this stream of research uncovers the correlations in cross-category purchase outcomes and marketing mix sensitivities from complementarity, consumer heterogeneity, state dependence and coincidences. The genesis is that if sensitivity to marketing mix variables is a common consumer trait, then one should expect to see similarities in sensitivity across multiple categories (Ainslie and Rossi 1998). For example, a low-income household might be price sensitive in many product categories. However, the categories studied are usually within the grocery shopping basket and are, by nature, closely related (e.g., toothbrush and toothpaste). The reality is that consumer purchases are never limited to a grocery context and customer behavior is likely to vary systematically across product categories as a function of more than the sources of cross-category variations listed above. For example, the joint purchase of beer and diapers would have been incorrectly picked up as mere coincidences by previous research. Hence, research that examines consumption across seemingly disparate categories would provide a more realistic and generalized approach in studying cross-category behavior. In order to study behavior in such a broad and comprehensive consumption context, managers need a more sophisticated approach that describes and provides a deeper understanding on consumers' underlying preferences or processes that govern choices.

As shown in Figure 2, there are many approaches, such as attitudinal or behavioral, that one can use to study disparate categories. Decades ago, researchers typically looked at choices at an aggregate level. Attitudinal research and survey studies on consumer "Values, Attitudes, Lifestyles" (VALS, VALS2) have long been interested in addressing such problems. While this stream of research often suffers from implementation difficulties such as smaller sample sizes, greater collection efforts, and sometimes self-report bias, they provide an intriguing angle to understanding person-factors (though mostly on *attitudes* and aggregated *discrete* segments or labels) from consumers' perspectives. On the behavioral side, techniques such as grouping or conglomeration are available to analyze data from aggregate responses and decompose the tabular frequencies into a set of latent classes or segments (Desarbo et al. 1993; Wedel and Kamakura 2000). A limitation of such an approach is that it relies on brute-force statistical fits rather than a utility-maximizing framework, and therefore is less theoretically realistic (Wedel et al., 1999). Furthermore, it imposes a fixed number of latent classes and assumes each person to be a member of one latent class. This is often too restrictive and difficult to interpret. In many applications, a continuous distribution of a parameter value that accommodates heterogeneity across consumers is more realistic (Andrews, Ainslie and Currim 2002; also see Andrews, Ansari and Currim 2002).

Most recently, there is a growing interest in understanding psychological processes that contribute to decision making (McGuire 1976). Over the past thirty years, a large stream of experimental studies show that consumer decision making is a highly complex process that challenges the assumption of a well-defined preference structure (or utility function) in modeling literature (Bettman 1979). New developments in neurosciences such as CAT scans and fMRI illustrate that different parts of the brain are active during different parts of mental life (including consciousness, emotions, choices and morality) and exact brain regions can be pinned down for certain types of decision making (Hedgcock and Rao 2009; Weller et al. 2009).

Despite the critical role of high-level latent processes in consumer decision making, there is little empirical research examining its impact on consumer choice with behavioral data. Incorporating these difficult-to-observe process parameters into well-defined quantitative models requires a continuous distribution of the latent variables. This can be achieved through latent trait analysis, which has received considerable attention in psychometrics and mathematical psychology. There are a few early marketing applications discussing latent or unobserved variables in survey research (Balasubramanian and Kamakura 1989), coupon redemption (Bawa et al. 1997) and cross-selling of financial services (Kamakura et al. 1991) in a single category context. Operationally, latent trait is the “person parameter” that has been defined in item response theory. It represents the strength of an attitude and captures parameter heterogeneity due to individual differences between persons, as opposed to parameter homogeneity in latent class approach. It has two unique advantages over traditional models: to the extent that marketing is applied psychology and applied sociology, the latent trait approach is more theoretically grounded by investigating the underlying decision process that impacts consumer choice; and empirically, a continuous distribution of person parameters usually leads to better fit.

3.3 The Empirical Model

In this research, I adopt a hierarchical multinomial processing tree model with Bayesian methods to examine the impact of polygamy on consumer choice while incorporating heterogeneity. Multinomial processing tree (MPT) models have been extensively used in cognitive psychology for memory testing, perception research and reasoning (see an overview by Batchelder and Riefer (1999)). MPT models are discrete choice models that are developed exclusively to explicitly measure and disentangle the impact of underlying or latent cognitive capacities with panel data resulting from multiple and confounded processes (Ansari, Vanhuele and Zemborain 2007). The "structural" parameters represent underlying psychological processes. Each MPT model is a re-parameterization of the decision outcome probabilities of the multinomial distribution, with each branch of the tree representing a different hypothesized sequence of processing stages and leading to a specific decision outcome. Hence, assumptions about the psychological processes in a given experimental paradigm can often be cast into the form of a processing tree structure in a natural manner (Klauer 2010).

Consistent with the choice modeling literature, the tree structure begins with a consumer choosing among categories (or types or channels), followed by brands and products. However, unlike choice models which rely on conditional probabilities to reach to the bottom of the hierarchy (i.e., product choice), my latent-trait MPT approach explicitly lets the latent trait determine the path to follow in traversing the tree structure. In addition, the latent trait for polygamy is active during every stage of the decision tree. I code polygamy separately for channel/type ($\boldsymbol{\theta}_{\mathrm{c}}$), brand ($\boldsymbol{\theta}_{\mathrm{b}}$) and product ($\boldsymbol{\theta}_{\mathrm{p}}$). Figure 3 shows the structure of the multinomial processing tree. Each product category is modeled by separate subtrees of the multinomial model. For a given product category, a consumer will first decide whether to choose multiple types/channels or a single one, then decide whether to choose multiple brands, and finally whether to choose multiple products[29]. Therefore, there will be up to a total of three latent trait parameters and eight mutually exclusive decision outcomes (end nodes) for each product category. My model building can be viewed as a three-step hurdling process: as illustrated in Figure 3, the model starts with observed individual level decision outcome frequencies, with the paths leading to the outcomes governed by the latent processes; then, it employs a Probit-link to transform individual parameters to population/prior, which is further specified using a hyperprior. Lastly, data augmentation is used for easier empirical estimation.

3.3.1 Person-Level Model

Specifically, for product category or subtree k, $k = 1, \dots, K$, $j = 1, \dots, J_k$, and consumer t, $t = 1, \dots, T$, the decision outcome/node C_{kj} is mutually exclusive and has a frequency n_{kjt}, which follows a multinomial distribution with parameters p_{kjt}, $j = 1, \dots,$ J_k,. For product category k, let N_k be the fixed number of responses. Across product categories k, the data are assumed to be distributed stochastically independent for each consumer t. In Table 7, an overview of the most important symbols is given for easy reference.

[29] I conducted robustness checks on the sequence of decision making process (e.g., product first, then brand and lastly type) and the results do not vary significantly.

Let p_{kjt} denote the choice probabilities of reaching the end node decision outcome C_{kj} by means of S structural parameters θ_s, $s = 1, \ldots, S$, each θ_s being probabilistic and free to vary in $(0, 1)$ (Ansari et al. 2007; Klauer 2010):

$$p_{kjt} \;=\; P\left(C_{kj}\middle|\,\theta_t\right), \tag{16}$$

Here, $\boldsymbol{\theta}_t$ is the vector of consumer t's parameter values θ_{st}, $s = 1, \ldots, S$. It represents a sequence of latent binary events which determine the path followed in traversing the tree. The choice probabilities $P(C_{kj} \mid \boldsymbol{\theta})$ sum to 1. We can use a simple EM algorithm for maximum-likelihood estimation of the model parameters (Hu and Batchelder 1994). This form can be characterized by means of the model's representation as a processing tree (e.g., Figure 3). Let the number of paths ending in decision outcome C_{kj} of subtree k be I_{kj}, and let the ith such path be denoted by B_{kji}. The probability that path B_{kji} is followed by consumer t in traversing the tree is given by:

$$P\left(\beta_{kji}\middle|\,\theta_t\right)= \prod_{s=1}^{S}\theta_{st}^{a_{skji}}(1-\theta)_{st}^{bskji}\ , \tag{17}$$

where a_{skji} and b_{skji} are the number of branches on path B_{kji} that are assigned to parameter θ_s and its complement $1-\theta_s$, respectively. The probabilities for a given node are then computed by adding the probabilities of all paths that terminate in the respective decision outcome:

$$P\left(C_{kj}\middle|\,\theta_t\right)=\sum_{i=1}^{I_{kj}}\prod_{s=1}^{S}\theta_{st}^{a_{skji}}(1-\theta)_{st}^{bskji}\ , \tag{18}$$

The vector of person-level decision outcome counts $\mathbf{n}_t = (n11_t, \ldots, n1J_1t, \ldots, nK1t, \ldots, nKJ_Kt)$ is modeled by a vector-valued random variable $\mathbf{N}_t$ that follows a product-multinomial distribution:

$$P\left(N_t \;=\; n_t|\theta_t\right)=\prod_{k=1}^{K}\left\{\binom{N_k}{n_{k1t}\cdots n_{kj_kt}}\prod_{j=1}^{J_K}[P\left(C_{kj}|\theta_t\right)]^{n_{kjt}}\right\}, \tag{19}$$

The model from Equation 19 defines the person-level model. In the next sections, I will specify the prior distribution, hyperprior distribution, and the Gibbs sampler required for the analysis.

3.3.2 Prior Distribution

Ansari et al. (2007) use a logit link to transform parameters from the interval (0, 1) to the real line and to model the transformed parameters by a multivariate normal distribution with arbitrary mean μ and arbitrary covariance matrix Σ to be estimated from the data. Klauer (2010) employs a similar approach through a probit link and a less informative hyperprior distribution with a Gibbs sampler.

Specifically, the person-level model is re-parameterized by means of new population-level parameters α_{st} linked to the original personal-level parameters θ_{st} via $\alpha_s = \Phi^{-1}(\theta_{st})$, $s = 1, \ldots, S$, $t = 1, \ldots, T$, where Φ is the cumulative distribution function of the standard normal distribution. Let us collect the parameters α_{st} in the vector $\boldsymbol{\alpha}_t$. Across individual consumer t, the parameter $\boldsymbol{\alpha}_t$ is assumed to follow a multivariate normal distribution with mean vector μ and covariance matrix Σ:

$$\boldsymbol{\alpha}_t \sim N(\mu,\Sigma). \tag{20}$$

That is, the person-level model is the multinomial-processing tree model with probit-transformed model parameters. It allows for separate parameter estimates for each person, but the population-level model constrains the individuals' parameters to be distributed according to a multivariate normal distribution with mean and covariance matrix to be estimated from the data.

3.3.3 Hyperprior Distribution

In the Bayesian framework, a hyperprior distribution is required for the population-level parameters of the prior distribution with mean μ, which is assumed to follow an independent normal distribution with mean zero and variance $p = 100$, and a covariance matrix, which is assumed to follow a scaled Inverse–Wishart distribution. Using a new set of scale parameters λ_s, $s = 1, \ldots, S$, they decompose $\Sigma = (\sigma_{kl})$ as follows:

$$\Sigma = \mathrm{Diag}(\lambda_s)Q\mathrm{Diag}(\lambda_s) \tag{21}$$

where $Q = (q_{kl})$ and $\sigma_{ss} = \lambda_s^2 q_{ss}$. Whereas the correlations $\rho_{kl} = \sigma_{kl}/\sqrt{(\sigma_{kk}\sigma_{ll})}$ are determined only by Q, that is, $\rho_{kl} = q_{kl}/\sqrt{(q_{kk}q_{ll})}$. Assuming an Inverse–Wishart distribution for Q with $S+1$ degrees of freedom and scale matrix set to the identity matrix I therefore maintains the desirable uniform distribution for the parameter correlations. The parameters of interest are $\boldsymbol{\alpha}_t$, μ, and Σ. The following hyperprior distribution results:

$$\begin{aligned} &\mu \sim N(\mathbf{0}_S, 100I), \\ &Q \sim \text{Inverse–Wishart}_{S+1}(I), \\ &\boldsymbol{\lambda} \sim N(\mathbf{1}_S, 100I), \end{aligned} \tag{22}$$

where $\mathbf{0}_S$ and $\mathbf{1}_S$ are vectors of dimension S with zero and one, respectively, in each cell.

3.3.4 Data Augmentation for the Gibbs Sampler

The proposed model includes two steps of data augmentation that are required for the Gibbs Sampler. First, I augment the decision outcome frequencies n_{kjt} by the path frequencies m_{kjit} and collect all path frequencies in the vector **m**. Second, a different random variable Z is assigned to each node. As shown in Figure 5, as the tree is traversed, the upper branch emanating from a given node is taken if the associated $Z > 0$ and the lower branch if $Z \leq 0$. Let Z follow an independent normal distribution with mean α_s with $\alpha_s = \Phi^{-1}(\theta_s)$ and variance 1. From a theory point of view, the decision outcomes nodes can be viewed as binary choice points with choices driven by unobserved latent variables Z_{slt} exceeding a given threshold or not. For example, the choice may indicate whether a consumer's polygamy is triggered and activated. Since each node is assigned to one of the processes postulated by the multinomial model, and the outcome of the process determines which choice is made in moving through the processing tree, they provide a substantive underpinning of latent processes beyond mere technical convenience.

Specifically, each person runs through N_k trials for product category (or subtree) k, $k = 1, \ldots, K$. Each such trial x, $x = 1, \ldots, N_k$, defines R_k random variables Z_{kxrt} , $r = 1, \ldots, R_k$, where R_k is the number of nodes or decision outcomes in product category k. The vector $\mathbf{Z}$ collects all Z_{kxrt} in a fixed order. Each node indexed by k and r is assigned one of the person-level parameters α_s. Let the number of nodes associated with parameter α_s in subtree k be o_{ks} . Across subtrees k, there are $n_{st} = \sum_k N_k o_{ks}$ random variables Z per consumer with mean α_{st}, consumer t's value on parameter α_s . An alternative way to index the $\sum_t \sum_s n_{st}$ elements of $\mathbf{Z}$ is therefore as Z_{slt} with Z_{slt} being the lth element of those elements of $\mathbf{Z}$ that are assigned parameter α_{st} as its mean.

Furthermore, it turns out that all conditional posterior distributions that are needed for the Gibbs sampler, other than the conditional posterior distribution of the individual Z_{slt}, do not depend on the order in which the paths occurred, nor on the order in which the n_{st} values of Z_{slt} were observed for each s and t (Klauer 2010). Therefore, we can work with order statistics$\boldsymbol{Z}_{\boldsymbol{st}}^{\mathbf{0}}$, in which the n_{st} variables Z_{slt} appear in ascending order. Let $\boldsymbol{Z}_o$ be the vector that stacks the order statistics$\boldsymbol{Z}_{\boldsymbol{st}}^{\mathbf{0}}$, $s = 1, \ldots, S$, $t = 1, \ldots, T$. The double data augmentation procedure by path frequencies $\mathbf{m}$ and by $\boldsymbol{Z}^o$ allows the posterior distribution of the model parameters to be expressed as a standard hierarchical linear regression with the given $\boldsymbol{Z}^{\mathbf{0}}_{slt}$ as the data, and therefore facilitates straightforward adaptation of the well-understood Gibbs sampler for analyzing standard hierarchical linear regression models in a Bayesian framework (e.g., Gelman and Hill 2007). Thus, the remainder of the model can be analyzed as though it was the following linear model:

$$\boldsymbol{Z}_t \sim N(X_t\boldsymbol{\alpha}_t, I),$$

$$\boldsymbol{\alpha}_t \sim N(\mu,\Sigma), \qquad (23)$$

with (μ,Σ) distributed according to the hyperprior specified above and X_t as a design matrix containing zeros and ones that simply assigns α_{st} as mean to each Z_{slt} , $l = 1, \ldots,$ n_{st} , $s = 1, \ldots ,S$, $t = 1, \ldots , T$.

To summarize, the double data augmentation and the probit link have two advantages. Technically, they allow replacing observed categorical responses by continuous data with an underlying linear Gaussian structure (Albert and Chib 1993). More importantly, they provide a substantive underpinning of latent processes.

3.3.5 The Gibbs Sampler

A Gibbs sampler is a Monte Carlo–Markov chain algorithm for sampling from the posterior distribution of the model parameters given the data **n**. Let us then re-parametrize the parameters $\boldsymbol{\alpha}_t$ as follows:

$$\boldsymbol{\alpha}_t = \mu+\boldsymbol{\beta}_t ,$$

$$\boldsymbol{\beta}_t = \mathrm{Diag}(\lambda_s)\boldsymbol{\gamma}_t . \qquad (24)$$

The parameter μ is the prior mean of the parameters and the parameters $\boldsymbol{\beta}_t$ are the individual-specific systematic deviations from it. The parameters λ_s are the scale parameters of the scaled Inverse–Wishart distribution and the parameter $\boldsymbol{\gamma}_t$ is an unscaled version of $\boldsymbol{\beta}_t$. Let $\boldsymbol{\gamma}$ be the vector that stacks the vectors $\boldsymbol{\gamma}_t$, $t = 1, \ldots ,T$. The Gibbs sampler cycles through blocks of parameters. For each block, one sample is drawn from the conditional distribution of the parameters of the block given the data and the remaining parameters. The parameter blocks for the Gibbs sampler are Q, $(\mathbf{Z}^o, \mathbf{m})$, $\boldsymbol{\gamma}$, $\boldsymbol{\lambda}$, and μ. The detailed conditional distributions are given below.

Conditional Distribution of Q

The conditional distribution of Q given the data and the other parameters depends only on the parameters $\boldsymbol{\gamma}_t$. Let S be the sum of cross-products of the $\boldsymbol{\gamma}_t$: $S - \sum_{t=1}^{T} \boldsymbol{\gamma_t \gamma_t}'$, then

$$Q \mid \mathbf{m}, \mathbf{Z}^o, \boldsymbol{\gamma}, \mu, \boldsymbol{\lambda}, \mathbf{n} \sim \text{Inverse–Wishart}_{T+S+1}(I + S), \tag{25}$$

Conditional Distribution of $(\mathbf{Z}^o, \mathbf{m})$

The conditional distribution of $(\mathbf{Z}^o, \mathbf{m})$ is sampled from by sampling the conditional distribution of $\mathbf{m}$ with $\boldsymbol{Z}^o$ integrated out, followed by sampling from the conditional distribution of $\boldsymbol{Z}^o$ given $\mathbf{m}$, $\mathbf{n}$, and the other parameters.

The conditional distribution of $\mathbf{m}$ given the data and the other parameters depends only on the data $\mathbf{n}$ and the parameters $\boldsymbol{\gamma}_t$, μ, and $\boldsymbol{\lambda}$. For each person and decision outcome C_{kj} , the path frequencies m_{kjit} , $i = 1, \ldots, I_{kj}$, follow a multinomial distribution with parameters n_{kjt} and p_i, $i = 1, \ldots, I_{kj}$, as defined in Equation 7 (note that $\theta_{st} = \Phi(\mu_s + \lambda_s \gamma_{st})$, hence $p_i = p_i(\mu, \gamma_t, \lambda)$). Thus, $\mathbf{m}$ follows a product-multinomial distribution:

$$\mathbf{m} \mid Q, \gamma, \mu, \lambda, \mathbf{n} \sim \otimes_{t=1}^{T} \otimes_{k=1}^{K} \otimes_{j=1}^{J} \text{Multinomial}\,(n_{kjt}, (p_i\,(\mu, \gamma_t, \lambda)_{i=1,\ldots,I_{kj}})), \tag{26}$$

Consider next the conditional distribution of $\boldsymbol{Z}^o$ given $\mathbf{m}$, $\boldsymbol{\gamma}$, $\boldsymbol{\lambda}$, μ, Q, and $\mathbf{n}$. To derive this distribution, consider first the conditional distribution of the (unordered) $\boldsymbol{Z}$. Let $\boldsymbol{P}$ be a sequence of paths, $\boldsymbol{P} = (P_{kxt})_{kxt}$, path P_{kxt} being a path of subtree k assigned to individual t 's trial x, $t = 1, \ldots, T$, $k = 1, \ldots, K$, $x = 1, \ldots, Nk$. Let $\xi_{\mathbf{m}}$ be the set of sequences of paths $\boldsymbol{P}$ consistent with path frequencies $\mathbf{m}$, that is, with m_{kj} *it* being the number of trials x with $P_{kxt} = B_{kji}$ for each k,j,i, and t. By definition of conditional probabilities, the density of $\boldsymbol{Z}$ is:

$$\mathrm{f}\,(\mathbf{Z} \mid \mathbf{m}, Q, \boldsymbol{\gamma}, \mu, \boldsymbol{\lambda}, \mathbf{n}) = \sum_{P \in \xi_{\mathbf{m}}} \mathrm{f}\,(\boldsymbol{Z} \mid \boldsymbol{P}, \mathbf{m}, Q, \boldsymbol{\gamma}, \mu, \boldsymbol{\lambda}, \mathbf{n})\ \mathrm{P}\,(\boldsymbol{P} \mid \mathbf{m}, Q, \boldsymbol{\gamma}, \mu, \boldsymbol{\lambda}, \mathbf{n}), \qquad (27)$$

The conditional distribution of $\boldsymbol{Z}^o$ given the data, the path frequencies $\mathbf{m}$, and the other parameters need to be generated only to the point that it is consistent with the path frequencies $\mathbf{m}$, and the order information is not required. Let $n_{st}^{+}=\sum_{k=1}^{K} \sum_{J=1}^{J_K} \sum_{i=1}^{I_{KJ}} a_{skji}\, m_{kjit}$ normal variates Z_{slt} with mean α_{st} truncated from below at zero, $n_{st}^{-}=\sum_{k=1}^{K} \sum_{J=1}^{J_K} \sum_{i=1}^{I_{KJ}} b_{skji}\, m_{kjit}$ normal variates Z_{slt} with mean α_{st} truncated from above at zero, and $n_{st} - n_{st}^{+} - n_{st}^{-}$ nontruncated normal variates Z_{slt} with mean α_s. It is sufficient to generate $n_{st}^{+}$and n_{st}^{-} truncated normal variates with mean α_{st} and variance one truncated at zero from below and above, respectively, as well as $n_{st} - n_{st}^{+} - n_{st}^{-}$ unconstrained normal variates with mean α_{st} and variance one for each parameter s and individual t.

Conditional Distribution of $\boldsymbol{\gamma}$

The different $\boldsymbol{\gamma}_t$, $t = 1, \ldots, T$, are conditionally independent, so that they can be sampled one after the other for a sample from the conditional distribution of $\boldsymbol{\gamma}$. For each person t, the conditional distribution of $\boldsymbol{\gamma}\ t$ can be derived as a Bayesian regression with data $\lambda_s^{-1}(Zslt - \mu_s)$, $s = 1, \ldots, S$, $l = 1, \ldots, n_{st}$, that are independently normally distributed with mean $\boldsymbol{\gamma}_t$ and variance λ_s^{-2} and with a normal prior for $\boldsymbol{\gamma}_t$, $\boldsymbol{\gamma}_t \sim N(\mathbf{0}_S, Q)$. Thus, the conditional distribution of $\boldsymbol{\gamma}_t$ given the data and the other parameters is multivariate normal with mean $\boldsymbol{g}_t$ and covariance matrix G_t given by

$$\boldsymbol{g}_t = G_t \mathrm{Diag}(\lambda_s) \boldsymbol{u}_t\,,$$

$$G_t = (Q^{-1} + \mathrm{Diag}(n_{st}\lambda_s{}^2))^{-1}, \tag{28}$$

where $\boldsymbol{u}_t$ is the vector of the sums $\sum_{l=1}^{nst}(Z_{slt} - \mu_s)$, $s = 1, \ldots, S$.

Conditional Distribution of $\boldsymbol{\lambda}$

The conditional distribution of $\boldsymbol{\lambda}$ can be derived as a Bayesian regression with data $\gamma_{st}^{-1}(Z_{slt} - \mu_s)$, $s = 1, \ldots, S$, $l = 1, \ldots, n_{st}$, $t = 1, \ldots, T$, that are independently normally distributed with mean λ_s and variance γ_{st}^{-2}(and with a normal prior for $\boldsymbol{\lambda}$, $\boldsymbol{\lambda} \sim N(\mathbf{1}_S, pI)$, where p is the variance of the hyperprior of λ_s (i.e., $p = 100$). Thus, the conditional distribution of $\boldsymbol{\lambda}$ given the data and the other parameters is multivariate normal with mean $\boldsymbol{h}$ and covariance matrix H given by

$$\boldsymbol{h} = H\boldsymbol{v},$$

$$H = \mathrm{Diag}\,(\,p^{-1} + \textstyle\sum_{t=1}^{T} n_{st}\,\gamma_{st}^2\,)^{-1}, \tag{29}$$

where $\boldsymbol{v}$ is the vector of the terms $p^{-1} + \sum_{t=1}^{T}\gamma_{st}\,\sum_{l=1}^{nst}(Z_{slt} - \mu_s)$, $s = 1, \ldots, S$.

Conditional Distribution of μ

The conditional distribution of μ can be derived as a Bayesian regression with data $Z_{slt} - \lambda_s\gamma_{st}$, $s = 1, \ldots, S$, $l = 1, \ldots, n_{st}$, $t = 1, \ldots, T$, that are independently normally distributed with mean μ_s and variance one, and with a normal prior for μ, $\mu \sim N(\mathbf{0}_S, pI)$. Thus, the conditional distribution of μ given the data and the other parameters is multivariate normal with mean $\boldsymbol{u}$ and covariance matrix U given by

$$\boldsymbol{u} = U\boldsymbol{w},$$

$$U = \mathrm{Diag}(\,p^{-1} + \textstyle\sum_{t=1}^{T} n_{st})^{-1}, \tag{30}$$

where $\boldsymbol{w}$ is the vector of the sums $\sum_{t=1}^{T}\sum_{l=1}^{nst}(Z_{slt} - \lambda_s\,\mu_s)$, $s = 1, \ldots, S$.

3.3.6 Implementation

Rough initial estimates of the parameters μ and Σ are obtained by means of the Monte Carlo EM (MCEM) algorithm. For the expectation step, the conditional distribution of $\boldsymbol{\beta}_t$, $t = 1, \ldots, T$, is sampled via a Gibbs sampler for given μ and Σ and given the data. The Gibbs sampler samples from the relevant conditional distributions specified above with μ and Q fixed at their current estimates, and with λs fixed to one, $s = 1, \ldots, S$, so that $\Sigma = Q$ and $\boldsymbol{\beta}_t = \boldsymbol{\gamma}_t$. In the maximization step, μ is then estimated as the mean of the sampled $\boldsymbol{\beta}_t$, and Σ as the covariance matrix of the sampled $\boldsymbol{\beta}_t$. Initial overdispersed values of parameters $\boldsymbol{\beta}_t$ and μ are then obtained by sampling from multivariate t -distributions with three degrees of freedom with mean given by $\mathbf{0}_S$ and the MCEM estimates of μ, respectively, and covariance matrix given by the MCEM estimate of Σ and by Σ/T, respectively. Initial values of $\boldsymbol{\lambda}$ were sampled from a uniform distribution on the interval $(0.5, 1.5)$, and initial values $\boldsymbol{\gamma}_t$ were set to $\gamma_{st} = \beta_{st}/\lambda_s$, using the initial overdispersed values of parameters $\boldsymbol{\beta}_t$.

3.4 Data

The data needed for this empirical study are categories that are seemingly disparate, or rather, snapshots of consumer life experiences that cover a wide range of product categories. One suitable dataset is the National Consumer Survey. Therefore, I use the Simmons National Consumer Survey, which is filled by a nationwide sample of 5,014 individuals in the United States in 2006. It is considered one of the broadest and deepest surveys of American consumer behavior available. Consumers were asked to report their product purchases and brand preferences for a wide range of categories. The selection criteria for the categories used in this essay are: 1) the product category is among the top 10 TNS/Kantar most advertised categories, 2) data in the category is complete, and 3) the combinations of the product categories pass the pretest of “disparateness”[30]. The categories that satisfy the criteria above are: Financial Investments (including fixed income, equity and others), Soft Drinks (including carbonated diet, carbonated non-diet, noncarbonated diet and noncarbonated non-diet) Automobiles (including SUVs, compact, midsized, full-sized, sports, pickups, vans, luxury cars.), and Cell Phone Plans (including pre-paid, family-share and individual-monthly). Note that these categories cover durable, high involvement, long purchase cycle options as well as nondurable, low involvement, FMCG options, thereby giving greater variations in degree of dissimilarity and distinctiveness (i.e. truly disparate). In each category, respondents report the up to four most recent purchases with respect to types, brands and products. Similarly, I further trim the data to include individuals who at least have one purchase in each respective category. Although the number of observations in these categories is as many as I would want to have, the consumer survey is, by far, the only study available in the field that captures consumption patterns across a variety of disparate categories. It allows greater examination of underlying psychological processes without compromising statistical power.

To ease the concern of a limited number of observations, I use a media diary that is filled by the same 5,014 individuals during the same time when the National Consumer Survey was issued in 2006[31]. The media diary is from Universal McCann’s Media in Mind Diary 2006 and consists of self-reported media activities -- i.e., computer (including Internet), television, radio, or print (newspapers and magazines). This media diary is conducted annually with a randomly-selected, nationwide sample in the United

[30] The pretest asks a random sample of 30 respondents to rate how similar or dissimilar they think the product categories are on a 1-7 scale. The combination of categories chosen has an average of 1.83 (with 1 being most disparate).

[31] The Media Consumption category is also pre-tested for “disparateness” with other product categories.

States, and is considered the largest survey on consumer media consumption conducted by any media agency. The timing intervals in the diary are defined by half-hour time slots. Thus, at any given time, a panelist could consume one or a combination of these alternatives (i.e., multiplexing). Respondents report their activities for each media channel every half hour for seven consecutive days, except for the time periods from 1AM am-3AM and 3AM-5AM, which are each recorded as two individual observations. I further trim the diary to include 1,775 individuals who consumed media activities at least once during any half-hour slot in the observation window. A sample data structure is presented in Figure 5. For each respondent, I also have selected demographic information including age, gender, household income, household size, and location information such as whether the respondent is from an urban or rural area.

Table 8 reports detailed descriptive statistics for each product category. In the Cell Phone Plan category, almost zero percent of consumers hold multiple types of plans, which is sensible because consumers rarely belong to both an individual plan and a family plan. Polygamy happens more at the brand level and product/service level (e.g., ring tones, caller IDs, etc.). Note that Financial Investment is a special category because information on brands and products are confidential. Nevertheless, respondents report detailed investment sub-types/formats. For instance, fixed income includes six formats: treasury bills, savings bonds, U.S. government bonds, municipal bonds, money markets and corporate bonds. Equity includes three formats: company stocks, common stocks and equity mutual funds. Others include four formats: other securities (e.g., futures and derivatives), investment collectibles, international investments, and trust funds. Specifically, 37.8 percent of the 3,014 respondents hold multiple types of financial investments, with an average of 2.17 types. 44.9 percent of the respondents hold multiple sub-types/formats, with an average of 3.06. Clearly, there is a considerable group of single-type investors exhibiting polygamy for different investment formats. Furthermore, I observe significant variation in polygamy across product categories (trees), and across tree levels. For example, in the Soft Drinks category, there are four major brands (Coco-Cola, Pepsi, Dr. Pepper and other brands) as in Dubé (2004) and 96 products/SKUs (23 Coco-Cola products, 20 Pepsi, 27 Dr. Pepper and 26 other brands). 76.5 percent of all 4,452 consumers purchase multiple types within the last seven days, with an average of 2.42 types, 72.8 percent purchase multiple brands with an average of 2.79, and 89 percent purchase multiple products with an average of 7.69. In a nutshell, summary statistics suggest that this data is sparse with large variations across categories (trees) and across decision stages (tree levels). It is also sensible and reflects reality (that less polygamy happens for specialty retailing products such as Cell Phone Plans, and more so for convenience products such as Soft Drinks).

3.5 Results

3.5.1 Model Selection and Goodness-of-Fit Measures

The deviance information criterion (DIC) is a Bayesian analogue of information measures such as Akaike's information criterion in that it comprises a term quantifying lack of model fit and a term penalizing model complexity. The latter term, *pD*, is of interest in its own right in that it is interpreted as the effective number of parameters. It is smaller than the actual number of parameters to the extent to which the model parameters are constrained by dependencies in the data or the prior. DIC can be computed on the basis of the output from the Gibbs sampler. A point estimate of the parameter estimates $\boldsymbol{\theta}_t$ is also required, and I used the maximum likelihood estimates from separate analyses conducted for each individual t. The model with the smallest DIC value strikes the best compromise between fit and complexity in the metric defined by DIC. I will report DIC in Section 6 where I compare the latent trait approach with the latent class approach.

3.5.2 Parameter Estimates

I obtain parameter estimates for each individual across all product categories (except for media) and summarize them in Table 9. Table 9 shows the posterior percentiles for the parameters (on the probability scale) and the posterior medians of the Probit-transformed parameters. The rows present the product categories or subtrees, whereas the columns present the latent trait parameter θ's at different levels of the tree. A high θ (close to 1) denotes a high level of polygamy. Several aspects of Table 9 are noteworthy. First, the posterior medians are able to reproduce the underlying population means with little bias. Second, there are variations of the magnitude of polygamy across tree levels, as well as across trees (product categories). Third, the relatively high standard deviations suggest evidence for large individual differences in the impact of polygamy. Let us take three real data records for illustration purposes: ID 1227670 is a young female from Los Angeles with all of her parameters close to 1 (e.g., 0.5, 0.7, 0.8, 0.7, 0.9,…). This suggests that she is a "Polygamist" that would love to enjoy offers of multiple products, brands and types. Managers should label her as a desirable candidate for cross-selling and attempt to provide a large assortment for her selection. In contrast, ID 1162260 is a senior male from New York City with low parameter estimates (e.g., 0.0, 0.1, 0.2, 0.1, 0.3,…). He is a "Monogamist" that exhibits high inertia in purchase patterns across consumption scenarios. Managers may want to avoid going through expensive cross-selling efforts but rather deepen a strong long-term relationship with him with just one type of product or service. Most consumers are like ID 1357204, who is being a "Wanderer" that is polygamous in some situations, but not in others (e.g., 0.1, 0.3, 0.5, 0.8, 0.9, …). My individual-level results on latent traits not only offer an empirical-based, theory-grounded process for understanding individual variations in cross-category decisions, but also provide a new basis for segmentation and profiling to generate important managerial insights on coordinating across categories and across different types of customers, as discussed in Section 3.6.

3.5.3 Model Comparison: Latent Trait versus Latent Class

I now compare the results from the latent-trait model with the results from the latent-class MPT model. The latent-class version of the multinomial processing tree is given as follows: $p_{kjt} = p_{kj}(\boldsymbol{\theta}_t)$, where $\boldsymbol{\theta}_t$ is the vector of the $\boldsymbol{S}$ parameter values by person t. Allowing for different parameters for each person t, the vector of person-wise category counts $(n11_t, \ldots, n1J_1t, \ldots, nK1t, \ldots, nKJ_Kt)'$ is still modeled by a vector-valued random variable **N** that follows a product-multinomial distribution

$$P\ (N_t = n_t|\theta_t) = \prod_{k=1}^{K}\left\{\binom{N_k}{n_{k1t}\ldots n_{kj_kt}}\prod_{j=1}^{J_K}[P_{kj}\ (\theta_t)]^{n_{kjt}}\right\}, \tag{31}$$

Let the model parameters follow a distribution with probability measure μ, then

$$P\ (N = n) = \int P\ (N = n|\eta)d\mu(\eta), \tag{32}$$

where $P\ (N = n|\eta)$is given by the right side of Equation 31, in which the fixed values $\boldsymbol{\theta}_t$ are replaced by the variable of integration, η, and n_t is replaced by n. Therefore, for T consumers, we have:

$$P\ ((N_1, \ldots, N_T) = (n_1, \ldots, n_t)) = \prod_{t=1}^{T}\{\int P\ (N_t = n_t|\eta)d\mu(\eta)\}, \tag{33}$$

Let μ be distributed over a finite number C of fixed parameter vectors $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_C$. If $\boldsymbol{\lambda}_c = \mu(\{\boldsymbol{\theta}_c\})$ is the size of class c, the model equation simplifies to:

$$P\ ((N_1, \ldots, N_T) = (n_1, \ldots, n_t)) = \prod_{t=1}^{T}\{\sum_{c=1}^{C}\lambda_c P\ (N_t = n_t|\theta_c)\}, \tag{34}$$

This means that each consumer t is assumed to belong to one of the C latent classes of proportional sizes $\boldsymbol{\lambda}_c$. In a latent-class multinomial model, the category counts jointly follow a mixture of product-multinomial distributions, and each category count considered individually follows a mixture of binomial distributions. Furthermore, it is well known that mixtures of binomial distributions with parameters p_c and $\boldsymbol{N}$ and mixture coefficients $\boldsymbol{\lambda}_c$ are identified if and only if N >= 2C – 1. A simple EM-algorithm can then be devised for the maximum-likelihood estimation of latent-class multinomial models.

Table 10 shows the model fit statistics for both the latent-trait model and the latent-class model. Smaller DIC values suggest that the proposed model outperforms the latent class model significantly. Following Klauer (2006), two test statistics, termed M1 and M2, are considered for mean structure testing, and another two test statistics, termed S1 and S2, for variance-covariance structure testing. All four statistics are asymptotically distributed as χ^2 when the degrees of freedom are larger than zero. Table 11 shows the detailed results from the latent class model. Parameter estimates for the Cell Phone Plans and Media Consumption category are not identified in the latent class framework because the probability is trivially close to zero (or one) so that there is not enough variation in the data for the model to distinguish multiple segments. In addition, not surprisingly, all the other categories seem to have two distinct classes: the polygamous class, and the single class. While I observe significant differences across tree levels and across trees, there are quite a number of places where the latent-class approach is not able to accurately capture the coefficients to reflect the true population mean (as indicated by the zero values), indicating a poor job of capturing underlying distribution with the discrete representation of consumer heterogeneity.

3.6 Segmentation and Prediction Analysis

I conducted a finite mixture analysis to segment the consumers based on the collection of individual $\boldsymbol{\theta}$ parameters ($\theta_{c,}$ $\theta_{b,}$ θ_p) across categories, with θ being free to vary between (0,1). Such continuous representation of consumer heterogeneity allows one to achieve value-based segmentation where consumers are grouped based on their decision processes rather than binary observed behavior, 0 or 1. It both provides richer theoretical support and better empirical fit with continuous distribution. Table 12 summarizes the segmentation results. The three segments “Polygamist”, “Monogamist” and “Wanderer” roughly each represent 20 percent, 10 percent and 70 percent of the data respectively. Profile analysis shows group differences are significant. As illustrated in the previous example in Section 5.2, the Polygamist segment shows high $\boldsymbol{\theta}$s across decision tree stages and categories, whereas the Monogamists show the opposite. The $\boldsymbol{\theta}$s for the Wanderer segment lie in between.

Next, I perform two types of out-of-sample predictions: customer-based and product/category-based. For the customer-based prediction, the idea is to find customers that behave similarly and use their parameters to predict the decisions of the holdout sample (15 percent). For product/category-based prediction, the assumption is that customers may exhibit similar behavior across multiple categories (Ainslie and Rossi 1998). For example, if a customer is price sensitive in the toothbrush category, then he may be sensitive to the toothpaste category, or even the clothing category. Specifically, I use parameters from three of the five categories to predict outcomes of the other two categories and report the average hit rates. Table 13 shows hit rate by segment using both types of prediction. In summary, while the latent trait model is not designed for prediction (but rather for assessing the underlying processes), the hit rates still seem reasonable (more than 60 percent), although it is much harder to predict decisions of the Wanderer segment as compared to the Polygamist and the Monogamist segments which exhibit more consistent behavior across categories. In addition, customer-based prediction yields better accuracy than product/category-based prediction. This finding relates back to the intuition that getting at the "why consumers did it" by looking at the underlying processes provides greater conceptual and empirical support as compared to the "what consumers did" question in the traditional behavioral segmentation approach.

3.7 Discussion and Conclusion

Much of marketing has focused on a consumer's choices and preferences in individual product categories or a set of closely related product categories. The reality is that consumers shop around a globe of categories that are much more diverse and complex than the traditionally defined "market basket". This research takes a first step in modeling a complete picture of consumer decision problems by examining consumption across seemingly disparate product categories to advance understanding in both dimensions of customer behavior: the breadth of their consumption portfolio and the depth of their latent decision processes. While traditional research on multi-category choice models and latent class suffer from data and modeling limitations that prohibit deeper investigation of the underlying process that governs consumer decision making, this research empirically examines consumer choices across seemingly disparate product categories using a latent trait hierarchical multinomial processing tree model. In doing so, this paper contributes to the consumer decision literature in three ways: 1) theoretically, the latent-trait approach provides rich support in examining the high level processes; 2) methodologically, the relative merits of models with continuous versus discrete representations of consumer heterogeneity are discussed; and 3) substantively, new insights on value-based targeting and profiling are presented with respect to managing across seemingly disparate product categories.

The power of the latent trait model lies in its ability to infer and assess the impact of underlying processes using behavioral data without necessarily augmenting survey or experiments on consumer attitudes. Segmentation and prediction analysis suggests that the approach of categorizing consumers as collections of latent process parameters provides better theoretical and empirical support for value/process based segmentation and targeting exercise.

The idea of modeling individual latent processes is not bound by a particular context, but is applicable to a broader phenomenon that is generally manifested across a wide range of settings and situations. It would be especially intriguing to study the impact of latent processes in the online world where firms may have access to large-scale behavioral data across categories and situations. Future research can look at how firms can improve current recommendation systems based on inferred consumer preferences across categories, and how brand constellations are formed in social media (e.g., a consumer may "like" many seemingly unrelated brands on Facebook).

5. REFERENCES

Ainslie, A. and P.E. Rossi (1998), “Similarities in Choice Behavior Across Multiple Categories,” *Marketing Science*, 17(2), 91-106.

Allenby, G. M. and P. E. Rossi (1999), “Marketing Models of Consumer Heterogeneity,” *Journal of Econometrics*, 89 (1-2), 57–78.

Andrews, R. L., A. Ainslie, and I. S. Currim (2002), “An Empirical Comparison of Logit Choice Models with Discrete Versus Continuous Representations of Heterogeneity,” *Journal of Marketing Research*, 39 (4), 479-487.

Andrews, R. L., A. Ansari, and I. S. Currim (2002), “Hierarchical Bayes vs. Finite Mixture Conjoint Analysis Models: A Comparison of Fit, Prediction, and Partworth Recovery,” *Journal of Marketing Research*, 39 (1), 87-98.

Ansari A., C. Mela and S. Neslin (2008), “Customer Channel Migration,” *Journal of Marketing Research*. 45(1), 60-76.

Ansari, A., M.Vanhuele and M. Zemborain(2008), “Heterogeneous Multinomial Processing Tree Models,” Working paper, Columbia University, New York.

Balasubramanian, S. and W. A. Kamakura (1989), “Measuring Consumer Attitudes Towards the Marketplace with Tailored Interviews,” *Journal of Marketing Research*, 26(3), 311-26.

Bawa.K., S.S. Srinivasan and R. K. Srivastava (1997), “Coupon Attractiveness and Coupon Proneness: A Framework for Modeling Coupon Redemption,” *Journal of Marketing Research*, 34(4), 517-525.

Batchelder, W.H., and D.M. Riefer (1999), “Theoretical and Empirical Review of Multinomial Processing Tree Modeling,” *Psychonomic Bulletin & Review*, 6(1), 57–86.

Belch G. E. & Belch M. A. (2001), *Advertising and Promotion: An Integrated Marketing Communication Perspectives*, US: McGraw-Hill.

Bell, D.R. and J.M. Lattin (1998), “Shopping Behavior and Consumer Preferences for Store Price Format: Why Large Basket Shoppers Prefer EDLP,” *Marketing Science*, 17, 1, 66-88.

Berry, S. T. (1994), "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25(2), 242-262.

Bettman, J. (1979), *An Information Processing Theory of Consumer Choice,* Reading, MA: Addison-Wesley Publishing Company.

Bettman, J., M. F. Luce and J. W. Payne (1998), “Constructive Consumer Choice Processes,” *The Journal of Consumer Research*, 25(3), 187-217.

Bhat, C.R. (2005), “A Multiple Discrete-Continuous Extreme Value Model: Formulation and Application to Discretionary Time-Use Decisions,” *Transportation Research Part B*, 39(8), 679-707.

Bucklin, R. E. and Gupta, S. (1992), “Brand Choice, Purchase Incidence, and Segmentation: An Integrated Modeling Approach,” *Journal of Marketing Research*, 29(2), 201-215.

Berns G.S., J.D. Cohen, and M.A. Mintun (1997), “Brain Regions Responsive to Novelty in the Absence of Awareness,” *Science*, 276(53), 1272-1275

Chan, T. (2006), “Estimating a Continuous Hedonic Choice Model with an Application to Demand for Soft Drinks,” *RAND Journal of Economics*, 37(2), 466-482.

Chiang, J. (1991), “A Simultaneous Approach to the Whether, What and How Much to Buy Questions,” *Marketing Science*. 10(4), 297–315.

Chib, S., P.B. Seetharaman and A. Strijnev (2002), “Analysis of Multi-Category Purchase Incidence Decisions Using IRI Market Basket Data,” *Advances in Econometrics*, 16, 55-90.

Chintagunta, P.K. (2002), “Investigating Category Pricing Behavior in a Retail Chain, *Journal of Marketing Research*,” 39(2), 141-154.

Chung, J. and V.R. Rao (2003), “A General Choice Model for Bundles with Multiple-Category Products: Application to Market Segmentation and Optimal Pricing for Bundles,” *Journal of Marketing Research*, 40 (2), 115–130.

Coombs, C. H. (1964), *A Theory of Data*. New York: John Wiley & Sons.

Craig, C. S., B. Sternthal, and C. Leavitt (1976), “Advertising Wearout: An Experimental Analysis,” *Journal of Marketing Research*, 13 (4), 365-372.

Danaher, P.J. and J. R. Rossiter (2011), “Comparing Perceptions of Marketing Communication Channels.” *European Journal of Marketing*, 45(1/2), 6-42.

Danaher, P. J. and R. T. Rust (1996), "Determining the Optimal Return on Investment for an Advertising Campaign," *European Journal of Operational Research*, 95, 511-521.

Deaton, A., and J. Muellbauer (1980), *Economics and Consumer Behavior*, Cambridge: Cambridge University Press.

DeSarbo, W., W. A. Kamakura and M. Wedel (2004) “Applications of Multivariate Latent Variable Models in Marketing,” in *Advances in Marketing Research and Modeling: The Academic and Industry Impact of Paul E. Green*, Boston, MA: Kluwer, 43-67.

DeSarbo, W., W. A. Kamakura, M. Wedel (2006), “Latent Structure Regression,” in *Handbook of Marketing Research*, Thousand Oaks, CA: Sage Publications, 394-417.

Dubé, J.-P. (2004), “Multiple Discreteness and Product Differentiation: Demand for Carbonated Soft Drinks,” *Marketing Science*, 23(1), 66–81.

Dubé, J.-P., G. J. Hitsch, E. Rossi (2009), “Do Switching Costs Make Markets Less Competitive?” *Journal of Marketing Research,* 46(4), 435–445.

Edwards, Y. and G. M. Allenby (2003), "Multivariate Analysis of Multiple Response Data," *Journal of Marketing Research*, 40(3), 321-334.

Gelman, A. and J. Hill (2007), *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.

Green, P.E., F.J. Carmone and D. P. Wachspress (1976), “Consumer Segmentation via Latent Class Analysis,” *The Journal of Consumer Research*, 3(3), 170-174.

Guadagni, P. M., J. D. C. Little (1983), “A Logit Model of Brand Choice Calibrated on Scanner Data,” *Marketing Science*, 2(3), 203–238.

Gupta, S. and P. K. Chintagunta (1994), “On Using Demographic Variables to Determine Segment Membership in Logit Mixture Models,” *Journal of Marketing Research*, 31(1), 128-136.

Hansen, K., V. P. Singh, P. K. Chintagunta. (2006), “Understanding Store Brand Purchase Behavior across Categories,” *Marketing Science*, 25(1) 75–90.

Hedgcock W., Rao A. R. (2009), “Trade-off Aversion as an Explanation for the Attraction Effect: A Functional Magnetic Resonance Imaging Study. *Journal Marketing Research,* 46(1), 1–13.

Heilman, C. and D. Bowman (2002), “Segmenting Consumers Using Multiple-Category Purchasing Data,” International Journal of Research in Marketing, 19 (3), 225–52.

Hendel, I. (1999), “Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns,” *Review of Economic Studies*, 66(2), 423-446.

Hu, X. and W.H. Batchelder (1994), “The Statistical Analysis of General Processing Tree Models with the EM Algorithm,” *Psychometrika*, 59(1), 21–47.

Kahn, B.E., M.U. Kalwani, and D.G. Morrison (1986), ``Measuring Variety Seeking and Reinforcement Behaviors Using Panel Data,'' *Journal of Marketing Research*, 23(2), 89-100.

Kamakura, W.A. and G. J. Russell (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research,* 26(4), 379-390.

Kamakura, W.A., Sridhar R. and R. K. Srivastava (1991), "Applying Latent Trait Analysis in the Evaluation of Prospects for Cross-Selling of Financial Services," International *Journal of Research in Marketing*, 8(4), 329-349.

Kim, J., G.M. Allenby and P. E. Rossi (2002), "Modeling Consumer Demand for Variety," *Marketing Science*, 21(3), 223-228.

Klauer, K.C. (2006), "Hierarchical Multinomial Processing Tree Models: a Latent-Class Approach," *Psychometrika*, 71(1), 1–31.

Klauer, K.C. (2010), "Hierarchical Multinomial Processing Tree Models: a Latent-Trait Approach," *Psychometrika*, 75(1), 70–98.

Langeheine, R. and J. Rost (1988), *Latent Trait and Latent Class Models*, New York: Plenum.

Lazarsfeld,P. F. (1950) , "The Logical and Mathematical Foundation of Latent Structure Analysis," in *Measurement and Prediction*, Princeton, N.J.: Princeton University Press.

Lord, F. M. and M. R. Novick (1968), *Statistical Theories of Mental Test Scores*, Reading, MA: Addison-Wesley.

Lucas, R. (1976), "Econometric Policy Evaluation: A Critique", in Brunner, K.; Meltzer, A., *The Phillips Curve and Labor Markets*, Carnegie-Rochester Conference Series on Public Policy, 1, New York: American Elsevier.

McFadden, D. (1980), "Economic Models for Probabilistic Choice Among Products," *Journal of Business*, 53(3), 513-529.

Manchanda, P., A. Ansari and S. Gupta (1999), "The Shopping Basket: A Model for Multicategory Purchase Incidence Decisions," *Marketing Science*, 18(2), 95-114.

Mantrala, M. K. (2002), "Allocating Marketing Resources," in *Handbook of Marketing*, Barton A. Weitz and Robin Wensley, eds. Thousand Oaks, CA: Sage Publications.

McGuire W.J. (1976), "Psychological Factors Influencing Consumer Choice," in *Selected Aspects of Consumer Behavior,* Washington University Press, 319 – 360.

Mehta, N. (2007), “Investigating Consumers’ Purchase Incidence and Brand Choice Decisions across Multiple Product Categories: A Theoretical and Empirical Analysis”. *Marketing Science.* 26(2), 196–217.

Mehta, N., Chen, X. and Narsimhan, O. (2010), “Examining Demand Elasticities in Hanemann's Framework: A Theoretical and Empirical Analysis,” *Marketing Science*, 29(3), 422-437.

Naik, P. A., & Raman, K. (2003), “Understanding the Impact of Synergy in Multimedia Communications,” *Journal of Marketing Research ,* 40(4), 375-388.

Naik, P. A., K. Raman, and R. S. Winer (2005), “Planning Marketing-Mix Strategies in the Presence of Interaction Effects: Empirical and Equilibrium Analysis,” *Marketing Science*, 24 (6), 25-34.

Pilotta, J. J., Schultz, D. E., Drenik, G., & Rist, P. (2004), “Simultaneous Media Usage”: A Critical Consumer Orientation to Media Planning,” *Journal of Consumer Behavior*, 3(3): 285–92.

Pilotta, J. J., & Schultz, D. E. (2005), “Simultaneous Media Experience and Synesthesia,” *Journal of Advertising Research*, 45 (1), 19-26.

Rust, R. T. and M. I. Alpert, (1984), “An Audience Flow Model of Television Viewing Choice,” *Marketing Science*, 3(2), 113-124.

Schultz, D. E. (2002), “Outdated Approaches to Planning Needs Revamping,” *Marketing News*, 36 (23), 6-7.

Schweidel, D.A., E.T. Bradlow, and P.S. Fader (2011), "Portfolio Dynamics for Customers of a Multiservice Provider," *Management Science*, 57(3), 471-486.

Shachar, R. and J. Emerson. (2000), “Cast Demographics, Unobserved Segments, and Heterogeneous Switching Costs in a Television Viewing Choice Model,” *Journal of Marketing Research*, 37(2) 173-186.

Seetharaman, P.B., A. Ainslie and P.K. Chintagunta (1999), “Investigating Household State Dependence Effects Across Categories,” *Journal of Marketing Research*, 36(4), 488-500.

Shachar, R., and J.W. Emerson (2000), “Cast Demographics, Unobserved Segments, and Heterogeneous Switching Costs in a Television Viewing Choice Model,” *Journal of Marketing Research , 37* (2), 173-186.

Smith, T.M., S. Gopalakrishna, and R. Chaterjee (2006), “A Three-Stage Model of Integrated Marketing Communications at the Marketing–Sales Interface,” *Journal of Marketing Research*, 43(4), 564–579.

Song, I. and P.K. Chintagunta (2006). “Measuring Cross-Category Price Effects with Aggregate Store Data,” *Management Science*, 52 (10), 1594-1609.

Song, I. and P.K. Chintagunta (2007). “A Discrete-Continuous Model for Multicategory Purchase Behavior of Households.” *Journal of Marketing Research*, *44*(4), 595-612.

Steckel, J., R. Winer, R. Bucklin, B. Dellaert, X. Drèze, G. Häubl, S. Jap, J. D. Little, T. Meyvis, A. Montgomery, A. Rangaswamy (2005), "Choice in Interactive Environments," *Marketing Letters*, 16 (3-4), 309-20.

Tversky, A. (1977), “Features of Similarity,” Psychological Review, 84(4), 327-352.

Vakratsas, D. and T. Ambler (1999), “How Advertising Works: What Do We Really Know?” *Journal of Marketing,* 63 (1), 26–43.

Wales, T. J. and A.D. Woodland, (1983). “Estimation of Consumer Demand Systems With Binding Non-negativity Constraints,” *Journal of Econometrics*, 21(3), 263-285.

Wedel, M. and W. A. Kamakura (2000) *Market Segmentation: Conceptual Methodological Foundations*, Second Edition. Boston: Kluwer Academic Publishers.

Wedel, M., W.A Kamakura, N. Arora, A. Bemmaor, J. Chiang, T. Elrod, R. Johnson, P. Lenk, S. Neslin, and C. S. Poulsen (1999), “Discrete and Continuous Representations of Unobserved Heterogeneity in Choice Modeling,” *Marketing Letters*, 10 (3), 219–32.

Weller, J., I. Levin, B. Shiv, and A. Bechara (2011), “The Effects of Insula Damage on Decision Making for Risky Gains and Losses, *Social Neuroscience,* Forthcoming.

Wendel, S., and B.G.C. Dellaert, (2005). “Situation Variation in Consumers' Media Channel Consideration,” *Journal of the Academy of Marketing Science,* 33(4), 575‒584.

Wilbur, K. C. (2008), “A Two-Sided, Empirical Model of Television Advertising and Viewing Markets,” *Marketing Science*, 27 (3), 356-378.

TABLE 1: MEDIA SWITCHING MATRIX

AGGREGATE SWITCHING MATRIX

	COMPUTER	TELEVISION	RADIO	PRINT	OUTSIDE OPTION
COMPUTER	**6,850** **18%**	5,274 14%	2,563 7%	1,260 3%	21,865 58%
TV	5,238 8%	**15,009** **22%**	5,359 8%	2,521 4%	39,642 58%
RADIO	2,674 7%	5,212 14%	**7,749** **21%**	1,537 4%	20,451 54%
PRINT	1,265 8%	2,537 16%	1,499 9%	**1,320** **8%**	9,600 59%
OUTSIDE	21,770 8%	39,750 14%	20,445 7%	9,597 3%	**188,220** **67%**

Note: Columns represent media channel that consumers switch TO. Rows represent the media channel that consumers switch FROM. Cells indicate the number of consumers and their proportion in the sample. Diagonal elements indicate state dependence (i.e., inertia)

TABLE 2: IN SAMPLE HIT RATE COMPARISON (Each cell contain the % accuracy rate)

(A)SINGLE MEDIA CONSUMPTION

Media Activities	*OVERALL HIT RATE*	Computer	TV	Radio	Print	Outside Good
Hanemann Model	***60.30%***	50.81%	31.87%	29.47%	18.66%	82.04%
Main model	***97.00%***	99.13%	99.27%	97.60%	86.67%	96.38%

(B) MULTIPLEXING MEDIA CONSUMPTION

Media Activities	Computer &TV	Computer &Radio	Computer &Print	TV &Radio	TV &Print	Radio &Print	Computer &TV &Radio	Computer &TV &Print	Computer &Radio &Print	TV &Radio &Print	Computer & TV &Radio & Print
Hanemann Model	.	.	.	.	.	.	.	.	.	.	.
Main model	99.07%	98.41%	83.33%	88.64%	98.81%	92.00%	100.00%	100.00%	100.00%	100.00%	100.00%

(C) ROBUSTNESS CHECK

Media Activities	***OVERALL HIT RATE***	Computer	TV	Radio	Print	Outside Good	Computer &TV	Computer &Radio	Computer &Print	TV &Radio	TV &Print	Radio &Print	Computer &TV &Radio	Computer &TV &Print	Computer &Radio &Print	TV &Radio &Print	Computer & TV &Radio & Print
Subsample1	***95.59%***	99.54%	100.00%	100.00%	40.87%	100.00%	100.00%	17.53%	15.00%	58.33%	41.86%	50.00%	100.00%	33.33%	100.00%	100.00%	100.00%
Subsample2	***95.08%***	87.47%	99.37%	96.53%	51.28%	100.00%	97.06%	22.73%	11.11%	20.83%	42.93%	29.41%	100.00%	100.00%	100.00%	0.00%	0.00%
Subsample3	***97.82%***	99.42%	100.00%	100.00%	41.71%	99.97%	100.00%	25.93%	10.53%	53.66%	46.67%	0.00%	0.00%	81.25%	0.00%	-	-

TABLE 3: ESTIMATES FOR THE PROPOSED MODEL

Structural Parameters	COMPUTER	TV	RADIO	PRINT
Baseline Preference Parameters	Coefficient(T-Stat)	Coefficient(T-Stat)	Coefficient(T-Stat)	Coefficient(T-Stat)
Intercept	-8.47***	-7.013***	-8.153***	-9.803***
	(-116.75)	(-146.6)	(-112.8)	(-77.31)
Attention Levels For				
Computer	7.74***	-0.335***	-0.421***	-0.239***
	(431.4)	(-14.6)	(-16.48)	(-6.45)
TV	-0.05***	2.089***	-0.095***	0.104***
	(-4.35)	(218.12)	(-6.64)	(6.51)
Radio	-0.096***	-0.044***	2.212***	0.144***
	(-9.37)	(-3.51)	(139.43)	(8.65)
Print	-1.41***	0.469***	0.651***	3.599***
	(-41.45)	(26.33)	(29.03)	(151.34)
Media Penetration				
Cable	-0.063***	0.078***	0.0036	0.07***
	(-3.98)	(5.8)	(0.21)	(2.8)
PC at Home	0.043**	-0.084***	-0.295***	0.178***
	(1.94)	(-4.8)	(-13.61)	(5.48)
PC at Work	0.132***	-0.099***	-0.0041	-0.051***
	(7.72)	(-7.18)	(-0.23)	(-1.96)
Consumer Demographics				
Age < 35	0.038*	-0.103***	0.114***	-0.106***
	(1.72)	(-5.02)	(4.65)	(-2.77)
Age > 65	0.097***	0.117***	-0.0387	0.305***
	(4.11)	(6.68)	(-1.56)	(9.95)
Hhincome < 35k	0.081***	0.126***	0.129***	-0.0098
	(3.19)	(6.6)	(5.2)	(-0.28)
HHsize <= 3	-0.106***	-0.0174	-0.14***	-0.105***
	(-3.99)	(-0.82)	(-4.91)	(-2.6)
Hhsize>=8	-0.0062	0.0087	0.0241	0.0364
	(-0.24)	(0.4)	(0.9)	(0.92)
Female	0.031***	-0.071***	0.056***	-0.059***
	(1.99)	(-5.53)	(3.33)	(-2.49)
Urban	-0.093***	-0.0196	0.049***	0.101***
	(-4.87)	(-1.22)	(2.44)	(3.56)
Day of the Week (Saturday as the base)				
Sunday	-0.012	0.070***	-0.122***	0.235***
	(-.34)	-3.08	(-3.55)	-5.39
Monday	0.411***	-0.009	0.100***	0.304***
	-13.22	(-0.40)	-3.17	-6.95
Tuesday	0.221***	-0.025	0.164***	0.244***
	-7.05	(-1.09)	-5.27	-5.49
Wednesday	0.231***	-0.082***	0.180***	0.163***
	-7.32	(-3.65)	-5.77	-3.63
Thursday	0.253***	-0.090***	0.147***	0.203***
	-8.08	(-3.78)	-4.69	-4.54
Friday	0.212***	-0.135***	0.161***	0.095***
	-6.69	(-5.63)	-5.13	-2.08
Time-of-the-Day (am1200 as base)				
am100	-1.511***	-1.801***	-1.097***	-2.354***
	(-12.96)	(-23.67)	(-11.5)	(-9.21)
am300	-1.794***	-2.102***	-1.126***	-2.87***
	(-13.67)	(-24.35)	(-11.8)	(-11.29)

am500	-0.1211	-0.22***	0.791***	1.14***
	(-1.31)	(-3.7)	(10.3)	(8.59)
am600	0.755***	0.317***	1.332***	2.001***
	(9.41)	(5.89)	(18.25)	(16.05)
am700	1.277***	0.432***	1.657***	2.272***
	(17.63)	(8.63)	(23.98)	(18.7)
am800	1.222***	0.272***	1.342***	2.196***
	(17.45)	(5.32)	(19.18)	(17.99)
am900	0.888***	0.0614	1.072***	1.897***
	(12.64)	(1.15)	(14.93)	(15.29)
am1000	0.801***	-0.0527	0.87***	1.527***
	(11.19)	(-0.94)	(11.77)	(11.94)
am1100	0.771***	-0.0663	0.845***	1.348***
	(10.99)	(-1.21)	(11.67)	(10.44)
pm1200	0.849***	0.18***	1.054***	1.525***
	(11.84)	(3.42)	(14.5)	(11.96)
pm100	0.872***	0.176***	0.873***	1.453***
	(12.41)	(3.36)	(11.96)	(11.23)
pm200	0.827***	0.181***	0.95***	1.081***
	(11.83)	(3.48)	(13.1)	(8.14)
pm300	0.894***	0.291***	1.08***	1.215***
	(12.57)	(5.58)	(14.78)	(9.22)
pm400	1.013***	0.426***	1.196***	1.381***
	(14.41)	(8.62)	(16.69)	(10.68)
pm500	1.165***	0.641***	1.433***	1.464***
	(15.72)	(12.66)	(19.65)	(11.35)
pm600	0.835***	0.775***	1.116***	1.38***
	(11.29)	(16.65)	(15.27)	(10.79)
pm700	0.905***	0.851***	1.008***	1.401***
	(12.36)	(18.52)	(13.43)	(11.06)
pm800	1.139***	1.097***	0.793***	1.374***
	(15.72)	(24.05)	(10.14)	(10.82)
pm900	1.151***	1.219***	0.737***	1.407***
	(15.84)	(27.03)	(9.32)	(11.05)
pm1000	1.396***	1.242***	0.723***	1.374***
	(19.39)	(27.43)	(9.12)	(10.68)
pm1100	0.588***	0.846***	0.352***	0.999***
	(7.33)	(16.97)	(4.16)	(7.29)
Satiation Parameters				
Alpha	0.999***	0.447***	0.479***	0.448***
	(4.66)	(12.75)	(2.95)	(7.93)
	note: Alpha for the outside good is close to 1			
Gamma	1 for all inside goods, 0 for the outside good			
Random Components (Standard Deviations)				
RC_Baseline	0.209***	0.109***	0.0017	0.665***
	(14.1)	(2.74)	(0.05)	(17.45)
RC_Satiation	0.298***	0.0534	0.176	0.208
	(7.99)	(0.88)	(0.17)	(1.11)
EC_Baseline	0.0055	(1 EC in the baseline utility)		
	(0.56)			
EC_Satiation	0.0077	(1 EC among satiation parameters)		
	(0.77)			
	RC: Random Coefficients; EC: Error Components			

TABLE 4: OUT-OF-SAMPLE HIT RATE COMPARISON (Each cell contains the % accuracy rate)

(A) SINGLE MEDIA CONSUMPTION

Media Activities	*OVERALL HIT RATE*	Computer	TV	Radio	Print	Outside Good
Hanemann Model	***60.08%***	50.95%	30.83%	27.27%	21.90%	82.07%
Main model	***95.90%***	87.59%	99.87%	97.85%	85.15%	95.22%

(B) MULTIPLEXING MEDIA CONSUMPTION (CONTINUED FROM ABOVE)

Media Activities	Computer &TV	Computer &Radio	Computer &Print	TV &Radio	TV &Print	Radio &Print	Computer &TV &Radio	Computer &TV &Print	Computer &Radio &Print	TV &Radio &Print	Computer & TV &Radio & Print
Hanemann Model	.	.	.	.	.	.	.	.	.	.	.
Main model	100.00%	98.08%	89.47%	91.43%	98.85%	94.29%	100.00%	100.00%	87.50%	100.00%	100.00%

TABLE 5: MARGINAL RATE OF SUBSTITUTION AT THE POINT OF ZERO CONSUMPTION

	Computer	TV	Radio	Print
Computer	-	0.233	0.729	3.793
TV	4.289	-	3.126	16.268
Radio	1.372	0.320	-	5.204
Print	0.264	0.061	0.192	-

Note: The column represents the media option that will be SUBSTITUTED FOR, The row represents the media option that will be SUBSTITUTED TO. These effects are computed using only the media-specific intercepts in the Baseline utilities.

TABLE 6: COUNTERFACTUAL RESULTS FROM PRINT MEDIA EXIT IN LOS ANGELES

Total # OBS reflects the total number of observations that have a specific media (displayed in the row) before the policy change

The columns denote the number of observations that now switch to the new media (column choice) -- i.e. switching away from the previously chosen media (on the same row).

The parentheticals reflect the row %

Current Period Impact - refers to media choices post-policy change when pre-policy change print was being consumed

AFTER / BEFORE	***Total # OBS***	**Outside good (%)**	**Computer (%)**	**TV (%)**	**Radio (%)**	**Computer & TV (%)**	**Computer & Radio (%)**	**TV & Radio (%)**	**Computer & TV & Radio (%)**
Print	234	126 (53.8)	77 (32.9)	23 (9.8)	8 (3.4%)	0	0	0	0
Radio & Print	40	0	22 (55.0)	0	0	3 (7.5)	15 (37.5)		0
TV & Print	175	0	1 (0.6)	29 (16.6 %)	0	77 (44.0)	0	68 (38.9)	0
TV & Radio & Print	85	0	0	0	11 (12.9)	0	54 (63.5)	20 (23.5)	0
Computer & Print	30	0	0	0	0	0	0	0	30 (100)
Computer & Radio & Print	21	0	0	0	0	0	0	0	21 (100)
Computer & TV & Print	18	0	0	0	0	0	0	0	18 (100)
Computer & TV & Radio & Print	76	0	0	0	0	0	0	0	76 (100)

TABLE 7: Notations and Variables for the Latent Trait Model

Symbol	Meaning	Number[a]	Dimension[b]
T, S, and K	Number of persons, person-level parameters, and subtrees, respectively	1	
J_k	Number of categories for subtree or category system k	K	
C_{kj}	Category j of subtree or category system k	$\sum_{k=1}^{K} J_k$	
n_{kjt}	Data: Number of times person t responded with category C_{kj}	$T\sum_{k=1}^{K} J_k$	
$\mathbf{n}_t$	Vector containing the n_{kjt} for person t	T	$\sum_{k=1}^{K} J_k$
I_{kj}	Number of paths ending in category C_{kj}	$\sum_{k=1}^{K} J_k$	
B_{kji}	i-th path ending in category C_{kj}	$\sum_{k=1}^{K}\sum_{j=1}^{J_k} I_{kj}$	
m_{kjit}	Augmented Data: Number of times path B_{kji} was followed by person t	$T\sum_{k=1}^{K}\sum_{j=1}^{J_k} I_{kj}$	
$\mathbf{m}$	Vector containing the m_{kjit}	1	$T\sum_{k=1}^{K}\sum_{j=1}^{J_k} I_{kj}$
θ_s	Person-level parameter	S	
$\boldsymbol{\theta}_t$	Vector containing person t's parameter values θ_{st}	T	S
a_{skji}	Number of branches on path B_{kji} assigned to θ_s	$S\sum_{k=1}^{K}\sum_{j=1}^{J_k} I_{kj}$	
b_{skji}	Number of branches on path B_{kji} assigned to $1-\theta_s$	$S\sum_{k=1}^{K}\sum_{j=1}^{J_k} I_{kj}$	
α_s	Probit-transformed person-level parameter	S	
$\boldsymbol{\alpha}_t$	Vector containing person t's parameter values α_{st}	T	S
μ	Population-level (prior) mean of $\boldsymbol{\alpha}_t$	1	S
Σ	Population-level (prior) covariance matrix of $\boldsymbol{\alpha}_t$	1	$S \times S$
$\boldsymbol{\lambda}$	Vector of scale parameters	1	S
Q	Unscaled population-level covariance matrix	1	$S \times S$
n_{st}	Number of nodes labeled θ_s across trees traversed by person t	$T \times S$	
Z_{slt}	Random variate associated with a θ_s node in one of t's tree traversals	$\sum_{t=1}^{T}\sum_{s=1}^{S} n_{st}$	
$\mathbf{Z}_t$	Vector containing person t's random variates Z_{slt}	T	$\sum_{s=1}^{S} n_{st}$
n_{st}^{+}	Number of Z_{slt} constrained to be positive for θ_s and person t	$T \times S$	
n_{st}^{-}	Number of Z_{slt} constrained to be less or equal to zero for θ_s and person t	$T \times S$	
β_{st}	Person-specific deviation of α_{st} from the population-level mean μ_s	$T \times S$	
γ_{st}	Unscaled version of β_{st}	$T \times S$	
$\boldsymbol{\beta}_t$ and $\boldsymbol{\gamma}_t$	Vectors containing person t's β_{st} and γ_{st}, respectively	T	S

[a]Number of different elements of this kind.

[b]Dimension of vector or matrix where applicable.

TABLE 8: Summary Statistics for Product Category Consumption

Category	Sample Size	No. of Types.	No. of Brands	No. of Products	% of Type Polygamy	% of Brand Polygamy	% of Product Polygamy
Cell Phone Plans	3,327	3	9	12	0.0%	3.7%	76.7%
Financial Investments	3,014	3	13	-	37.8%	44.9%	-
Automobile	3,646	8	42	477	50.8%	49.8%	61.3%
Soft Drinks	4,452	4	4	96	76.5%	72.8%	89%
Media Consumption	4,218	4	4	-	93.1%	95.8%	-

TABLE 9: Parameter Estimates from the Latent Trait Model

Category (Tree)	θ_c (PM[a])	SD[b]	θ_b (PM[a])	SD[b]	θ_p (PM[a])	SD[b]
Cell Phone Plans	.042	.057	.048	.098	.840	.201
Financial Investments	.242	.193	.345	.165	-	-
Automobiles	.537	.235	.515	.257	.731	.239
Soft Drinks	.884	.713	.872	.293	.948	.246
Media Consumption	.912	.925	.957	.938	-	-

Note: [a]PM = Posterior Median (mean across simulated data sets).
[b]25 Posterior Percentile
[C]75 Posterior Percentile
[d]Standard Deviation of posterior (mean across simulated data sets).

TABLE 10: Model Fit and Comparison

Category	DIC	
	Latent Trait	Latent Class
Financial Investments	**3907.75**	5307.94
Automobile	**2953.54**	9063.23
Cell Phone Plans	**3384.22**	4665.33
Soft Drinks	**5422.93**	10822.35
Media Consumption	**5821.31**	10239.43

TABLE 11: Parameter Estimates from the Latent Class Model

Financial Investments (2 classes)				
Parameter	Class1, Weight 0.579		Class 2, Weight 0.421	
	Coefficient	95% CI	Coefficient	95% CI
θ_1	0.000	[-0.049 0.049]	0.893	[0.878 0.908]
θ_2	0.043	[0.038 0.047]	1.000	[0.941 1.059]
Goodness-of-Fit	statistics:			
M1	0	M2	0	
S1	0.279	S2	0.759	
Automobile (2 classes)				
Parameter	Class1, Weight 0.600		Class 2, Weight 0.400	
	Coefficient	95% CI	Coefficient	95% CI
θ_1	0.847	[0.831 0.864]	0.000	[-0.052 0.052]

θ_2	0.830	[0.813 0.847]	0.000	[-0.052 0.052]
θ_3	1.000	[0.957 1.043]	0.034	[0.019 0.05]
Goodness-of-Fit				
M1	0	M2	0.000	
S1	0.774	S2	1.508	
Soft Drinks (2 classes)				
Parameter	Class1, Weight 0.741		Class 2, Weight 0.259	
	Coefficient	95% CI	Coefficient	95% CI
θ_1	0.885	[0.87 0.9]	0.422	[0.378 0.466]
θ_2	0.982	[0.957 1.007]	0.000	[-0.108 0.108]
θ_3	1.000	[0.958 1.042]	0.575	[0.525 0.624]
Goodness-of-Fit				
M1	0	M2	0.002	
S1	0.149	S2	0.832	
Cell Phone Plans (Not Enough Variations to Distinguish Multiple Segments)				
Media Consumption (Not Enough Variations to Distinguish Multiple Segments)				

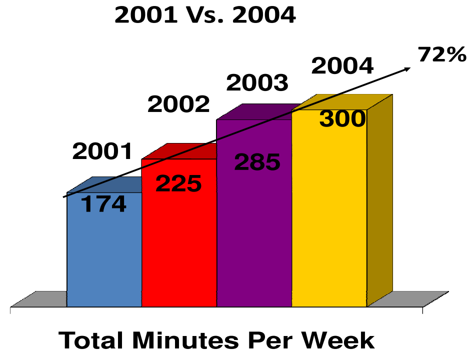
TABLE 12: Results from Latent Trait Segmentation

Parameters	Class 1: Polygamist (19.2%)	Class 2: Wanderer (71.5%)	Class 3: Monogamist (9.3%)
θ_1_cell	0.803	0.466	0.041
θ_2_cell	0.763	0.393	0.038
θ_3_cell	0.872	0.000	0.000
θ_1_auto	0.863	0.469	0.304
θ_2_auto	0.885	0.469	0.284
θ_3_auto	0.892	0.552	0.000
θ_1_finance	1.036	0.478	0.165
θ_2_finance	1.115	0.489	0.210
θ_1_softdrinks	1.182	0.571	0.000
θ_2_softdrinks	0.896	0.637	0.000
θ_3_softdrinks	0.948	0.000	0.000
θ_1_media	0.917	0.566	0.341
θ_2_media	0.964	0.593	0.238

TABLE 13: Out-of-Sample Prediction Results

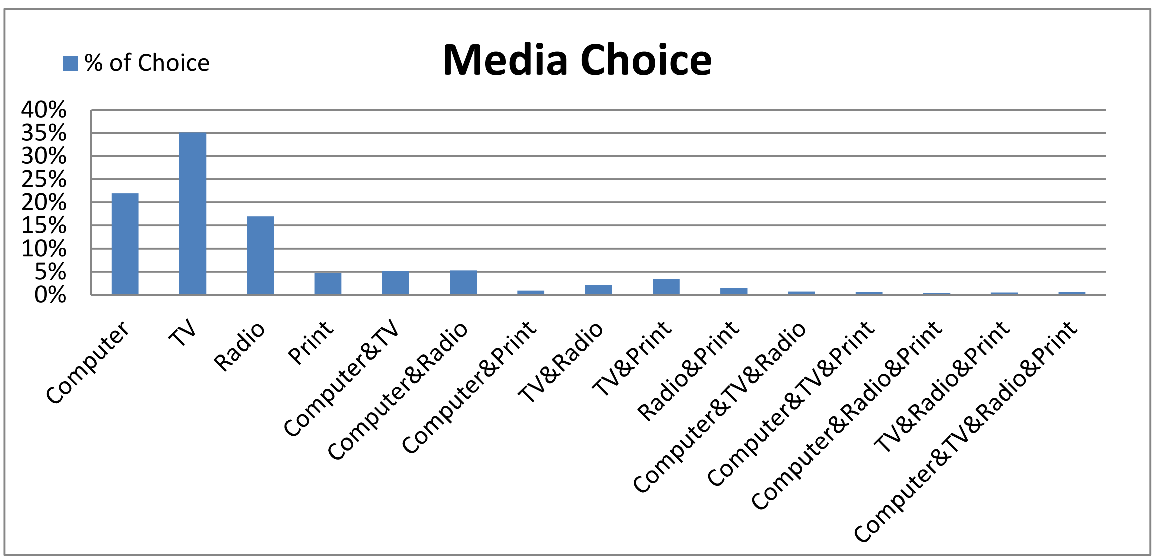
Segment	Polygamist	Wanderer	Monogamist
Customer-based Prediction	74.7%	61.2%	65.9%
Product-based Prediction	65.4%	52.0%	63.3%

FIGURE 1: Growth in Joint Consumption of Television and Computer Media



Source: Universal McCann's Media in Minds Survey2001-2004.

FIGURE 2: Summary Statistics for Single and Multimedia Consumption



"Print" includes newspaper and other publications such as magazines

FIGURE 3: Multinomial Process Tree Representation of the Latent Trait Model

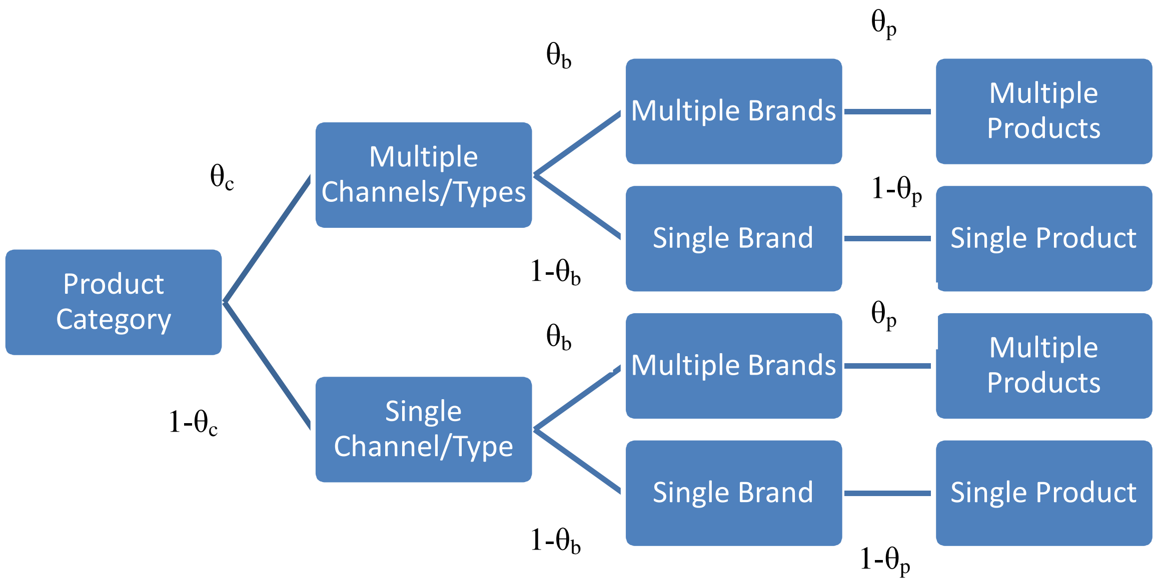


FIGURE 4: Multinomial Process Tree Representation of the Latent Trait Model with Augmented Data

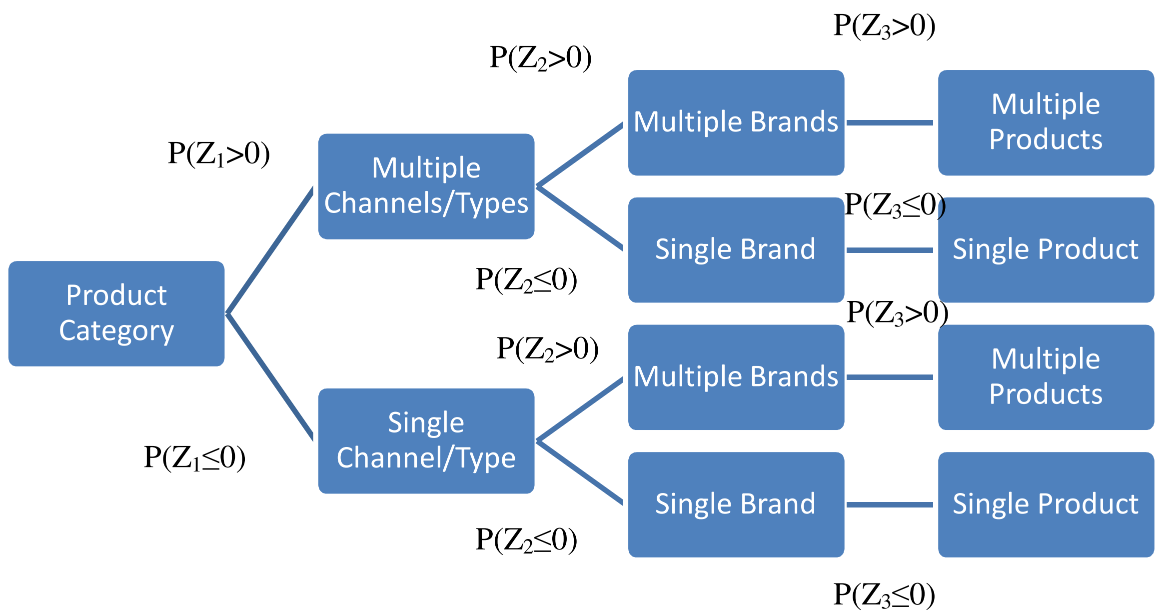


FIGURE 5: Sample Data Structure (Media Consumption Category)

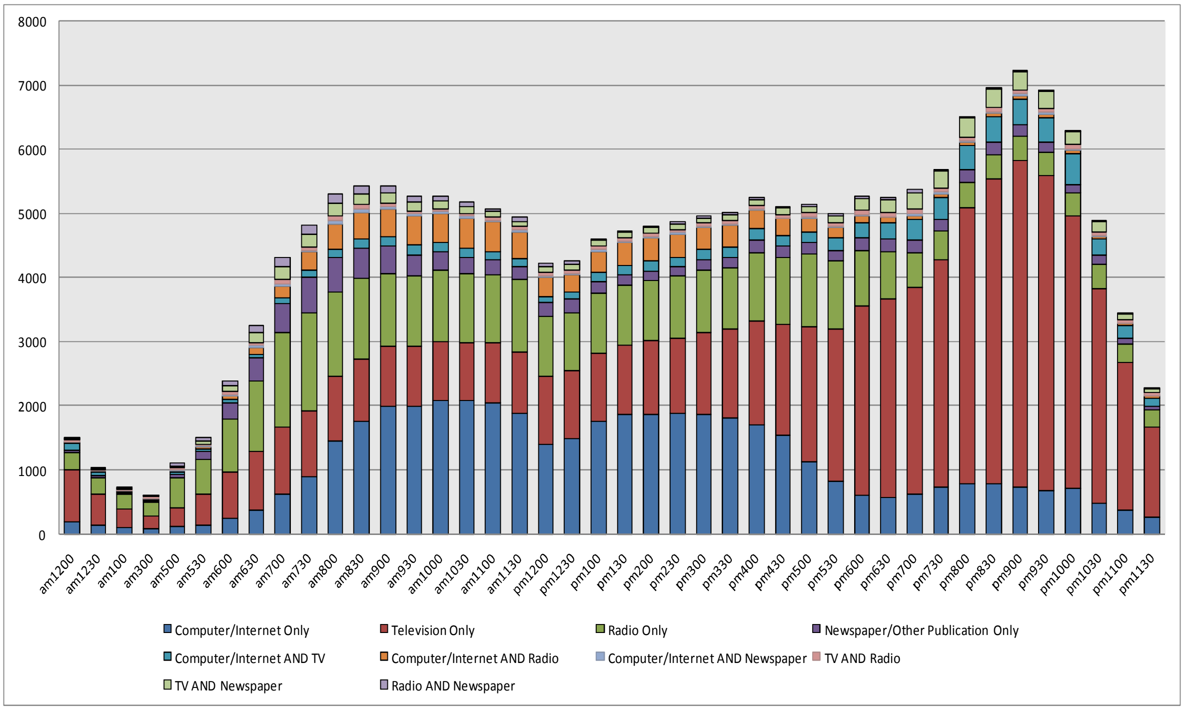
	full_day	full_time	id	multiplex_~l	multiplex_~d
74	2	pm430	1146945	0	0
75	2	pm500	1146945	0	0
76	2	pm530	1146945	0	0
77	2	pm600	1146945	0	0
78	2	pm630	1146945	0	0
79	2	pm700	1146945	0	0
80	2	pm730	1146945	0	0
81	2	pm800	1146945	0	0
82	2	pm830	1146945	0	0
83	2	pm900	1146945	0	0
84	2	pm930	1146945	0	0
85	3	am100	1146945	0	0
86	3	am1000	1146945	1	1
87	3	am1030	1146945	1	3
88	3	am1100	1146945	1	1
89	3	am1130	1146945	1	5
90	3	am1200	1146945	0	0
91	3	am1230	1146945	0	0

Disaggregate Choice Data

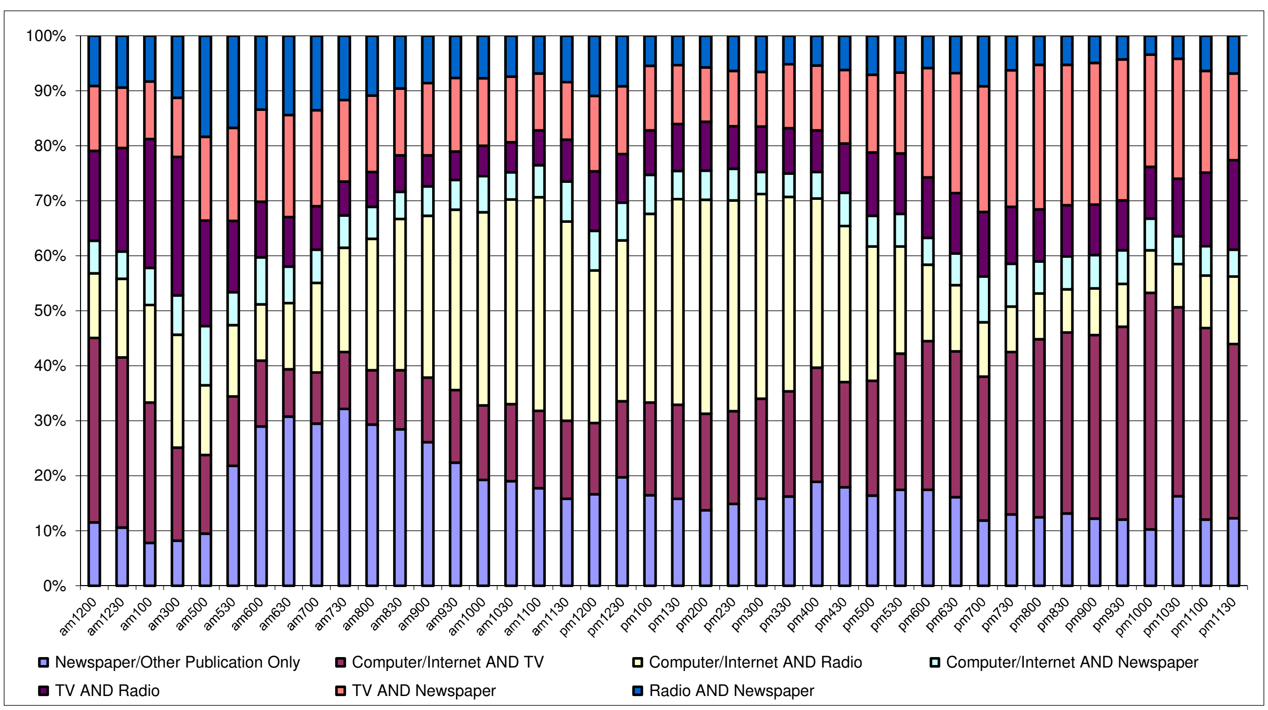
id	Channel Polygamy	Brand Polygamy
1128380	11	24
1128381	2	10
1128396	1	7
1128397	4	8
1128406	17	21
1128407	2	2
1128417	4	5
1128427	5	9
1128497	8	13
1128498	16	26

Aggregate Count Data

A1.FIGURE 1: MEDIA CHOICES BY TIME-OF-DAY (AGGREGATED ACROSS HOUSEHOLDS)



A1.FIGURE 2: AVERAGESHARE OF MEDIA USAGE BY TIME-OF-DAY AT THE HOUSEHOLD LEVEL



A1.TABLE 1: Media Choices by Day-Of-Week (Aggregated Across Households)

Day of week	Internet Only	TV Only	Radio Only	Print Only	Internet AND TV	Internet AND Radio	Internet AND Print	TV AND Radio	TV AND Print	Radio AND Print
Sunday	1,234	5,663	1,558	901	605	155	88	234	655	188
Monday	3,510	4,874	2,321	604	990	869	169	327	477	233
Tuesday	3,702	4,489	2,445	578	740	941	145	310	461	220
Wednesday	3,648	4,435	2,451	569	685	902	128	287	472	201
Thursday	3,744	4,268	2,504	579	674	968	134	292	424	189
Friday	3,475	4,210	2,442	559	613	875	118	282	405	168
Saturday	1,293	5,026	2,202	637	589	233	96	222	378	188
Grand Total	20,606	32,965	15,923	4,427	4,896	4,943	878	1,954	3,272	1,387

Cells contain the number of households

A1.TABLE 2: Results from Single Discrete-Continuous Model of Demand (Hanemann Model)

Structural Parameters	COMPUTER	TV	RADIO	PRINT
Baseline Preference Parameters	Coefficient(T-Stat)	Coefficient(T-Stat)	Coefficient(T-Stat)	Coefficient(T-Stat)
Intercept	-4.415***	-3.088***	-4.105***	-5.582***
	(-42.14)	(-54.29)	(-43.02)	(-34.2)
Attention Levels For				
Computer	3.794***	-0.218***	-0.301***	-0.249***
	(46.57)	(-5.36)	(-5.17)	(-2.36)
TV	-0.031	1.353***	-0.087***	0.089***
	(-1.05)	(111.47)	(-3.21)	(3.31)
Radio	-0.047***	-0.029	1.457***	0.118***
	(-2.33)	(-1.1)	(72.36)	(3.62)
Print	-0.657***	0.34***	0.463***	2.494***
	(-2.85)	(14.18)	(14.89)	(47.84)
Media Penetration				
Cable	-0.046**	0.058***	-0.001	0.039
	(-1.82)	(2.27)	(-0.02)	(0.88)
PC at Home	0.034	-0.062**	-0.23***	0.127
	(0.58)	(-1.82)	(-4.08)	(1.63)
PC at Work	0.091***	-0.079***	0.001	-0.009
	(2.61)	(-3.03)	(0.03)	(-0.18)
Consumer Demographics				
Age < 35	0.024	-0.076***	0.083	-0.084
	(0.38)	(-1.97)	(1.51)	(-1.24)
Age > 65	0.063*	0.081***	-0.029	0.254***
	(1.68)	(2.49)	(-0.51)	(4.22)
Hhincome < 35k	0.050	0.096***	0.097	0.003
	(1.17)	(2.62)	(1.61)	(0.04)
HHsize <= 3	-0.067**	-0.014	-0.093	-0.093
	(-1.84)	(-0.34)	(-1.28)	(-1.08)
Hhsize>=8	-0.014	0.011	0.029	0.021
	(-0.14)	(0.27)	(0.38)	(0.34)
Female	0.028	-0.051***	0.036	-0.038
	(0.89)	(-2.05)	(0.87)	(-0.85)
Urban	-0.059	-0.013	0.032	0.077
	(-1.21)	(-0.4)	(0.66)	(1.35)
Day of the Week (Saturday as Base)				
Sunday	-0.012	0.048***	-0.097***	0.187***
	(-0.44)	(2.43)	(-3.16)	(3.85)
Monday	0.277***	-0.0097801	0.08***	0.262***
	(8.58)	(-0.43)	(2.3)	(5.15)
Tuesday	0.153***	-0.021147	0.137***	0.207***
	(4.08)	(-0.96)	(4.14)	(4.09)
Wednesday	0.156***	-0.064***	0.142***	0.157***
	(4.02)	(-2.88)	(4.33)	(3.19)
Thursday	0.185***	-0.069***	0.119***	0.185***
	(5.95)	(-3.27)	(3.82)	(3.9)
Friday	0.154***	-0.099***	0.129***	0.083**
	(6.56)	(-4.69)	(4.26)	(1.91)
Time-of-the-Day (am1200 as base)				
am100	-0.943***	-1.075***	-0.462***	-1.465***
	(-9.38)	(-17.68)	(-7.42)	(-6.3)
am300	-1.177***	-1.333***	-0.484***	-1.695***

	(-8.02)	(-16.78)	(-7.04)	(-7.63)
am500	-0.123	-0.217***	0.671***	1.015***
	(-1.06)	(-3.29)	(8.13)	(6.02)
am600	0.603***	0.261***	1.135***	1.786***
	(6.97)	(4.87)	(15.3)	(11.41)
am700	0.972***	0.346***	1.365***	2.008***
	(11.34)	(7.04)	(19.64)	(13.3)
am800	0.934***	0.22***	1.143***	1.946***
	(11.27)	(4.54)	(16.47)	(12.89)
am900	0.721***	0.053	0.91***	1.709***
	(8.81)	(1.1)	(13.59)	(11.34)
am1000	0.66***	-0.042	0.754***	1.413***
	(8.03)	(-0.88)	(11.08)	(9.13)
am1100	0.627***	-0.062	0.766***	1.253***
	(7.47)	(-1.3)	(11.17)	(7.98)
pm1200	0.681***	0.128***	0.883***	1.401***
	(7.83)	(2.53)	(12.32)	(9.04)
pm100	0.704***	0.12***	0.778***	1.352***
	(8.46)	(2.57)	(11.21)	(8.88)
pm200	0.67***	0.128***	0.846***	1.018***
	(8.06)	(2.79)	(11.99)	(6.42)
pm300	0.723***	0.216***	0.923***	1.143***
	(8.7)	(4.74)	(13.27)	(7.39)
pm400	0.806***	0.321***	1.029***	1.268***
	(9.86)	(7.12)	(14.58)	(8.39)
pm500	0.902***	0.493***	1.207***	1.348***
	(11.2)	(11)	(17.32)	(9.13)
pm600	0.675***	0.577***	0.959***	1.267***
	(7.87)	(13.13)	(13.35)	(8.22)
pm700	0.717***	0.618***	0.86***	1.303***
	(7.24)	(14.02)	(11.17)	(8.2)
pm800	0.9***	0.787***	0.701***	1.286***
	(10.19)	(18.35)	(9.27)	(8.34)
pm900	0.9***	0.874***	0.643***	1.291***
	(10.23)	(20.69)	(8.79)	(8.41)
pm1000	1.072***	0.915***	0.626***	1.264***
	(13.83)	(21.58)	(8.79)	(8.46)
pm1100	0.484***	0.656***	0.329***	0.921***
	(6.92)	(16.4)	(4.9)	(7.07)

A1.TABLE 3 Results from Our Main Model (Half-hour Time Interval)

Structural Parameters	COMPUTER	TV	RADIO	PRINT
Baseline Preference Parameters	Coefficient(T-Stat)	Coefficient(T-Stat)	Coefficient(T-Stat)	Coefficient(T-Stat)
Intercept	-8.748***	-6.633***	-8.08***	-9.228***
	(-111.83)	(-140.37)	(-105.05)	(-73.24)
Attention Levels For				
Computer	8.269***	-0.381***	-0.439***	-0.309***
	(382.49)	(-23.73)	(-23.82)	(-10.95)
TV	-0.149***	2.11***	-0.086***	0.148***
	(-16.89)	(295.35)	(-7.43)	(11.6)
Radio	-0.127***	-0.086***	2.539***	0.163***
	(-17.94)	(-9.06)	(255.26)	(11.79)
Print	-2.425***	0.235***	0.437***	3.375***
	(-81.39)	(21.29)	(28.68)	(155.37)
Media Penetration				
Cable	-0.073***	0.092***	0.052***	0.134***
	(-5.86)	(8.55)	(3.56)	(5.49)
PC at Home	0.0126	-0.083***	-0.19***	0.308***
	(0.71)	(-5.95)	(-10.05)	(9.7)
PC at Work	0.188***	-0.118***	-0.0184	0.0324
	(13.57)	(-10.73)	(-1.19)	(1.29)
Consumer Demographics				
Age < 35	0.0165	-0.144***	0.059***	-0.172***
	(0.96)	(-8.68)	(2.76)	(-4.44)
Age > 65	0.075***	0.145***	0.025	0.433***
	(3.86)	(10.53)	(1.16)	(15.03)
Hhincome < 35k	0.101***	0.147***	0.185***	-0.0391
	(4.84)	(9.75)	(8.58)	(-1.13)
HHsize <= 3	-0.058***	0.0148	-0.077***	-0.101***
	(-2.75)	(0.89)	(-3.2)	(-2.61)
Hhsize>=8	-0.184***	0.0196	0.0329	-0.0168
	(-8.63)	(1.13)	(1.43)	(-0.43)
Female	0.0013	-0.095***	0.032***	-0.083***
	(0.11)	(-9.29)	(2.23)	(-3.64)
Urban	-0.03***	-0.0145	-0.0018	0.077***
	(-2.03)	(-1.13)	(-0.1)	(2.78)
Day of the Week (Saturday as Base)				
Sunday	-0.0347	0.127***	-0.124***	0.384***
	(-1.12)	(7.14)	(-4.15)	(9.52)
Monday	0.14***	-0.0271	0.11***	0.229***
	(5.38)	(-1.49)	(4.09)	(5.45)
Tuesday	0.076***	-0.062***	0.167***	0.18***
	(2.93)	(-3.31)	(6.27)	(4.21)
Wednesday	0.061***	-0.14***	0.157***	0.095***
	(2.33)	(-7.48)	(5.88)	(2.19)
Thursday	0.118***	-0.124***	0.117***	0.155***
	(4.53)	(-6.58)	(4.37)	(3.58)
Friday	0.063***	-0.184***	0.158***	0.0706
	(2.39)	(-9.7)	(5.89)	(1.6)
Time-of-the-Day (am1200 as base)				
am100	-1.228***	-2.02***	-1.153***	-2.202***
	(-10.03)	(-26.58)	(-11.21)	(-10.43)
am300	-1.441***	-2.326***	-1.068***	-3.248***
	(-10.7)	(-27.25)	(-10.89)	(-13.18)

am500	-0.395***	-0.539***	0.372***	-0.314**
	(-3.41)	(-7.54)	(4.07)	(-1.94)
am530	-0.218***	-0.401***	0.617***	0.576***
	(-1.97)	(-6)	(6.97)	(4.01)
am600	0.187**	-0.179***	0.949***	0.490***
	(1.91)	(-2.93)	(11.29)	(3.56)
am630	0.239***	-0.0528	1.209***	1.176***
	(2.59)	(-0.91)	(14.95)	(9.16)
am700	0.374***	0.0722	1.287***	1.511***
	(4.37)	(1.28)	(16.21)	(12)
am730	0.418***	0.0694	1.359***	0.911***
	(5.1)	(1.23)	(17.31)	(7.15)
am800	0.426***	0.0393	1.082***	1.322***
	(5.39)	(0.7)	(13.62)	(10.53)
am830	0.439***	-0.0107	1.007***	1.516***
	(5.62)	(-0.19)	(12.7)	(11.99)
am900	0.383***	-0.0315	0.87***	0.867***
	(4.93)	(-0.54)	(10.84)	(6.7)
am930	0.407***	-0.0649	0.779***	0.994***
	(5.24)	(-1.12)	(9.74)	(7.59)
am1000	0.419***	-0.0437	0.755***	0.587***
	(5.41)	(-0.75)	(9.44)	(4.36)
am1030	0.388***	-0.083	0.742***	0.781***
	(5.01)	(-1.41)	(9.2)	(5.77)
am1100	0.384***	-0.0321	0.805***	0.923***
	(4.96)	(-0.55)	(10.03)	(6.88)
am1130	0.368***	-0.0174	0.834***	0.293***
	(4.71)	(-0.3)	(10.34)	(2.1)
pm1200	-0.235***	-0.404***	-0.0658	-0.998***
	(-2.1)	(-5.97)	(-0.64)	(-4.98)
pm1230	0.399***	0.16***	0.728***	0.75***
	(5.09)	(2.82)	(8.84)	(5.44)
pm100	0.394***	0.153***	0.749***	0.296***
	(5.05)	(2.7)	(9.18)	(2.06)
pm130	0.371***	0.138***	0.797***	0.395***
	(4.75)	(2.43)	(9.74)	(2.71)
pm200	0.405***	0.17***	0.808***	0.458***
	(5.19)	(3.02)	(9.92)	(3.24)
pm230	0.384***	0.259***	0.773***	0.113
	(4.91)	(4.66)	(9.43)	(0.78)
pm300	0.411***	0.268***	0.793***	0.545***
	(5.26)	(4.89)	(9.73)	(3.84)
pm330	0.4***	0.333***	0.887***	0.793***
	(5.09)	(6.18)	(10.87)	(5.83)
pm400	0.393***	0.357***	0.924***	0.257**
	(4.97)	(6.68)	(11.34)	(1.83)
pm430	0.347***	0.473***	1.073***	0.518***
	(4.29)	(9.07)	(13.19)	(3.68)
pm500	0.348***	0.567***	1.026***	0.754***
	(4.21)	(11.03)	(12.55)	(5.53)
pm530	0.337***	0.759***	0.897***	0.1507
	(3.99)	(15.22)	(10.74)	(1.07)
pm600	0.357***	0.774***	0.815***	0.585***
	(4.21)	(15.6)	(9.66)	(4.27)
pm630	0.328***	0.901***	0.713***	0.902***
	(3.92)	(18.34)	(8.24)	(6.93)
pm700	0.436***	0.891***	0.649***	0.2153

	(5.26)	(18.21)	(7.33)	(1.58)
pm730	0.415***	1.11***	0.588***	0.584***
	(5.04)	(23.11)	(6.48)	(4.29)
pm800	0.473***	1.198***	0.601***	0.873***
	(5.76)	(25.08)	(6.66)	(6.66)
pm830	0.486***	1.257***	0.595***	0.262**
	(5.89)	(26.44)	(6.54)	(1.92)
pm900	0.514***	1.243***	0.506***	0.589***
	(6.2)	(26.12)	(5.55)	(4.3)
pm930	0.397***	1.227***	0.589***	0.0834
	(4.83)	(25.56)	(6.47)	(0.59)
pm1000	0.36***	0.984***	0.387***	0.417***
	(4.16)	(20.09)	(4.13)	(2.91)
pm1030	0.269***	0.758***	0.264***	0.278**
	(3)	(14.88)	(2.77)	(1.89)
pm1100	0.1281	0.376***	0.0516	-0.2671
	(1.34)	(6.92)	(0.52)	(-1.63)
pm1130	0.33***	0.2***	0.902***	0.704***
	(4.14)	(3.54)	(11.1)	(5.17)
am1230	0.362***	0.16***	0.752***	0.914***
	(4.56)	(2.8)	(9.09)	(6.81)
Satiation Parameters				
Alpha	0.999***	0.449***	0.210***	0.291***
	(5.05)	(12.34)	(30.11)	(10.32)
	note: Alpha for the outside good is approximately 1			
Gamma	1 for all inside goods, 0 for the outside good			
Random Components (Standard Deviations)				
RC_Baseline	0.118***	0.041	0.020	0.207***
	(14.85)	(1.14)	(0.99)	(5.80)
RC_Satiation	0.708***	2.388***	0.043	1.158
	(16.27)	(15.82)	(0.32)	(1.45)
EC_Baseline	0.004	(1 EC in the baseline utility)		
	(0.87)			
EC_Satiation	0.022	(1 EC among satiation parameters)		
	(1.23)			
	RC: Random Coefficients; EC: Error Components			

A1.TABLE 4A: Marginal Rate of Substitution at the Point of Zero Consumption: Baseline Constants

	Computer	TV	Radio	Print	Computer&TV	Computer&Radio	Computer&Print	TV&Radio	TV&Print	Radio&Print	Computer&TV&Radio	Computer&TV&Print	Computer&Radio&Print	TV&Radio&Print	Computer&TV&Radio&Print
Computer	-	0.233	0.729	3.793	1.11E+03	3.47E+03	1.81E+04	8.10E+02	4.22E+03	1.32E+04	3.86E+06	2.01E+07	6.28E+07	1.46E+07	6.98E+10
TV	4.289	-	3.126	16.268	4.77E+03	1.49E+04	7.76E+04	3.47E+03	1.81E+04	5.65E+04	1.66E+07	8.62E+07	2.69E+08	6.28E+07	2.99E+11
Radio	1.372	0.320	-	5.204	1.53E+03	4.77E+03	2.48E+04	1.11E+03	5.78E+03	1.81E+04	5.30E+06	2.76E+07	8.62E+07	2.01E+07	9.58E+10
Print	0.264	0.061	0.192	-	2.93E+02	9.16E+02	4.77E+03	2.14E+02	1.11E+03	3.47E+03	1.02E+06	5.30E+06	1.66E+07	3.86E+06	1.84E+10
Computer&TV	9.00E-04	2.10E-04	6.56E-04	3.41E-03	-	3.13E+00	1.63E+01	7.29E-01	3.79E+00	1.19E+01	3.47E+03	1.81E+04	5.65E+04	1.32E+04	6.28E+07
Computer&Radio	2.88E-04	6.71E-05	2.10E-04	1.09E-03	3.20E-01	-	5.20E+00	2.33E-01	1.21E+00	3.79E+00	1.11E+03	5.78E+03	1.81E+04	4.22E+03	2.01E+07
Computer&Print	5.53E-05	1.29E-05	4.03E-05	2.10E-04	6.15E-02	1.92E-01	-	4.48E-02	2.33E-01	7.29E-01	2.14E+02	1.11E+03	3.47E+03	8.10E+02	3.86E+06
TV&Radio	1.23E-03	2.88E-04	9.00E-04	4.68E-03	1.37E+00	4.29E+00	2.23E+01	-	5.20E+00	1.63E+01	4.77E+03	2.48E+04	7.76E+04	1.81E+04	8.62E+07
TV&Print	2.37E-04	5.53E-05	1.73E-04	9.00E-04	2.64E-01	8.24E-01	4.29E+00	1.92E-01	-	3.12	9.16E+02	4.77E+03	1.49E+04	3.47E+03	1.66E+07
Radio&Print	7.59E-05	1.77E-05	5.53E-05	2.88E-04	8.43E-02	2.64E-01	1.37E+00	6.15E-02	3.20E-01	-	2.93E+02	1.53E+03	4.77E+03	1.11E+03	5.30E+06
Computer&TV&Radio	2.59E-07	6.04E-08	1.89E-07	9.82E-07	2.88E-04	9.00E-04	4.68E-03	2.10E-04	1.09E-03	3.41E-03	-	5.20E+00	1.63E+01	3.79E+00	1.81E+04
Computer&TV&Print	4.98E-08	1.16E-08	3.63E-08	1.89E-07	5.53E-05	1.73E-04	9.00E-04	4.03E-05	2.10E-04	6.56E-04	1.92E-01	-	3.13E+00	7.29E-01	3.47E+03
Computer&Radio&Print	1.59E-08	3.71E-09	1.16E-08	6.04E-08	1.77E-05	5.53E-05	2.88E-04	1.29E-05	6.71E-05	2.10E-04	6.15E-02	3.20E-01	-	2.33E-01	1.11E+03
TV&Radio&Print	6.83E-08	1.59E-08	4.98E-08	2.59E-07	7.59E-05	2.37E-04	1.23E-03	5.53E-05	2.88E-04	9.00E-04	2.64E-01	1.37E+00	4.29E+00	-	4.77E+03
Computer&TV&Radio&Print	1.43E-11	3.34E-12	1.04E-11	5.43E-11	1.59E-08	4.98E-08	2.59E-07	1.16E-08	6.04E-08	1.89E-07	5.53E-05	2.88E-04	9.00E-04	2.10E-04	-

Note: The column represents the media option that will be SUBSTITUTED FOR, The row represents the media option that will be SUBSTITUTED TO. The highlighted cells indicate strong substitution effects such that consumers will give up more quantities of the column media to obtain one additional unit of the row media.

A1.TABLE 4B: Marginal Rate of Substitution at the Point OF Zero Consumption: Attention for Computer

	Computer	TV	Radio	Print	Computer&TV	Computer&Radio	Computer&Print	TV&Radio	TV&Print	Radio&Print	Computer&TV&Radio	Computer&TV&Print	Computer&Radio&Print	TV&Radio&Print	Computer&TV&Radio&Print
Computer	-	3215.70	3501.69	2919.59	1.40	1.52	1.27	4897.11	4083.05	4446.18	2.13	1.78	1.93	6217.98	2.70
TV	0.00	-	1.09	0.91	0.00	0.00	0.00	1.52	1.27	1.38	0.00	0.00	0.00	1.93	0.00
Radio	0.00	0.92	-	0.83	0.00	0.00	0.00	1.40	1.17	1.27	0.00	0.00	0.00	1.78	0.00
Print	0.00	1.10	1.20	-	0.00	0.00	0.00	1.68	1.40	1.52	0.00	0.00	0.00	2.13	0.00
Computer&TV	0.72	*2299.39*	*2503.89*	*2087.66*	-	1.09	0.91	3501.69	2919.59	3179.25	1.52	1.27	1.38	4446.18	1.93
Computer&Radio	0.66	*2111.60*	*2299.39*	*1917.16*	0.92	-	0.83	3215.70	2681.15	2919.59	1.40	1.17	1.27	4083.05	1.78
Computer&Print	0.79	*2532.60*	*2757.83*	*2299.39*	1.10	1.20	-	3856.83	3215.70	3501.69	1.68	1.40	1.52	4897.11	2.13
TV&Radio	0.00	0.66	0.72	0.60	0.00	0.00	0.00	-	0.83	0.91	0.00	0.00	0.00	1.27	0.00
TV&Print	0.00	0.79	0.86	0.72	0.00	0.00	0.00	1.20	-	1.09	0.00	0.00	0.00	1.52	0.00
Radio&Print	0.00	0.72	0.79	0.66	0.00	0.00	0.00	1.10	0.92	-	0.00	0.00	0.00	1.40	0.00
Computer&TV&Radio	0.47	*1509.90*	*1644.18*	*1370.87*	0.66	0.72	0.60	2299.39	1917.16	2087.66	-	0.83	0.91	2919.59	1.27
Computer&TV&Print	0.56	*1810.94*	*1971.99*	*1644.18*	0.79	0.86	0.72	2757.83	2299.39	2503.89	1.20	-	1.09	3501.69	1.52
Computer&Radio&Print	0.52	*1663.04*	*1810.94*	*1509.90*	0.72	0.79	0.66	2532.60	2111.60	2299.39	1.10	0.92	-	3215.70	1.40
TV&Radio&Print	0.00	0.52	0.56	0.47	0.00	0.00	0.00	0.79	0.66	0.72	0.00	0.00	0.00	-	0.00
Computer&TV&Radio&Print	0.37	*1189.16*	*1294.91*	*1079.66*	0.52	0.56	0.47	1810.94	1509.90	1644.18	0.79	0.66	0.72	2299.39	-

Note: The column represents the media option that will be SUBSTITUTED FOR, The row represents the media option that will be SUBSTITUTED TO. The highlighted cells indicate strong substitution effects such that consumers will give up more quantities of the column media to obtain one additional unit of the row media.

The underlined italicized estimates indicate potential complementarity, i.e., where multiplexing is preferred to single media option.

A1.TABLE 4C: Marginal Rate of Substitution at the Point OF Zero Consumption: Attention for Television

	Computer	TV	Radio	Print	Computer&TV	Computer&Radio	Computer&Print	TV&Radio	TV&Print	Radio&Print	Computer&TV&Radio	Computer&TV&Print	Computer&Radio&Print	TV&Radio&Print	Computer&TV&Radio&Print
Computer	-	0.12	1.05	0.86	0.12	1.10	0.90	0.13	0.11	0.94	0.14	0.11	0.99	0.12	0.12
TV	8.49	-	8.87	7.28	1.05	9.33	7.65	1.10	0.90	8.00	1.16	0.95	8.41	0.99	1.04
Radio	0.96	0.11	-	0.82	0.12	1.05	0.86	0.12	0.10	0.90	0.13	0.11	0.95	0.11	0.12
Print	1.17	0.14	1.22	-	0.14	1.28	1.05	0.15	0.12	1.10	0.16	0.13	1.16	0.14	0.14
Computer&TV	8.07	0.95	8.44	6.92	-	8.87	7.28	1.05	0.86	7.61	1.10	0.90	8.00	0.94	0.99
Computer&Radio	0.91	0.11	0.95	0.78	0.11	-	0.82	0.12	0.10	0.86	0.12	0.10	0.90	0.11	0.11
Computer&Print	1.11	0.13	1.16	0.95	0.14	1.22	-	0.14	0.12	1.05	0.15	0.12	1.10	0.13	0.14
TV&Radio	7.72	0.91	8.07	6.62	0.96	8.49	6.96	-	0.82	7.28	1.05	0.86	7.65	0.90	0.95
TV&Print	9.42	1.11	9.84	8.07	1.17	10.35	8.49	1.22	-	9.09	1.28	1.05	9.33	1.10	1.16
Radio&Print	1.06	0.12	1.11	0.91	0.13	1.17	0.96	0.14	0.11	-	0.14	0.12	1.05	0.12	0.13
Computer&TV&Radio	7.34	0.87	7.68	6.30	0.91	8.07	6.62	0.95	0.78	6.92	-	0.82	7.28	0.86	0.90
Computer&TV&Print	8.95	1.05	9.36	7.68	1.11	9.84	8.07	1.16	0.95	8.44	1.22	-	8.87	1.05	1.10
Computer&Radio&Print	1.01	0.12	1.05	0.87	0.12	1.11	0.91	0.13	0.11	0.95	0.14	0.11	-	0.12	0.12
TV&Radio&Print	8.57	1.01	8.95	7.34	1.06	9.42	7.72	1.11	0.91	8.07	1.17	0.96	8.49	-	1.05
Computer&TV&Radio&Print	8.15	0.96	8.51	6.98	1.01	8.95	7.34	1.05	0.87	7.68	1.11	0.91	8.07	0.95	-

Note: The column represents the media option that will be SUBSTITUTED FOR, The row represents the media option that will be SUBSTITUTED TO. The highlighted cells indicate strong substitution effects such that consumers will give up more quantities of the column media to obtain one additional unit of the row media.

The underlined italicized estimates indicate potential complementarity, i.e., where multiplexing is preferred to single media option.

A1.TABLE 4D: Marginal Rate of Substitution at the Point OF Zero Consumption: Attention for Radio

	Computer	TV	Radio	Print	Computer&TV	Computer&Radio	Computer&Print	TV&Radio	TV&Print	Radio&Print	Computer&TV&Radio	Computer&TV&Print	Computer&Radio&Print	TV&Radio&Print	Computer&TV&Radio&Print
Computer	-	0.95	0.10	0.79	1.04	0.11	0.87	0.10	0.82	0.09	0.11	0.90	0.09	0.09	0.10
TV	1.05	-	0.10	0.83	1.10	0.12	0.91	0.11	0.87	0.09	0.12	0.95	0.10	0.09	0.10
Radio	10.06	9.54	-	7.91	10.51	1.10	8.71	1.04	8.26	0.87	1.15	9.09	0.95	0.90	1.00
Print	1.27	1.21	0.13	-	1.33	0.14	1.10	0.13	1.04	0.11	0.15	1.15	0.12	0.11	0.13
Computer&TV	0.96	0.91	0.10	0.75	-	0.10	0.83	0.10	0.79	0.08	0.11	0.87	0.09	0.09	0.09
Computer&Radio	*9.13*	*8.67*	0.91	***7.18***	9.54	-	7.91	0.95	7.51	0.79	1.04	8.26	0.87	0.82	0.90
Computer&Print	*1.16*	*1.10*	0.11	0.91	1.21	0.13	-	0.12	0.95	0.10	0.13	1.04	0.11	0.10	0.11
TV&Radio	*9.62*	*9.13*	0.96	*7.57*	10.06	1.05	8.33	-	7.91	0.83	1.10	8.71	0.91	0.87	0.95
TV&Print	*1.22*	*1.16*	0.12	0.96	1.27	0.13	1.05	0.13	-	0.10	0.14	1.10	0.12	0.11	0.12
Radio&Print	*11.62*	*11.02*	***1.16***	*9.13*	12.14	1.27	10.06	1.21	9.54	-	1.33	10.51	1.10	1.04	1.15
Computer&TV&Radio	*8.74*	*8.30*	0.87	*6.88*	9.13	0.96	7.57	0.91	7.18	0.75	-	7.91	0.83	0.79	0.87
Computer&TV&Print	*1.11*	*1.05*	0.11	0.87	1.16	0.12	0.96	0.11	0.91	0.10	0.13	-	0.10	0.10	0.11
Computer&Radio&Print	*10.55*	*10.02*	*1.05*	*8.30*	11.02	1.16	9.13	1.10	8.67	0.91	1.21	9.54	-	0.95	1.04
TV&Radio&Print	*11.12*	*10.55*	*1.11*	*8.74*	11.62	1.22	9.62	1.16	9.13	0.96	1.27	10.06	1.05	-	1.10
Computer&TV&Radio&Print	*10.10*	*9.59*	*1.00*	*7.94*	10.55	1.11	8.74	1.05	8.30	0.87	1.16	9.13	0.96	0.91	-

Note: The column represents the media option that will be SUBSTITUTED FOR, The row represents the media option that will be SUBSTITUTED TO. The highlighted cells indicate strong substitution effects such that consumers will give up more quantities of the column media to obtain one additional unit of the row media.

The underlined italicized estimates indicate potential complementarity, i.e., where multiplexing is preferred to single media option.

A1.TABLE 4E: Marginal Rate of Substitution at the Point OF Zero Consumption: Attention for Print

	Computer	TV	Radio	Print	Computer&TV	Computer&Radio	Computer&Print	TV&Radio	TV&Print	Radio&Print	Computer&TV&Radio	Computer&TV&Print	Computer&Radio&Print	TV&Radio&Print	Computer&TV&Radio&Print
Computer	-	0.15	0.13	0.01	0.63	0.52	0.03	0.08	0.00	0.00	0.33	0.02	0.01	0.00	0.01
TV	6.55	-	0.83	0.04	4.10	3.42	0.18	0.52	0.03	0.02	2.14	0.11	0.09	0.01	0.06
Radio	7.86	1.20	-	0.05	4.91	4.10	0.21	0.63	0.03	0.03	2.56	0.13	0.11	0.02	0.07
Print	149.77	22.86	19.06	-	93.67	78.11	4.10	11.92	0.63	0.52	48.85	2.56	2.14	0.33	1.34
Computer&TV	1.60	0.24	0.20	0.01	-	0.83	0.04	0.13	0.01	0.01	0.52	0.03	0.02	0.00	0.01
Computer&Radio	1.92	0.29	0.24	0.01	1.20	-	0.05	0.15	0.01	0.01	0.63	0.03	0.03	0.00	0.02
Computer&Print	36.55	5.58	4.65	0.24	22.86	19.06	-	2.91	0.15	0.13	11.92	0.63	0.52	0.08	0.33
TV&Radio	12.56	1.92	1.60	0.08	7.86	6.55	0.34	-	0.05	0.04	4.10	0.21	0.18	0.03	0.11
TV&Print	239.46	36.55	30.48	1.60	149.77	124.89	6.55	19.06	-	0.83	78.11	4.10	3.42	0.52	2.14
Radio&Print	287.18	43.84	36.55	1.92	179.61	149.77	7.86	22.86	1.20	-	93.67	4.91	4.10	0.63	2.56
Computer&TV&Radio	3.07	0.47	0.39	0.02	1.92	1.60	0.08	0.24	0.01	0.01	-	0.05	0.04	0.01	0.03
Computer&TV&Print	58.45	8.92	7.44	0.39	36.55	30.48	1.60	4.65	0.24	0.20	19.06	-	0.83	0.13	0.52
Computer&Radio&Print	70.09	10.70	8.92	0.47	43.84	36.55	1.92	5.58	0.29	0.24	22.86	1.20	-	0.15	0.63
TV&Radio&Print	459.16	70.09	58.45	3.07	287.18	239.46	12.56	36.55	1.92	1.60	149.77	7.86	6.55	-	4.10
Computer&TV&Radio&Print	112.07	17.11	14.26	0.75	70.09	58.45	3.07	8.92	0.47	0.39	36.55	1.92	1.60	0.24	-

Note: The column represents the media option that will be SUBSTITUTED FOR, The row represents the media option that will be SUBSTITUTED TO. The highlighted cells indicate strong substitution effects such that consumers will give up more quantities of the column media to obtain one additional unit of the row media.

The underlined italicized estimates indicate potential complementarity, i.e., where multiplexing is preferred to single media option.

A1.TABLE 4F: Marginal Rate of Substitution at the Point OF Zero Consumption: Home PC Ownership

	Computer	TV	Radio	Print	Computer&TV	Computer&Radio	Computer&Print	TV&Radio	TV&Print	Radio&Print	Computer&TV&Radio	Computer&TV&Print	Computer&Radio&Print	TV&Radio&Print	Computer&TV&Radio&Print
Computer	-	1.14	1.40	0.87	1.09	1.34	0.84	1.53	0.95	1.17	1.46	0.91	1.12	1.28	1.22
TV	0.88	-	1.23	0.77	0.96	1.18	0.74	1.34	0.84	1.03	1.29	0.80	0.99	1.12	1.08
Radio	0.71	0.81	-	0.62	0.78	0.96	0.60	1.09	0.68	0.84	1.04	0.65	0.80	0.91	0.87
Print	1.15	1.30	1.61	-	1.25	1.54	0.96	1.75	1.09	1.34	1.67	1.04	1.29	1.46	1.40
Computer&TV	0.92	<u>*1.04*</u>	<u>*1.29*</u>	0.80	-	1.23	0.77	1.40	0.87	1.08	1.34	0.84	1.03	1.17	1.12
Computer&Radio	0.74	0.85	<u>*1.04*</u>	0.65	0.81	-	0.62	1.14	0.71	0.87	1.09	0.68	0.84	0.95	0.91
Computer&Print	<u>*1.20*</u>	<u>*1.36*</u>	<u>*1.68*</u>	<u>*1.04*</u>	1.30	1.61	-	1.82	1.14	1.40	1.75	1.09	1.34	1.53	1.46
TV&Radio	0.66	0.74	0.92	0.57	0.71	0.88	0.55	-	0.62	0.77	0.96	0.60	0.74	0.84	0.80
TV&Print	<u>*1.05*</u>	<u>*1.20*</u>	<u>*1.48*</u>	0.92	1.15	1.41	0.88	1.61	-	1.23	1.54	0.96	1.18	1.34	1.29
Radio&Print	0.85	0.97	<u>*1.20*</u>	0.74	0.93	1.15	0.71	1.30	0.81	-	1.25	0.78	0.96	1.09	1.04
Computer&TV&Radio	0.68	0.78	0.96	0.60	0.74	0.92	0.57	1.04	0.65	0.80	-	0.62	0.77	0.87	0.84
Computer&TV&Print	<u>*1.10*</u>	<u>*1.25*</u>	<u>*1.54*</u>	0.96	1.20	1.48	0.92	1.68	1.04	1.29	1.61	-	1.23	1.40	1.34
Computer&Radio&Print	0.89	<u>*1.01*</u>	<u>*1.25*</u>	0.78	0.97	1.20	0.74	1.36	0.85	1.04	1.30	0.81	-	1.14	1.09
TV&Radio&Print	0.78	0.89	<u>*1.10*</u>	0.68	0.85	1.05	0.66	1.20	0.74	0.92	1.15	0.71	0.88	-	0.96
Computer&TV&Radio&Print	0.82	0.93	<u>*1.15*</u>	0.71	0.89	1.10	0.68	1.25	0.78	0.96	1.20	0.74	0.92	1.04	-

Note: The column represents the media option that will be SUBSTITUTED FOR, The row represents the media option that will be SUBSTITUTED TO. The highlighted cells indicate strong substitution effects such that consumers will give up more quantities of the column media to obtain one additional unit of the row media.

The underlined italicized estimates indicate potential complementarity, i.e., where multiplexing is preferred to single media option.

A1.TABLE 4G: Marginal Rate of Substitution at the Point OF Zero Consumption: Cable TV Ownership

	Computer	TV	Radio	Print	Computer&TV	Computer&Radio	Computer&Print	TV&Radio	TV&Print	Radio&Print	Computer&TV&Radio	Computer&TV&Print	Computer&Radio&Print	TV&Radio&Print	Computer&TV&Radio&Print
Computer	-	0.87	0.94	0.88	0.92	1.00	0.93	0.87	0.81	0.87	0.92	0.86	0.93	0.81	0.86
TV	1.15	-	1.08	1.01	1.06	1.15	1.07	1.00	0.93	1.01	1.06	0.99	1.07	0.93	0.99
Radio	1.07	0.93	-	0.94	0.99	1.06	1.00	0.92	0.87	0.93	0.98	0.92	0.99	0.86	0.92
Print	1.14	0.99	1.07	-	1.06	1.14	1.06	0.99	0.92	1.00	1.05	0.98	1.06	0.92	0.98
Computer&TV	*1.08*	0.94	*1.01*	0.95	-	1.08	1.01	0.94	0.88	0.94	1.00	0.93	1.01	0.87	0.93
Computer&Radio	*1.00*	0.87	0.94	0.88	0.93	-	0.94	0.87	0.81	0.88	0.92	0.87	0.93	0.81	0.86
Computer&Print	*1.07*	0.93	*1.00*	0.94	0.99	1.07	-	0.93	0.87	0.94	0.99	0.92	1.00	0.87	0.92
TV&Radio	*1.16*	*1.00*	*1.08*	*1.01*	1.07	1.15	1.08	-	0.94	1.01	1.06	1.00	1.07	0.93	0.99
TV&Print	*1.23*	*1.07*	*1.16*	*1.08*	1.14	1.23	1.15	1.07	-	1.08	1.14	1.06	1.15	1.00	1.06
Radio&Print	*1.15*	*1.00*	*1.07*	*1.00*	1.06	1.14	1.07	0.99	0.93	-	1.06	0.99	1.06	0.92	0.98
Computer&TV&Radio	*1.09*	0.94	*1.02*	0.95	1.00	1.08	1.01	0.94	0.88	0.95	-	0.94	1.01	0.88	0.93
Computer&TV&Print	*1.16*	***1.01***	*1.09*	***1.02***	1.07	1.16	1.08	1.00	0.94	1.01	1.07	-	1.08	0.94	1.00
Computer&Radio&Print	*1.08*	0.93	*1.01*	0.94	1.00	1.07	1.00	0.93	0.87	0.94	0.99	0.93	-	0.87	0.92
TV&Radio&Print	*1.24*	***1.08***	*1.16*	***1.09***	1.15	1.23	1.16	1.07	1.00	1.08	1.14	1.07	1.15	-	1.06
Computer&TV&Radio&Print	*1.16*	***1.01***	*1.09*	***1.02***	1.08	1.16	1.09	1.01	0.94	1.02	1.07	1.00	1.08	0.94	-

Note: The column represents the media option that will be SUBSTITUTED FOR, The row represents the media option that will be SUBSTITUTED TO. The highlighted cells indicate strong substitution effects such that consumers will give up more quantities of the column media to obtain one additional unit of the row media.

The underlined italicized estimates indicate potential complementarity, i.e., where multiplexing is preferred to single media option.

A1.TABLE 4H: Marginal Rate of Substitution at the Point OF Zero Consumption: Workplace PC Use

	Computer	TV	Radio	Print	Computer&TV	Computer&Radio	Computer&Print	TV&Radio	TV&Print	Radio&Print	Computer&TV&Radio	Computer&TV&Print	Computer&Radio&Print	TV&Radio&Print	Computer&TV&Radio&Print
Computer	-	1.26	1.15	1.45	1.10	1.00	1.27	1.27	1.60	1.46	1.11	1.40	1.27	1.61	1.41
TV	0.79	-	0.91	1.15	0.88	0.80	1.01	1.00	1.27	1.15	0.88	1.11	1.01	1.27	1.12
Radio	0.87	1.10	-	1.26	0.96	0.88	1.11	1.10	1.40	1.27	0.97	1.22	1.11	1.40	1.23
Print	0.69	0.87	0.79	-	0.76	0.69	0.88	0.87	1.10	1.00	0.77	0.97	0.88	1.11	0.97
Computer&TV	0.91	*1.14*	*1.04*	*1.31*	-	0.91	1.15	1.15	1.45	1.32	1.00	1.27	1.15	1.46	1.27
Computer&Radio	1.00	*1.26*	*1.14*	*1.44*	1.10	-	1.26	1.26	1.59	1.45	1.10	1.40	1.27	1.60	1.40
Computer&Print	0.79	0.99	0.90	*1.14*	0.87	0.79	-	1.00	1.26	1.15	0.87	1.10	1.00	1.27	1.11
TV&Radio	0.79	1.00	0.91	*1.14*	0.87	0.79	1.00	-	1.26	1.15	0.88	1.11	1.01	1.27	1.11
TV&Print	0.62	0.79	0.72	0.91	0.69	0.63	0.79	0.79	-	0.91	0.69	0.88	0.80	1.00	0.88
Radio&Print	0.69	0.87	0.79	*1.00*	0.76	0.69	0.87	0.87	1.10	-	0.76	0.96	0.88	1.10	0.97
Computer&TV&Radio	0.90	*1.14*	*1.03*	*1.31*	1.00	0.91	1.14	1.14	1.44	1.31	-	1.26	1.15	1.45	1.27
Computer&TV&Print	0.71	0.90	0.82	*1.03*	0.79	0.72	0.91	0.90	1.14	1.04	0.79	-	0.91	1.15	1.00
Computer&Radio&Print	0.78	0.99	0.90	*1.14*	0.87	0.79	1.00	0.99	1.26	1.14	0.87	1.10	-	1.26	1.10
TV&Radio&Print	0.62	0.78	0.71	0.90	0.69	0.62	0.79	0.79	1.00	0.91	0.69	0.87	0.79	-	0.88
Computer&TV&Radio&Print	0.71	0.90	0.81	*1.03*	0.78	0.71	0.90	0.90	1.14	1.03	0.79	1.00	0.91	1.14	-

Note: The column represents the media option that will be SUBSTITUTED FOR, The row represents the media option that will be SUBSTITUTED TO. The highlighted cells indicate strong substitution effects such that consumers will give up more quantities of the column media to obtain one additional unit of the row media.

The underlined italicized estimates indicate potential complementarity, i.e., where multiplexing is preferred to single media option.

A1.TABLE 5: COUNTERFACTUAL RESULTS FROM PRINT MEDIA EXIT IN NEW YORK CITY

AFTER / BEFORE	*Total # OBS*	Outside good (%)	Computer (%)	TV (%)	Radio (%)	Computer & TV (%)	Computer & Radio (%)	TV & Radio (%)	Computer & TV & Radio (%)
Print	380	204 (53.7)	127 (33.4)	28 (7.4)	21 (5.5)	0	0	0	0
Radio & Print	65	0	35 (53.8)	0	0	12 (18.5)	18 (27.7)	0	0
TV & Print	212	0	0	23 (10.8)	0	121 (57.1)	0	68 (27.7)	0
TV & Radio & Print	115	0	0	0	13 (11.3)	0	73 (63.5)	29 (25.2)	0
Computer & Print	65	0	0	0	0	0	0	0	65 (100.0)
Computer & Radio & Print	35	0	0	0	0	0	0	0	35 (100.0)
Computer & TV & Print	23	1 (4.3)	0	0	0	0	0	0	22 (95.7)
Computer & TV & Radio & Print	104	1 (1.0)	0	0	0	0	0	0	103 (99.0)