

ERIC M. SCHWARTZ

Department of Marketing
The Wharton School of the University of Pennsylvania
726.7 Jon M. Huntsman Hall
3730 Walnut Street
Philadelphia, PA 19104-6340 USA

516-314-8387 (mobile)
215-898-2534 (fax)
eric Schw@wharton.upenn.edu
www.ericmichaelschwartz.com
Updated October 20, 2012

Education

Ph.D. University of Pennsylvania, Wharton School, Marketing, May 2013, expected

M.S. University of Pennsylvania, Wharton School, Statistics, May 2013, expected

B.A. University of Pennsylvania, College of Arts and Sciences, Mathematics and Spanish, May 2008

Research interests

Topics

Direct and interactive marketing
Test-and-learn marketing
Digital media advertising
Customer relationship management
Word-of-mouth

Methods

Adaptive marketing experiments
Dynamic programming
Reinforcement learning
Machine learning
Bayesian analysis

Publications

Berger, Jonah, and Eric M. Schwartz (2011), "What Drives Immediate and Ongoing Word of Mouth?" *Journal of Marketing Research*, 48 (5), 869-880.

– Featured in *Insights from MSI* (Fall 2010)

– Formerly *MSI Working Paper* [10-105].

Under review

Schwartz, Eric M., Eric T. Bradlow, Peter S. Fader (2012), "Model Selection Using Database Characteristics: Classification Methods and an Application to the 'HMM and Its Children,'" invited for 3rd round at *Marketing Science*.

Dissertation

Committee: Eric T. Bradlow (co-advisor), Peter S. Fader (co-advisor),
Raghuram Iyengar, and Christophe Van den Bulte

"Adaptive Marketing Experiments and the Attribute-based Multi-Armed Bandit" (Essay 1)

"Bandit's Paradise: Customer Acquisition through Online Display Advertising" (Essay 2)
– *Job Market Paper*

"Reducing Rapid Churn: How Does Treatment of New Customers Affect CLV?" (Essay 3)

Work in progress

“Conjoint for Profit: Adaptively Inducing Choice” with Raghuram Iyengar and Leonard Lodish

Other working papers and technical notes

Schwartz, Eric M., Eric T. Bradlow, Peter S. Fader, Yao Zhang, ““Children of the HMM’: Modeling Longitudinal Customer Behavior at Hulu.com,” Wharton Marketing Department Working Paper, August 2011.

Schwartz, Eric M., Peter S. Fader, Bruce G. S. Hardie, “Incorporating Covariates into Beta-Bernoulli Models,” Wharton Marketing Department Working Paper, January 2010.

Conference presentations and invited talks

“Adaptive Marketing Experiments and the Attribute-Based Multi-Armed Bandit”
Marketing in Israel Conference (December 2011)*
Rotterdam School of Management / Erasmus School of Economics (January 2012)*
Tilburg University (January 2012)*
Marketing Science Conference, Boston (June 2012)

“Attribute-Based Bandit Problem: A Reinforcement Learning Approach”
Marketing Science Conference, Houston (June 2011)

“Children of the HMM: Modeling Longitudinal Customer Behavior at Hulu.com”
Marketing Science Conference, Cologne (June 2010)

“What Do People Talk About and Why?”
Marketing Science Conference, Cologne (June 2010), presented by Jonah Berger

“Just Test It! Experimentation in Retail”
Jay H. Baker Retailing Initiative Board Meeting (November 2009)*

“Incorporating Covariates into Beta-Bernoulli Models”
Marketing Science Conference, Ann Arbor (June 2009)
John D. C. Little Festschrift (June 2009), presented by Peter Fader*

*Denotes invited talk.

Reviewing

Ad hoc reviewer for *Marketing Science*, *Marketing Letters*, and *Journal of Applied Econometrics*

Teaching and advising

Teaching assistant

Applied Probability Models in Marketing (MBA, undergraduate), 2010-2011
Introduction to Marketing (MBA for Executives, San Francisco), 2011
Marketing Metrics (Executive Education), 2009-2012
Sales Force Allocation (Executive Education), 2009-2012

Guest lecturer

Google Marketing Academy (Executive Education) 2012
Applied Probability Models in Marketing (MBA, undergraduate) (2008-2009)

Advisor

Master Thesis in Applied Mathematics and Computational Sciences, Misung Son, 2012
“Variance Formulas for the Beta-Geometric/Beta-Bernoulli Model” with Peter Fader

MBA Independent Study in Statistics, Jonathan Karmel, 2011
“From Linear Algebra to Hierarchical Generalized Linear Models”

Teacher training

Center for Teaching and Learning, Teacher Development Program (2011)

Service

Wharton Lunch and Learn Program (2010-2012)
Wharton Doctoral Council (2008-2012)
Wharton Doctoral Programs Orientation Speaker (2010)
Council of Undergraduate Deans, Student Representative (2007)
Student Committee on Undergraduate Education (2005-2008)

Honors, awards, and grants

AMA-Sheth Foundation Doctoral Consortium, Fellow (2011)
Workshop on Quantitative Marketing and Structural Econometrics, Fellow (Duke, 2010)
Russell Ackoff Award for Doctoral Student Research, Recipient (2009-2012)
Jay H. Baker Retailing Initiative Research Grant, Recipient (2009)
Lauder Center for International Business Education and Research Grant, Recipient (2009)
University of Pennsylvania Class of 1939 Fellowship, Recipient (2008-2009)
INFORMS Marketing Science Doctoral Consortium, Fellow (2009-2012)
Wharton Doctoral Fellowship, Recipient (2008-2012)
Summa Cum Laude, Dean’s List, University of Pennsylvania, GPA: 3.9/4.0 (2004-2008)
Benjamin Franklin Scholar, University of Pennsylvania (2004-2008)

Professional affiliations

INFORMS Society for Marketing Science
American Marketing Association
American Statistical Association
Phi Beta Kappa

Relevant coursework

Graduate-level Courses

Empirical Models in Marketing	David Bell, Maria Ana Vitorino
Economics and OR Models in Marketing	Jagmohan Raju
Measurement and Data Analysis	Raghuram Iyengar
Measurement and Data Analysis (2007)	Eric Bradlow, Wes Hutchinson, Robert Meyer
Applied Probability Models in Marketing	Peter Fader
Marketing Strategy	George Day, Christophe Van den Bulte
Research Methods in Marketing	Wes Hutchinson
Consumer Behavior: Information Processing	Geeta Menon
Dynamic Programming	Maria Rieders
Stochastic Models	Maria Rieders
Probability and Optimization	Michael Steele
Introduction to Optimization	Sergei Savin
Bayesian Statistics and Computation	Shane Jensen
Mathematical Statistics	Dylan Small
Numerical Methods in Economics*	Ulrich Doraszelski
Empirical Industrial Organization*	Ulrich Doraszelski
Observational Statistics	Dylan Small
Applied Econometrics I*	Dylan Small
Econometrics I: Fundamentals*	Kyungchul Song
Microeconomics	Richard Kihlstrom

Undergraduate-level Courses in Mathematics and Economics

Real Analysis
Linear and Abstract Algebra
Partial Differential Equations
Complex Analysis,
Microeconomic Theory Honors Seminar

* Denotes course not taken for credit.

Computer and natural languages

Fluent: Matlab, R, Spanish, SQL
Proficient: Catalan, SAS

References

Eric T. Bradlow
K. P. Chao Professor
Professor of Marketing, Statistics and Education
Vice-Dean and Director of Wharton Doctoral Programs
Co-Director of the Wharton Customer Analytics Initiative
Department of Marketing
761 Jon M. Huntsman Hall
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6340
ebradlow@wharton.upenn.edu
215.898.8255

Peter S. Fader
Frances and Pei-Yua Chia Professor
Professor of Marketing
Co-Director of the Wharton Customer Analytics Initiative
Department of Marketing
749 Jon M. Huntsman Hall
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6340
faderp@wharton.upenn.edu
215.898.1132

Raghuram Iyengar
Associate Professor of Marketing
Department of Marketing
756 Jon M. Huntsman Hall
The Wharton School
University of Pennsylvania
3730 Walnut Street
Philadelphia, PA 19104-6340
riyengar@wharton.upenn.edu
215.898-2391

Christophe Van den Bulte
Professor of Marketing
Department of Marketing
759 Jon M. Huntsman Hall
The Wharton School
University of Pennsylvania
3730 Walnut Street
Philadelphia, PA 19104-6340
vdbulte@wharton.upenn.edu
215.898.6532

Extended abstracts of dissertation research

Essay 1: Adaptive Marketing Experiments and the Attribute-based Multi-Armed Bandit

Changes in the interactive marketing environment have increasingly led firms to turn to business experiments such as A/B/C or multivariate tests. Many marketing activities – sending emails, serving online advertising, designing websites, recommending products – can be framed as adaptive experiments to continuously “*test and learn.*” Such contexts make up a rich class of direct/interactive marketing problems structured around the question: which targeted marketing action should we take, when, with which customers, and in which contexts? But as this practice becomes part of regular business operations, such continuous testing should be done profitably. The purpose of this research is to illustrate how to manage business experiments to simultaneously “*earn and learn.*” Profitable experimentation illustrates the dilemma of exploration versus exploitation, where a manager chooses between an action that is thought to be performing well currently (exploit) and an uncertain action with the hope of benefiting from that learning in the future (explore). This tension is at the heart of the classic sequential decision-making problem, known as the multi-armed bandit, which comprises the core theoretical thread of this research.

To resolve this tension, this paper makes two contributions to practice and theory. First, we greatly broaden the types of such problems marketers can address with the multi-armed bandit framework. More specifically, we introduce two key features that make the bandit problem much more applicable to marketing decisions: (a) we bring in an attribute structure, which allows for shared learning and greater efficiency across a wide array of tactics at the manager’s disposal (as in any attribute-based or regression-based modeling approach like conjoint analysis); and (b) we allow for a heterogeneous response model within the bandit framework. Second, not only do we quantify the overall improvements in algorithm performance, but we also qualify it by investigating a range of moderating factors related to the practical implementation of various test-and-learn algorithms. We develop a comprehensive taxonomy that shows how various aspects of the managerial problem (e.g., decision context, data sparseness) can affect the relative performance of different multi-armed bandit methods. For instance, three of these issues are: (a) How rare is the event of interest (e.g., clicks, transactions)? (b) How small are the effect sizes differentiating the actions? and (c) How big is the batch of simultaneous allocations that are made for each decision period (e.g., ad impressions, customers)?

While existing methods in the literature typically include numerical experiments, they tend to focus on illustrating the theoretical results (e.g., asymptotic properties) using “toy problems” with a limited range of alternative actions and limited contextual factors. These analyses usually lack the breadth and richness of a real-world setting. By contrast, our numerical experiments are intended to be highly realistic, reflecting challenging conditions, such as many actions, extremely rare events of interest, and large batches of simultaneous decisions. Plus, we manipulate a range of conditions in order to learn about their impact on test-and-learn method performance. As a result of this research, we provide evidence of when each method is likely to perform well and a recommendation of when to use which method. By drawing upon bandit methods across operations research (*dynamic programming*), statistics (*index strategies, randomized probability matching* from clinical trials), and computer science (*upper confidence bound policies* from reinforcement learning), and systematically manipulating a variety of aspects of the problem in numerical experiments, we complete the first such empirical investigation. We have already performed extensive simulations along these various dimensions of the problem including a wide array of algorithms (old and new) to manage the bandit problems.

*Essay 2: Bandit's Paradise: Customer Acquisition through Online Display Advertising
(Job Market Paper)*

Despite the common concerns about the (presumed) ineffectiveness of display ads, advertisers still use them extensively for the purpose of acquiring new customers. As advertisers try to maximize their return on investment in purchasing online media, their challenge is to determine which ads to serve and on which websites (i.e., contextual advertising) to deliver them. But how should firms allocate their ad impressions so that they can simultaneously learn how to grow profits in the future and improve profits while learning? The setting of online display advertising presents five interesting challenges, and no existing method addresses all of them. (1) Advertisers often try dozens of different ads. (2) The various ads are interrelated, since they can be described by attributes such as creative design, message and size/format. Thus, observing one ad's performance can suggest how similar ads will perform. (3) The way those attributes affect the ad's performance depends on the context, such as the website on which it appears. (4) While much previous work has focused on click-through rates as the outcome measure of interest, the bottom-line return on investment is not necessarily linked to clicks, but rather to conversions (i.e., new customers actually acquired). But that is quite a rare event, since 100,000 ad impressions may lead to only one new customer. (5) Finally, as is common in media buying, the advertiser's key decision is what percentage of the next batch of already purchased impressions should be allocated to each ad (i.e., the weights that the publisher uses to rotate the ads).

To overcome all of these challenges, we apply a Bayesian approach to manage the *attribute-based multi-armed bandit*, accounting for *unobserved heterogeneity* and *batched* decision making. This approach is based on a principle called *randomized probability matching*. We demonstrate the business value of running the adaptive near-optimal experiment through live implementation in a controlled field experiment with a large international retail bank that focuses on direct marketing. In addition, we can perform counterfactual policy simulation, since we have generated random variation in advertising. We find that this approach outperforms benchmark methods from the literature and heuristic stopping rules for experiments used in practice. Finally, we quantify the expected loss – new customers not acquired – that advertisers experience by optimizing click-through instead of conversions directly.

We have already performed two field experiments focused on acquiring customers for a single financial product using ads covering nearly 100 creative-by-size combinations and more than 200 million ad impressions delivered across a mixture of 15 websites and ad networks. We tested our algorithm with a live experiment in collaboration with the advertising bank by setting the weights with which the ad delivery system randomly rotated each ad for each media placement. The setting of these rotation weights is exactly what advertisers can commonly control, and our algorithms adjust these weights (away from uniform weights) in a fully data-driven model-based manner. As a result, we present the largest scale live field experiment of these test-and-learn methods to date (in the online display ad context and in others). Results suggest that there are substantial gains in not only considering are different creative attributes explain ad performance (*attributed-based approach*), but also that the importance of those creative attributes differs by the type of website where the ad appears (*unobserved heterogeneity*). We also offer insight into how using past performance to categorize media (e.g., websites, placements, and publishers), instead of using approximations based on the demographics of the sites' visitors, can lead to more informed ad allocations and future media buying.

Essay 3: Reducing Rapid Churn: How Does Treatment of New Customers Affect CLV?

An increasing number of business models rest on converting a high volume of trials, traffic, or installations into repeat and paying customers, while recognizing that a vast majority of them will never make another transaction, visit, or play again. Within this setting, the firm's key to monetizing its user/customer base (e.g., preventing churn and extracting value) hinges on what the firm does at the moment its customers are acquired. Despite much work aiming to optimize customer relationship management (CRM) actions to maximize expected customer lifetime value (CLV), surprisingly little attention is paid to such rapid churn environments and even less is paid to what to do about it. However, given the decreasing costs of testing different CRM tactics, the rapid churn environment is a natural setting to continuously test-and-learn which methods of treating new customers turn out to generate more profitable and longer lasting customers. In this research, we do so using a multi-armed bandit approach to CRM that maximizes long-run customer value. While models of repeat purchasing can provide expectations of what customers will do in a future period, those same models do not immediately tell the marketer what to do. Facing such predictions of customer behavior, managers ask, "Should we target our customers expected to be most valuable or the ones that seem less valuable?" Managers may instead prefer a third choice: to target the customers for whom those actions will have the greatest return on investment. To answer this question, firms could run experiments using different marketing activities such as A/B or multivariate tests (with stratified randomized ensuring balance within each level of expected customer value). But business settings with rapid customer churn do not give firms the luxury of learning how effective its targeted marketing activities will be by interacting repeatedly with the same customer (since that customer is likely gone soon after the first contact). However, there is a plus side to this problem: the short life-span of customers shrinks the time-lag between action and learning. That is, after a brief period of time, we can begin to evaluate the potential success of the intervention.

The purpose of this work is to address this key managerial question: how should marketers profitably learn about how to manage their customer relationships and learn about their impact on future customer value? From a substantive perspective, we contribute to the literature on bandit problems by introducing the realistic business consideration of the future value of actions (in terms of customer value), beyond the value of learning about those actions' effectiveness. Various challenges arise when using future customer value as a payoff function, so we make methodological contributions, as well. Since future value is not immediately observed, there is still a lag between taking the action and fully learning its effect on the customer's future value. Hence, there are two ways to learn about a marketing action's effect on a customer's expected future value: directly and indirectly. That is, to learn about the action's impact over the next six-months, you can randomly assign customers to the action then wait six months for the realized value (directly), or you can see the intermediate transaction patterns and update expectations at each step along the way (indirectly, model-based). While the difference between direct and indirect learning is intuitive, we formalize it using reinforcement learning, and quantify the benefits from adopting this framework to solve this marketing problem.