Marketing Letters (2020) 31:361–370
https://doi.org/10.1007/s11002-020-09535-7

How can machine learning aid behavioral marketing research?

Linda Hagen[1] · Kosuke Uetake[2] · Nathan Yang[3] · Bryan Bollinger[4] · Allison J. B. Chaney[5] · Daria Dzyabura[6] · Jordan Etkin[5] · Avi Goldfarb[7] · Liu Liu[8] · K. Sudhir[9] · Yanwen Wang[10] · James R. Wright[11] · Ying Zhu[12]

Published online: 16 July 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract
Behavioral science and machine learning have rapidly progressed in recent years. As there is growing interest among behavioral scholars to leverage machine learning, we present strategies for how these methods that can be of value to behavioral scientists using examples centered on behavioral research.

Keywords Behavioral science · Big data · Semi-supervised learning · Supervised learning · Unsupervised learning

1 Introduction

Machine learning (ML)—a field of computer science dedicated to developing learning algorithms, often using big data, to generate predictions needed to make decisions (Agarwal et al. 2018)—has made its way into marketing research. At the same time, behavioral science—broadly defined as a branch of science that studies human behavior—has gained prominence across many sub-disciplines of applied business and policy research. Behavioral science is now using big data for theory testing across many domains, such as business, health and wellness, and education. For example, a growing number of studies leverage big data obtained from field experiments (Berger and Milkman 2012) and pooled experiments (Hartford et al. 2016). Given this newly available data, behavioral scientists can benefit from machine learning to tackle

✉ Linda Hagen
Linda.Hagen@marshall.usc.edu

✉ Kosuke Uetake
kosuke.uetake@yale.edu

✉ Nathan Yang
ncy6@cornell.edu

Extended author information available on the last page of the article

important methodological challenges, such as heterogeneous treatment effects, stimulus sampling, and missing outcomes. With the right analytical tools, various types of big data can complement behavioral research: visual information including images posted by firms and customers (Hagen 2020; Liu et al. 2019), fMRI brain scans (Pereira et al. 2009), user spatial patterns (Blanchard et al. 2020; Hui et al. 2009), smart mirror feeds (Lu et al. 2016), and satellite imagery of consumers' living environment (Bollinger et al. 2019); textual data from company communications (Anand et al. 2020), customer reviews (Chakraborty et al. 2019), and online user-generated content (Puranam et al. 2017; Netzer et al. 2012); as well as high-frequency data about education (Lu et al. 2017), entertainment (Nevskaya and Albuquerque 2019), and exercise (Aral and Nicolaides 2017) behavior. The availability of these data provides new opportunities to study behavioral research topics with ML innovations.

Motivated by these opportunities, this article offers a primer on machine learning for behavioral scientists. We first provide a high-level overview of the type of data typically used in different ML methods, followed by more detailed descriptions of well-established ML methodologies (e.g., supervised learning, unsupervised learning, semi-supervised learning) that can be of value to behavioral researchers. After establishing this technical background, we offer detailed examples of how various ML methodologies can be used for specific behavioral research topics.

2 Overview of ML methods

We first provide a brief overview of ML methodologies most likely to be useful to behavioral researchers, focusing on the general intuition and objective behind each as well as illustrating behavioral research they may complement. For technical details, we recommend Dzyabura and Yoganarasimhan's (2018) book chapter.

Common across most ML procedures is the process of training a statistical learning model to make predictions for making data-driven decisions. Typical ML procedures require the researcher to allocate some data for algorithm training, while the remaining data is held-out for testing. The general objective for a canonical ML problem is to fit a model that makes a prediction using the predictors from training data and see how well the predictions match with reality when the fitted model is applied to the testing data. With a set of predictors from the data and specific assumptions about the model for fitting the data, a ML algorithm will generate predictions.

While the number of ML methods is vast, they can be grouped into three broad categories based on the type of training data typically required: supervised learning, unsupervised learning, and semi-supervised learning. A separate class is reinforcement learning (i.e., computational learning from interaction), which we only discuss briefly in the conclusion.

2.1 Broad comparison of ML methods by way of example

Before discussing each method type, it will be helpful to think about the key differences across these methods while abstracting away from the fine-grained details about the models. To frame our comparisons (Table 1), we will refer to some hypothetical behavioral research that aims to study the relationship between

Table 1 Summary of ML methods in example

Method	Data requirements	Example: extracting personality
Supervised	{Outcomes, Predictors}	{Personality traits, Social media posts}
Unsupervised	{Predictors}	{Social media posts}
Semi-supervised	Some obs. ➔ {Outcomes, Predictors} Most obs. ➔ {Predictors}	Some obs. ➔ {Personality traits, Social media posts} Most obs. ➔ {Social media posts}

customer personality and social media activity. The data available to the researcher might include participants' responses to online personality quizzes embedded on the social media platform (or perhaps personality inventories collected in the lab) and textual data scraped from their social media posts. The variables of interest are the personality traits, so we will refer to this variable as the outcome. For a more detailed survey of the analysis of text data in marketing, we refer the reader to Berger et al. (2019).

Supervised learning involves some variables labeled as predictors and other variables as outcomes. In this context, it could be used to identify what aspects of the textual content of users' social media posts (i.e., predictors) predict certain personality traits (i.e., outcomes).

Unsupervised learning, in contrast, only involves predictors. Instead of predicting a specified outcome, predictors are used to group entities of interest by similarity. In this context, a method like K-means clustering could be used to segment individuals based on the similarity of the content of their social media posts.

Finally, semi-supervised learning could be used to first fit a model to predict personality traits (i.e., outcomes) from social media posts (i.e., predictors) using labeled data from users from whom both are available, and then apply a predictive model that accommodates for observations with and without personality trait scores (e.g., users who did not complete the quiz).

In this example, some of the benefits and caveats of each methodology become apparent. Supervised learning, for instance, requires that the outcome of interest be very clearly defined, which may be difficult without a priori knowledge of the underlying behavioral theory (in this example, which trait is a relevant outcome, and what constitutes a valid measure of the trait). Conversely, unsupervised learning is very open-ended and does not necessarily yield readily usable predictions, as the segments obtained via clustering still require human interpretation to make meaning of what the grouping may represent (here, understanding the dimensions that define each cluster of social media personality type). Semi-supervised learning may offer an appealing solution to researchers who have a small-scale labeled dataset from the lab, but want to integrate into their analysis a large sample of unlabeled data (in this case, augmenting a laboratory sample with scraped from social media).

2.2 Supervised learning

This type of ML requires training data to be labeled, which means that the model is fitted using an outcome of interest (i.e., each observation's label). Two commonly used

supervised learning algorithms include least absolute shrinkage and selection operator (LASSO) and random forests.

LASSO is in many ways similar to standard linear regression and classification, though it contains additional constraints. These constraints help the algorithm perform variable selection, such that an important output from this algorithm in addition to the prediction itself are indications of which predictors to keep and which to drop through regularization that helps prevent overfitting (Zhu et al. 2019). In other words, unlike linear regression, where the researcher must predetermine a subset of variables to test as predictors, the algorithm considers all variables in a large set and identifies the subset of suitable predictors. The data-driven approach to dropping predictors is especially helpful when the number of predictors approaches the number of observations, which can lead to inaccurate predictions. Model complexity issues are likely to arise in behavioral studies for which the researcher collects many variables for each participant (e.g., demographics, psychographics, information collected from Internet-of-Things devices), but the number of participants is small. The use of LASSO may help a behavioral researcher by facilitating variable selection, which can help prioritize data collection efforts in the future if each variable is costly to collect (be it in terms of money, time, or logistical challenges). To implement the LASSO, one can use the R-package glmnet.

Another popular method is random forest. It is a technique for classification/regression that makes use of many decision trees, where each tree represents a "human" decision maker. Each decision tree alone may not be accurate, but when pooled together, prediction accuracy can improve. Furthermore, this algorithm can accommodate for any number of interactions between predictors when training the model, which is often of interest in behavioral research. In particular, the researcher does not have to assume a priori which variable interactions to consider. In addition to the prediction, post-estimation output of this algorithm also includes an variable importance ranking, which ranks the variables based on predictive value, that is, by how much model fit would deteriorate if a variable were to be omitted. For this reason, this algorithm can also be used for reducing model complexity as in the LASSO but requires more human discretion than the LASSO in terms of choosing a ranking threshold below which to drop predictors. To implement the random forest, one can use the R-package randomForest.

2.3 Unsupervised learning

This branch of ML makes use of training data that does not contain labels, which means that the model is not fitted using outcomes, but instead fitted using only predictors with the objective to group entities by their similarity. A familiar problem that these methods can solve is dimension reduction (e.g., grouping high-dimensional stimuli), akin to factor analysis or multidimensional scaling methods. Some unsupervised learning algorithms include clustering and autoencoders.

Clustering is a technique that aims to find subgroups or clusters in the data using the predictors alone. Such techniques are already commonly used in behavioral research. K-means clustering is an iterative approach for partitioning the data into K distinct and non-overlapping clusters. The researcher sets the number of clusters desired, and the algorithm identifies clusters for which within-cluster variation is as small as possible.

Iterations continue until each observation to the cluster can be assigned to a cluster centroid. Tasks that can be accomplished using K-means clustering include grouping similar images, text excerpts, or historical trajectories of observed behavior together. K-means clustering can be implemented using the R-command kmeans.

Autoencoders can be thought of as another dimension reduction technique. It is a neural network that aims to reconstruct the algorithm's input data, like images and text. The process of reconstructing the input involves a feature space, which has a lower dimensionality than the input space (e.g., a picture with thousands of pixels translated into several key features). Consequently, the feature space is a simpler, more useful, descriptive summary of the original high-dimensional data. Autoencoders can be implemented using the R-package autoencoder.

2.4 Semi-supervised learning

Semi-supervised learning brings together labeled and unlabeled data, and uses them to fit a classifier simultaneously (Chapelle et al. 2006). This approach is ideal if the researcher has partially censored data, where the outcome is only observed for a fraction of the observations. Such data structure might exist if the researcher collects data both passively (e.g., wearable technology, web searching behavior) and actively (e.g., surveys, lab responses), whereby observations from passively collected (big) data may not necessarily contain the exact outcome of interest unlike in actively collected data. These methods can be used when predictive factors need to be inferred from the data. In particular, a likelihood function used to fit the predictive model can consist of a composite of a linear regression (for when the outcome is observed) and a type-1 Tobit model (for when the outcome is not observed).

3 Integration of behavioral science with ML

This section highlights a few detailed examples of how ML methodologies may support behavioral research. While these examples do not cover all possible research streams that can benefit from ML methods, they showcase particular behavioral science challenges that ML can help overcome.

3.1 Supervised learning for estimating heterogeneous treatment effects

Many behavioral science questions involve the impact of judgments relative to the self or some idiosyncratic reference point (e.g., social comparison against others; gains and losses relative to some reference point; deviation from some norm). Such comparisons may influence consumer behavior and real-world outcomes, like educational attainment (Brewer and Weber 1994) and health/wellness (Lockwood et al. 2005). In controlled laboratory experiments, heterogeneity in subjects' self-view or current state can sometimes be mitigated (e.g., inducing subjects to think of a particular target that makes them feel inferior; endowing all participants with the same amount of money before they gain or lose anything), but in real-world settings, subjects enter in different states that can produce heterogeneous treatment effects. For instance, the identical social comparison target may constitute upward social comparison for one research subject, but downward comparison

for another, and thus produce vastly different treatments for these two individuals (Pelham and Wachsmuth 1995). Furthermore, variation in the group composition itself might further amplify this heterogeneity (Uetake and Yang 2019).

A challenge that such heterogeneous treatment effects introduce is that group means used to compare two supposedly categorically different treatments may be diluted by the potentially large variation within treatment groups. Traditional solutions in behavioral science that address heterogeneous treatment effects include non-parametric approaches like propensity-score matching. But such methods often require a priori knowledge of which combinations of moderators need to be included.

To accommodate for potential heterogeneous treatment effects, behavioral science can use causal forest estimation techniques in conjunction with randomized control trials (Wager and Athey 2018). These methods essentially use the supervised learning technique to establish propensity trees (rather than propensity scores) for matching treatment and control groups. Unlike the traditional propensity score matching, causal forests allow the data to determine the dimensions most important to consider when selecting the nearest neighbors. Causal forests can be implemented using the R-package grf, and this technique has been used in recent quantitative marketing research to study heterogeneous treatment effects (Ascarza 2018; Guo et al. 2019; Kim et al. 2019).

3.2 Unsupervised learning for stimulus sampling

Behavioral researchers frequently study the effect of some composite judgment (e.g., product esthetics; a brand's "warmth") on related behavior or other judgments. A typical approach in behavioral science is to identify one particular manifestation that epitomizes the construct of interest (e.g., one specific target rated as physically attractive in a pretest) to use as an experimental stimulus. However, this practice may lead to rather arbitrary stimulus selection, as any number of stimuli may embody the construct of interest.

Unsupervised learning techniques can be used to group real-world stimuli, which can aid researchers in determining the most relevant features that jointly give rise to the composite judgment. For example, to study the effect of esthetics on brand attitudes, researchers may recruit test subjects to respond to a large set of real visual ads, some of which may later serve as experimental stimuli. The visual ads can be considered high-dimensional data, and unsupervised learning will help reduce their dimensionality, for example, by clustering them into groups of particularly similar types of visual ads (Sanakoyeu et al. 2018). This dimension-reduction may then help researchers identify salient visual stimuli that can be used interchangeably in later experiments; it may inform their thinking about what key esthetic ingredients differentiate the clusters and may direct stimulus design; and it may even guide their theoretical thinking about the nature of the esthetics construct in question.

An attractive feature of deep unsupervised learning is that researchers can leverage pre-trained neural networks to take full advantage of machine learning architecture that has been refined over the years using large amounts of data from publicly available repositories. This computational advantage is less prevalent in classification exercises that use mean-field approximations (Tanaka and Titterington 2004). Using an autoencoder algorithm, researchers can organize a large set of stimuli in an agnostic manner prior to the experiment. Different types of visual stimuli can then be assigned across individuals to evaluate which types elicit positive responses.

3.3 Semi-supervised learning for evaluating dynamic behavioral phenomena

Many consumer behaviors evolve over time (e.g., information search, expertise, enjoyment, motivation). Lab experiments can assign some individuals into a treatment group and observe the dynamics of their behavior or outcomes over time, but a nchallenge of longitudinal studies is the exacerbated burden on the participant and issues with (potentially differential) attrition.

A large body of work on dynamic behavioral phenomena include goals and motivation. Experiments in this area of research have offered important insights about how to spur goal-directed action (Fishbach and Touré-Tillery 2013), but there is an increasing desire to study the same phenomenon in the field (Gordon et al. 2019; Kettle et al. 2016; Suher et al. 2019; Wang et al. 2016), especially given that a lot of consumer technology embedded into people's everyday lives allows for easy personal quantification of goal-directed behavior (Etkin 2016). Studying goals and motivation via field experiments is desirable because as researchers can observe individuals in a real-world context over a long period of time—across many months rather than days.

However, longitudinal, involved designs often increase compliance failure and imperfect response rates, and participants may not record all information relevant for tracking goal progress (Knäuper et al. 2018). Missing information about outcomes may become even more pronounced as the study's temporal window expands. For instance, in a study on health goal pursuit, many of the health outcome labels (e.g., daily weight for fitness-related goals) may be missing for many time periods. Semi-supervised learning may help researchers predict these missing health outcomes using data based on passively collected goal-directed behavior (e.g., step-counter data from wearables).

4 Discussion

There are multiple promising opportunities for integrating machine learning innovations into behavioral science research. In particular, ML can help with various behavioral research designs, including heterogeneous treatment effects, stimulus sampling, and imputing missing data for studying complex phenomena that evolve over time. Having illustrated some strategies for this integration, we now provide more speculative ideas about other potential synergies.

First, the ML methodologies we discuss do not cover the entire spectrum of algorithms available. One notable omission is reinforcement learning (Sutton and Barto 1998), where agents learn via interactions with an environment. Because of reinforcement learning's ability to capture trial-and-error search and delayed reward, it accommodates for phenomena that are otherwise computationally difficult or impossible to integrate into estimable models, such as present bias in behavioral models for intertemporal choice with self-control (Fedus et al. 2019).

Second, our discussion did not address the use of ML to formulate personalized treatments for each test subject. Personalized and hyperlocal design may help improve statistical power by using data from past experiments to recommend more optimal stimuli options for experimenters when many are available (e.g., pictures or text excerpts). These recommendations may narrow down options and even help choose the right (personalized) stimulus for each participant, and thereby reduce noise in the manipulation. However, an

important concern about ML-based personalization is its tendency to homogenize recommendations, especially in systems where users provide feedback on the recommendations—meaning that, over time, recommendations become increasingly more similar within the same population (Chaney et al. 2018). One potential consequence is that personalization may artificially generate similarity in behavioral outcomes.

Third, there is a growing desire to develop explainable ML, that is, research on *how* ML systems make predictions and execute their actions (Rai 2020). A major shortcoming of many ML algorithms is the inability to interpret the prediction outputs. For example, it is typically not straightforward to determine the direction of the prediction and often difficult to assess what drives them. Hypothesis testing usually being the goal in behavioral science, being able to understand the "why" behind each ML prediction would further increase the usefulness of ML to researchers in this domain. One of the recent innovations in this stream of research includes the integration of conscious thought into the priors used in ML algorithms (Bengio 2019) such that agents can make decisions based on statements about reality which are true, likely, and/or useful for these decisions (i.e., higher level of abstraction that humans often use in their thought processes).

Ultimately, this primer offers suggestions for new interdisciplinary research. As these fields continue to evolve, we anticipate a growing number of behavioral research topics that can be aided by these new methods.

References

Agarwal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines*. Harvard Business Review Press.

Anand, P., Thomas, M., Pillai, K. G., Meng-Lewis, Y. (2020). Linguistic analysis of psychological distancing: reading between the lines for unexpressed bad news. Working paper.

Aral, S., & Nicolaides, C. (2017). Exercise contagion in a global social network. *Nature Communications, 8*, 1–8.

Ascarza, E. (2018). Retention futility: Targeting high-risk customers might be ineffective. *Journal of Marketing Research, 55*, 80–98.

Bengio, Y. (2019). The consciousness prior. Working paper.

Berger, J., & Milkman, K. L. (2012). What makes online content viral. *Journal of Marketing Research, 49*(2), 192–205.

Berger, J., Humphreys, A., Ludwig, S., Moe, W., Netzer, O., & Schweidel, D. (2019). Uniting the tribes: Using text for marketing insight. *Journal of Marketing* Articles-in-Advance.

Blanchard, S., Dyachenko, T., & Kettle, K. (2020). Locational choices: Studying consumer preference for proximity to others. *Journal of Marketing Research* forthcoming.

Bollinger, B., Gillingham, K., Kirkpatrick, J., Sexton, S. (2019). Visibility and peer influence. Working Paper.

Brewer, M., & Weber, J. (1994). Self-evaluation effects of interpersonal versus intergroup social comparison. *Journal of Personality and Social Psychology, 66*, 268–275.

Chakraborty, I., Kim, M., Sudhir, K. (2019). Attribute sentiment scoring with online text reviews: Deep learning and accounting for attribute self-selection. Working Paper.

Chaney, A. J. B., Stewart, B. M., & Engelhardt, B. E. (2018). How algorithmic confounding in recommendation systems increases homogeneity and decreases utility. In *Proceedings of the 12th ACM Conference on Recommender Systems* (pp. 224–232).

Chapelle, O., Schölkopf, B., & Zien, A. (Eds.). (2006). *Semi-supervised learning*. Cambridge: MIT Press.

Dzyabura, D., & Yoganarasimhan, H. (2018). Machine learning and marketing. In D. Hanssens & N. Mizik (Eds.), *Handbook of marketing analytics: Methods and applications in marketing, Public policy, and litigation support*. Cheltenham: Edward Elgar Publishing.

Etkin, J. (2016). The hidden cost of personal quantification. *Journal of Consumer Research, 42*, 967–984.

Fedus, W., Gelada, C., Bengio, Y., Bellmare, M., Larochelle, H. (2019). Hyperbolic discounting and learning over multiple horizons. Working paper.

Fishbach, A., & Touré-Tillery, M. (2013). Goals and motivation. In R. Biswas-Diener & E. Diener (Eds.), *Noba textbook series: Psychology*. Champaign: DEF Publishers.

Gordon, M., Althoff, T., & Leskovec, J. (2019). Goal-setting and achievement in activity tracking apps: A case study of MyFitnessPal. *WWW, 13-17*(2019), 1–12.

Guo, T., Sriram, S., Manchanda, P. (2019). The effect of information disclosure on industry payments to physicians. Working paper.

Hagen, L. (2020). Pretty healthy food: How and when aesthetics enhance perceived healthiness. *Journal of Marketing* (forthcoming).

Hartford, J., Wright, J. R., & Leyton-Brown, K. (2016). Deep learning for predicting human strategic behavior. *Thirtieth Annual Conference on Neural Information Processing Systems*, 1–9.

Hui, S., Bradlow, E., & Fader, P. (2009). Testing behavioral hypotheses using an integrated model of grocery store shopping path and purchase behavior. *Journal of Consumer Research, 36*, 478–493.

Kettle, K., Trudel, R., Blanchard, S., & Häubl, G. (2016). Repayment concentration and consumer motivation to get out of debt. *Journal of Consumer Research, 43*, 460–477.

Kim, S. Y., Lewis, M., and Wang, Y. (2019). Physical store openings and product purchase and return behaviors: A quasi-experimental approach using the causal Forest method. Working Paper.

Knäuper, B., Carriare, K., Frayn, M., Ivanova, E., Xu, Z., Ames-Bull, A., Islam, F., Lowensteyn, I., Sadikaj, G., Luszczynska, A., Grover, S., & McGill CHIP Healthy Weight Program Investigators. (2018). The effects of if-then plans on weight loss: Results of the McGill CHIP healthy weight program randomized controlled trial. *Obesity, 26*, 1285–1295.

Liu, L., Dzyabura, D., & Mizik, N. (2019). Visual listening in: Extracting brand image portrayed on social media. *Marketing Science* Articles-in-Advance.

Lockwood, P., Wong, C., McShane, K., & Dolderman, D. (2005). The impact of positive and negative fitness exemplars on motivation. *Basic and Applied Social Psychology, 27*, 1–13.

Lu, S., Xiao, L., & Ding, M. (2016). A video-based automated recommender (VAR) system for garments. *Marketing Science, 35*, 484–510.

Lu, T., Bradlow, E., Hutchinson, W. (2017). Binge consumption of online content. Working paper.

Netzer, O., Feldman, R., Goldenberg, J., & Fresko, M. (2012). Mine your own business: Market-structure surveillance through text mining. *Marketing Science, 31*, 521–543.

Nevskaya, Y., & Albuquerque, P. (2019). How should firms manage excessive product use? A continuous-time demand model to test reward schedules, notifications, and time limits. *Journal of Marketing Research, 56*, 379–400.

Pelham, B., & Wachsmuth, J. (1995). The waxing and waning of the social self: Assimilation and contrast in social comparison. *Journal of Personality and Social Psychology, 69*, 825–838.

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: A tutorial overview. *Neuroimage, 45*, 199–209.

Puranam, D., Narayan, V., & Kadiyali, V. (2017). The effect of calorie posting regulation on consumer opinion: A flexible latent Dirichlet allocation model with informative priors. *Marketing Science, 36*, 726–746.

Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science, 48*, 137–141.

Sanakoyeu, A., Bautista, M., & Ommer, B. (2018). Deep unsupervised learning of visual similarities. *Pattern Recognition, 78*, 331–343.

Suher, J., Huang, S.-c., & Lee, L. (2019). Planning for multiple shopping goals in the marketplace. *Journal of Consumer Psychology, 29*, 642–651.

Sutton, R., & Barto, A. (1998). *Reinforcement learning: An introduction*. Cambridge: MIT Press.

Tanaka, K., & Titterington, D. M. (2004). Probabilistic image processing based on the Q-Ising Model by means of the mean-field method and loopy belief propagation. In *Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04)* (pp. 1–4).

Uetake, K., & Yang, N. (2019). Inspiration from the "biggest loser": Social interactions in a weight loss program. *Marketing Science* Articles-in-Advance.

Wager, S., & Athey, S. (2018) Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association, 113*(523), 1228–1242. https://doi.org/10.1080/01621459.2017.1319839

Wang, Y., Lewis, M., Cryder, C., & Sprigg, J. (2016). Enduring effects of goal achievement and failure within customer loyalty programs: A large-scale field experiment. *Marketing Science, 35*, 565–575.

Zhu, Y., Yu, Z., & Cheng, G. (2019). High dimensional inference in partially linear models. *Proceedings of Machine Learning Research, 89*, 2760–2769.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Linda Hagen[1] · Kosuke Uetake[2] · Nathan Yang[3] · Bryan Bollinger[4] · Allison J. B. Chaney[5] · Daria Dzyabura[6] · Jordan Etkin[5] · Avi Goldfarb[7] · Liu Liu[8] · K. Sudhir[9] · Yanwen Wang[10] · James R. Wright[11] · Ying Zhu[12]

Bryan Bollinger
bryan.bollinger@stern.nyu.edu

Allison J. B. Chaney
ajb.chaney@duke.edu

Daria Dzyabura
ddzyabura@nes.ru

Jordan Etkin
jordan.etkin@duke.edu

Avi Goldfarb
agoldfarb@rotman.utoronto.ca

Liu Liu
liu.liu-1@colorado.edu

K. Sudhir
k.sudhir@yale.edu

Yanwen Wang
yanwen.wang@sauder.ubc.ca

James R. Wright
james.wright@ualberta.ca

Ying Zhu
yiz012@ucsd.edu

[1] University of Southern California, Los Angeles, CA, USA

[2] Yale University, New Haven, CT, USA

[3] Cornell University, Ithaca, NY, USA

[4] New York University, New York City, NY, USA

[5] Duke University, Durham, NC, USA

[6] New Economic School, Moscow, Russia

[7] University of Toronto, Toronto, Canada

[8] University of Colorado Boulder, Boulder, CO, USA

[9] Yale University, New Haven, CT, USA

[10] University of British Columbia, Vancouver, Canada

[11] University of Alberta, Edmonton, Canada

[12] University of California San Diego, San Diego, CA, USA