Ryan Dew

Columbia Business School – 3022 Broadway, Uris Hall 5C – New York, NY 10027
+1 (610) 883-0929 • ryan.dew@columbia.edu • www.rtdew.com

Education

Columbia University **New York, NY**
Columbia Business School
Ph.D., Marketing, Expected 2018
M.Phil., Marketing, 2016

University of Pennsylvania **Philadelphia, PA**
College of Arts and Sciences
B.A., Mathematics, 2013
Academic honors: Summa cum laude, Phi Beta Kappa

Research Interests

Substantive: customer relationship management, customer analytics, data-driven design, decision support, preference measurement, creativity

Methodological: machine learning, Bayesian nonparametrics, unstructured data (e.g. text, images), big data, scalable inference, Bayesian econometrics

Dissertation

Title: Machine Learning Methods for Data-Driven Decisions

Committee: Asim Ansari (advisor), Olivier Toubia, Oded Netzer, David Blei

Publications and Manuscripts Under Review

Dew, Ryan and Asim Ansari (Forthcoming), "Bayesian Nonparametric Customer Base Analysis with Model-based Visualizations," *Marketing Science*.

Dew, Ryan, Yang Li, and Asim Ansari, "Dynamic Preference Heterogeneity," revision invited at *Journal of Marketing Research*.

Research in Progress

Dew, Ryan, Asim Ansari, and Olivier Toubia, "Letting Logos Speak: Deep, Probabilistic Models for Logo Design."

Dew, Ryan and Oded Netzer, "Customer-Centric Data Fusion."

Dew, Ryan and Asim Ansari, "Scalable Decision Support Systems for Robust CRM."

Conference Presentations

Marketing Science, Los Angeles, CA, June 2017
"Dynamic Heterogeneity: A Bayesian Nonparametric Approach"

Marketing Dynamics, Hamburg, Germany, July 2016
"Gaussian Process Dynamic Choice Models"

AMA Advanced Research Techniques Forum, Boston, MA, June 2016
"Bayesian Semiparametric Framework for Understanding and Predicting Customer Base Dynamics"

Marketing Science, Shanghai, China, June 2016
"Gaussian Process Dynamic Choice Models"

Data Science Day (Poster Session), Columbia University, April 2016
"Model-based Dashboards for Customer Analytics"

Marketing Science, Baltimore, MD, June 2015
"Bayesian Semiparametric Modeling of Cohort Lifecycles"

Grants, Honors, and Awards

AMA-Sheth Foundation Doctoral Consortium Fellow, 2017

ISMS Doctoral Consortium Fellow, 2017

Amanda and Harold J. Rudolph Fellowship, Columbia Business School, 2016

Deming Center Doctoral Fellowship, Columbia Business School, 2016

ISMS Doctoral Consortium Fellow, 2016

Quantitative Marketing and Structural Econometrics Workshop, 2015

ISMS Doctoral Consortium Fellow, 2015

Adobe Digital Marketing Research Award (with Kinshuk Jerath and Miklos Sarvary), 2014

Doctoral Program Fellowship, Columbia Business School, 2013-2018

Phi Beta Kappa, University of Pennsylvania, 2013

Teaching Interests

Marketing Analytics
Big Data and Computational Marketing
Marketing Research
Applied Probability Models in Marketing
Machine Learning and Bayesian Methods in Marketing

Teaching Experience

Teaching Assistant

MBA:

Marketing, *MBA Core*, Spring 2014, Fall 2014, Spring 2015, Fall 2015, Spring 2016
Marketing Strategy, *EMBA Core*, Fall 2014, Summer 2015, Fall 2015, Spring 2016, Fall 2016
Digital Marketing, *MBA Elective*, Fall 2015-2016
Pricing, *MBA Elective*, Spring 2015, Spring 2016
Marketing for Organic Revenue Growth, *EMBA Elective*, Winter 2015, Winter 2016
The Psychology and Economics of Consumer Finance, *MBA Elective*, Winter 2014

Doctoral:

Causal Inference, *Ph.D. Seminar*, Fall 2015, Fall 2016
Empirical Models in Marketing, *Ph.D. Seminar*, Spring 2015

Tutorials

Estimating Bayesian Models with Stan, for *Bayesian Methods in Marketing*, Fall 2015

Introduction to Programming in R, for *Empirical Models in Marketing*, Spring 2015

Conjoint Analysis, for *Marketing Strategy*, Fall 2014, Fall 2015, Spring 2016, Fall 2016

Work Experience

Electronic Arts — **Redwood City, CA**
Advanced Analytics Intern — *2013*

Wharton Customer Analytics Initiative — **Philadelphia, PA**
Research Assistant — *2012-2013*

Self-run Tutoring Service — **Philadelphia, PA and New York, NY**
Private Tutor — *2010-2014*
Tutored undergraduate mathematics, statistics, economics, and English writing.

Doctoral Coursework

Marketing:

Empirical Models in Marketing	Asim Ansari
Mathematical Models in Marketing	Rajeev Kohli
Analytical Models in Marketing	Kinshuk Jerath
Bayesian Methods in Marketing	Asim Ansari
Advanced Empirical Methods	Asim Ansari, Olivier Toubia, Oded Netzer, Scott Shriver
Consumer Behavior I	Eric Johnson
Consumer Behavior II	Michel Pham and Bernd Schmitt
Marketing Decisions and Methods	Donald Lehmann

Economics:

Econometrics I	Jushan Bai
Econometrics II	Christoph Rothe
Economic Theory I-II	Geoffrey Heal
Economic Theory III-IV	Paolo Siconolfi
Industrial Organization	Andrea Prat
Empirical Methods in MS/OM	Marcelo Olivares
Economics and Optimization of Online Marketplaces	Gabriel Weintraub

Statistics and Machine Learning:

Foundations of Graphical Models	David Blei
Truth in Data	David Blei
Gaussian Processes and Kernel Methods	John Cunningham
Probabilistic Models of Discrete Data	David Blei
Applied Multivariate Statistics	Kamel Jedidi
Causal Inference	Jose Zubizarreta

Statistical Methodology	Andreas Buja (Penn)
Bayesian Methods and Computation	Shane Jensen (Penn)

Languages

Computer: R, Python, Julia, Stan, Mathematica, SQL (basic)

Human: English (native), Spanish (intermediate), Mandarin (beginner)

References

Asim Ansari (*Advisor*)
William T. Dillard Professor of Marketing
Columbia University
`maa48@gsb.columbia.edu`
(212) 854-7647

David Blei (*Committee member*)
Professor of Computer Science and Statistics
Columbia University
`david.blei@columbia.edu`

Oded Netzer (*Committee member*)
Associate Professor of Business
Columbia University
`onetzer@gsb.columbia.edu`
(212) 854-9024

Olivier Toubia (*Committee member*)
Glaubinger Professor of Business
Columbia University
`ot2107@gsb.columbia.edu`
(212) 854-8243

Selected Abstracts

Bayesian Nonparametric Customer Base Analysis with Model-based Visualizations

Ryan Dew and Asim Ansari

Based on dissertation essay 1, forthcoming at Marketing Science.

Modern marketers are responsible for understanding and managing customer spending behavior across many different products. Dynamics in spending result from both predictable customer-level effects, which are characterized by interpurchase time, customer lifetime, and past purchase frequency, as well as calendar time effects, which are driven by managerial actions such as product changes and promotions, and by general trends and random shocks. Understanding these dynamics in spending is further complicated by a lack of knowledge of all of the factors that influence spending for a given product, a problem exacerbated in large multiproduct firms by information asymmetries that can exist between the product teams that execute marketing actions and the marketing analytics team responsible for customer base analysis. A comprehensive understanding of customer base dynamics therefore requires a modeling framework that flexibly integrates both known and unknown calendar time determinants of spending with the individual-level effects that robustly predict spend activity. In this paper, we develop a Bayesian nonparametric framework based on Gaussian process priors to understand and predict customer spending. Our model separates out calendar time effects from individual-level dynamics by modeling both sets of factors as unknown latent functions that jointly determine spend propensity. The primary output of our Gaussian Process Propensity Model (GPPM) is a set of estimated curves that provides a visual and easily comprehensible representation of purchasing dynamics, which we call the model-based dashboard. We illustrate the utility of our modeling framework on data from two popular free-to-play mobile video games. We show how the GPPM's model-based dashboard can be useful for assessing patterns and disruptions in spending. We also show how the GPPM exhibits superior forecasting ability compared to existing customer base analysis benchmarks, including hazard and buy-till-you-die models.

Most recent version available online at `http://ssrn.com/abstract=2692307`

Dynamic Preference Heterogeneity

Ryan Dew, Yang Li, and Asim Ansari

Based on dissertation essay 2, invited for revision at Journal of Marketing Research.

Consumers' preferences change over time, often in tandem with population trends, but frequently exhibiting individual-specific idiosyncrasies. In this paper, we develop a novel Bayesian nonparametric framework based on Hierarchical Gaussian Processes (HGP) for modeling such dynamic heterogeneity. Our specification generalizes previous approaches for estimating dynamics that have been used in the marketing literature by flexibly capturing both the evolution of population trends and individual-level departures from those trends. This allows for sharing of statistical information across individuals, and within individuals over time, and provides rich individual-level insights and efficient inferences regarding preference evolution. We showcase our HGP specification in a choice modeling context, using simulations and real data from two CPG categories. We find that commonly used heterogeneity specifications can lead to significant biases or inefficiencies when dynamic heterogeneity is present, and are unable to represent managerially relevant individual-level parameter trajectories. Our application uncovers robust evidence of dynamic heterogeneity during the Great Recession, and important individual-specific trends that can be leveraged by retailers for optimal targeted pricing.

Most recent version available online at `http://ssrn.com/abstract=2915632`

Letting Logos Speak: Deep, Probabilistic Models for Logo Design

Ryan Dew, Asim Ansari, and Olivier Toubia

Dissertation essay 3.

Logos serve a fundamental role in branding as the visual figurehead of the brand. Yet, due to the difficulty of using unstructured image data, prior research on logo design has been largely limited to non-quantitative studies. In this work, we explore logo design from a data-driven perspective. In particular, we aim to answer several key questions: first, to what degree can logos represent a brand's personality? Second, what are the key visual elements in logos that elicit brand and firm relevant associations, such as brand personality traits? Finally, given text describing a firm's brand or function, can we suggest features of a logo that elicit the firm's desired brand identity? To answer these questions, we develop both a novel logo tokenization algorithm, that uses modern image processing tools to decompose unstructured pixel-level image data into meaningful visual features, and a novel probabilistic model, which we term Guided Deep Gamma Trees (GDGT), which links those visual tokens with textual descriptions of firms. Our model extends existing models for text and recommendation systems, such as Latent Dirichlet Allocation and Poisson Factorization, by learning associations between text and visual features at differing levels of abstraction, via sets of latent factors that explain feature co-occurrences. Moreover, our model adds an additional layer of supervision that guides the learned latent factors to also explain consumers' perceptions of brand personality. The outcome of GDGT is a synthesis of both textual and image data, that allows interpretable textual topics to be used to answer questions about visual meaning, and that allows us to infer prototypical logos corresponding to given textual descriptions. We apply our modeling framework on a dataset of hundreds of logos, textual descriptions from firms' websites, third party descriptions of firms, and consumer evaluations of brand personality to explore the above questions.

Customer-Centric Data Fusion

Ryan Dew and Oded Netzer

Marketers face a deluge of data on all aspects of their customers, including what they are buying, how they are using their products, and what they are saying about their purchases. We seek to address the problem of understanding general patterns across many modes of customer behavior, by modeling consumers through a set of latent, inferred personas that are related to behaviors across all domains of interest. Our approach is based on mixed membership models and hierarchical Dirichlet processes, which are commonly used in machine learning for natural language processing. By adapting these methods as a scaffolding for a CRM system, we can allow firms to leverage data across domains, through existing probabilistic models of customer behavior, even in the presence of missing data, to better understand and predict behavior in any particular domain. We illustrate the expressive capacity and managerial utility of such a framework through an application to rental car transaction and textual feedback data.

This work was awarded the Deming Center's 2016 Doctoral Fellowship.

Scalable Decision Support Systems for Robust CRM

Ryan Dew and Asim Ansari

Based on dissertation essay 1.

Modern machine learning methods provide a powerful basis on which to build decision support systems for customer relationship management, yet many of the most flexible and comprehensive of these methods are computationally complex or slow to estimate on very large datasets. In particular, hierarchical models that capture the individual-level heterogeneity that is crucial for CRM and targeted marketing can suffer from scalability problems. In this work, we develop and apply methods for scalable, approximate inference using stochastic gradient methods, that allows robust and hierarchical decision support systems, such as the recently developed Gaussian process propensity model (GPPM), to be scaled to very large datasets. In particular, we show that use of stochastic gradient inference methods allows accurate and fast estimation of complex CRM models like the GPPM, which in turn facilitates two previously difficult forms of analysis: first, models can be fit to more subsets of the customer base, allowing for nuanced understanding of differences across acquisition channel and customer characteristics. Second, the model can be automatically adjusted to new settings, where more expressive structure is needed to understand customer spending behavior.