

February 28, 2011

3<sup>rd</sup> Year PhD Student Marketing

Co-Advisors: Eric Bradlow and Peter Fader

**Project Title:** Beyond the Basic Bandit Problem in Interactive Marketing

**Project Summary:**

Interactive marketing problems implemented by firms to stimulate demand and subsequent sales are often stated as, “Which marketing actions should our firm send to which customers and when?” The current industry practice of A/B or multivariate testing aims to learn about the efficacy of such tactical actions, ranging from a particular email ad copy to a promotional offer or discount. My research aims to improve the current practice of marketing experimentation by running, instead, an optimal *sequence* of tests. Facing similar problems, researchers often turn to dynamic optimization methods. One such method, which has been used in the marketing literature, solves the basic *multiarmed bandit problem*. I review these problems, but I also argue for the benefits of using methods beyond the basic bandit framework. I draw on the computer science field of *Reinforcement Learning* (RL) as a more general framework that includes the *Markov Decision Process* (MDP) associated with basic bandit problems. RL formalizes how a firm learns about its customers and the effectiveness of its marketing actions by repeatedly interacting with the market. The algorithms from this approach efficiently strike a balance between *exploring* marketing actions the firm is uncertain about and *exploiting* the information that has been collected thus far.

One marketing problem that requires innovation beyond the basic bandit problem is the following: after a series of tests, how can firms learn about a tactic they have not yet tested? These methods can incorporate a conjoint-like design of a set of actions that firms considers selecting. Then the algorithms accommodate those correlations among actions to share information across actions, even to give a best guess at performance of actions not yet tried. Another marketing problem that is often overlooked in sequential tests is drilling down to the individual level. Instead of testing which website design works best on aggregate, for example, I am interested in developing algorithms and models to learn the best actions to take with particular groups of individuals based on their based behavior.

The RL approach links with many areas of the marketing literature, but these connections are not known to most researchers in either field. First, it shifts the descriptive and predictive work on the value of customers to a prescriptive and normative stream of work on how to manage customer relationships. Second, these methods have commonalities with problems involving adaptive clinical trials,

structural dynamic discrete choice models, and econometric design of field and quasi-experiments. Third, just as these methods have benefited clinical trials researchers, so too could they benefit consumer behavior researchers by reducing sample size requirements and providing effect new approaches for testing sequences of treatments.

This work is basis for my dissertation. Over the next year I plan to both perform extensive numerical validation of the methods and collaborate with a company to put these methods to the test in real markets. The Monte Carlo experiments will demonstrate under which conditions the general algorithms I propose out-perform existing methods in marketing. This will be done for a range of settings including simple bandit problems (e.g., no inter-relations among actions, small action space) so that exact dynamic programming solutions are available. But the numerical experiment will also include a range of settings where I manipulate the true degree of correlation among the actions available to a firm and the underlying heterogeneity of the customer base.

As I continue to find the best industry partner for collaboration, I continue about which industries and specific companies are better fits for this work. It's crucial that the firm can target individual customers by tracking individual-level behavior and responses to those actions over time. Those targeting decisions can be done in multiple decision periods, so that within a decision epoch different customers receive different marketing tactics. To secure a collaborator to run these sequences of field experiments, I have spoken with firms ranging from nonprofits, e-retailers, online gaming firms, fermium subscription businesses, and interactive media firms, among others. Many of those contacts have come thanks to the work of the Wharton Customer Analytics Initiative.

My research centers on the problem of making sequential decisions in uncertain environments. This has a long tradition in operations research, computer science, economics, and marketing. My research sits at the intersection these fields to bring a new marketing twist on a classic problem. I intend to bring these together with field experiments to both develop a usable tool for firms and generate both novel methodological and substantive insights.