Eric Schwartz 2010 Ackoff Doctoral Student Fellowship

ericschw@wharton.upenn.edu

February 26, 2010

2nd Year PhD Student, Marketing

Co-Advisors: Eric Bradlow and Peter Fader

Exploration and Exploitation in Interactive Media

Project Summary

A challenging tension many managers face is the tradeoff between exploring actions with uncertain outcomes and exploiting their knowledge of more certain options. When marketers are advertising on an online video site, they select which ad copy to serve to a viewer. But they often test out a few versions, and based on those results, they either choose the "best" version or keep testing. Firms practicing customer relationship management or direct marketing constantly experiment with different way s of targeting customers, for instance, via using A/B testing on the Internet (see the recent rise in popular press business books on the subject).

But how can experimentation in interactive media be profitable in the long-run? What tools should firms use to map observed customer behavior onto informed experiments?

My research aims to unite a set of related methodological areas related to dynamic optimization and experimentation to tackle this problem in the context of interactive media. Under the supervision of Professors Eric Bradlow and Peter Fader and the collaboration of corporate participants in the Wharton Interactive Media Initiative (WIMI), I have gained access to data from the online video provider Hulu.com and have been in touch with an interactive CRM company willing to run field experiments. This link to practitioners is central to connecting the methodological and substantive aspects of my research by performing field tests.

First, I discuss the potential methodological contribution of this work, and then I follow with a discussion of the substantive issues I will explore. I have already begun to draw upon the following methodologies in my current research:

(1) hidden Markov models and probability models to capture non-stationary customer behavior,
(2) dynamic programming to (approximately) solve multi-armed bandit problems and partially observable Markov decision processes in order to generate optimal resource allocation policies, and
(3) sequential adaptive experiments or optimal treatment regimes, so the firm's policy itself serves as an informed policy experiment.

While these literatures are interrelated, there is great room for advancement, tightening those links, and making the problems feasible for real-time experiments with companies. For instance, hidden Markov models (HMMs) are well-established in marketing (cf. Netzer et al. 2008). On the other hand, there are limited applications of partially observable Markov decision processes (POMDPs) in marketing (Montoya et al. 2009), although they are common in computer science. Resource allocation problems formulated as bandit problems have long been studied in economics, operations research, computer science, and marketing, but recent computational advances for good approximations have brought these problems into

a new light. Using these tools, the field of marketing has only started to scratch the surface of applying these results to live experiments (Simester et al. 2005; Hauser et al. 2009).

These methods capture and quantify the tradeoff between exploration and exploitation of available actions. Instead of simply choosing the action that has performed best on average based on past experience, it is often optimal to choose some actions that may have a slightly lower average in the past, but have a higher potential upside (i.e., higher risk and uncertainty). This is a fundamental result of bandit problems (Gittins 1979).

The area of optimal adaptive experiments, however, remains nearly untapped by marketers. While sequential sampling and efficient experimental design have been thoroughly studied in marketing, a new area in statistics, especially applied to clinical trials, has been emerging. Such experiments are designed not only to identify effects of manipulations, but also to optimize an desired outcome (Murphy et al. 2007; Murphy 2003). The results come in the form of optimal decision rules, or contingency plans, for instance, a treatment regime to reduce drinking days in a sample of alcoholics. The mathematical underpinning of these optimal sequential experiments is approximate dynamic programming, so the methodological link to the above areas is clear.

Interactive media provides an outstanding domain not only for developing and testing these methods, but also for contributing to new substantive findings in marketing and risk management. For instance, online video providers are optimizing by testing quantity of ads, types of ads, formats of ads, and price of ad-free service (HBR Case on Hulu 2009). When choosing among a set of possible ads to serve to a visitor, in certain cases, the best strategy is not necessarily to exploit the current knowledge and serve the one that has the highest expected reward; instead, it could be to serve the one with the highest expected mean plus some "option value," which depends on the uncertainty of the potential reward. So, a more uncertain option should be explored.

The combination of these methods with this domain opens up a variety of interesting questions. In the short run in an ad-supported online video environment, more ad impressions are clearly profitable, but what may be the long-run effect of ads on viewing? Do more online video ads make viewers less valuable in the long-run (i.e., increase their attrition rates, shorten session length, and/or decrease ads viewed over time)? On the other hand, do certain types of ads or new media ad formats have a positive on future value of viewers? Is there a "sweet spot" (i.e., optimal) quantity of ad exposures for certain types of ads or viewers? If so, how should a firm explore that optimal level of ad quantity?

Another key topic is customer "engagement" with websites, social media, and applications among other forms of media. Nevertheless, it is a loosely defined yet commonly used notion. Quantifying this metric with dynamic models of customer behavior is important for research and practice. With this metric, firms can answer questions such as, is it better to target users when they are most engaged, moderately engaged, or less engaged?

This problem of making sequential decisions in uncertain environments has a long tradition in operations research, computer science, economics, and marketing. My research interests sit at the intersection of various methods and the emergence of new media. I intend to bring these together with live field experimentation with firms to both develop a usable tool for firms and generate novel insights for theory and practice. At the center of the questions and areas discussed in this proposal is the key tradeoff between exploration and exploitation in interactive media.

Budget Information

Anticipated Expenses

Most of my expenses will be travel and registration for conferences. I have already established contacts with two firms willing to share their data and potentially run field experiments. In addition, I may travel to other universities (like MIT or Columbia) to meet with professors or attend courses. Since the work will involve optimal experimentation, it may be natural to test the adaptive experiments in the Wharton Behavioral Lab, so this could be an additional cost. Other expenses will be purchasing software and books.