Anna (Anya) Shchetkina

phone: +1-215-594-98-24
email: annashch@wharton.upenn.edu website: anyashchetkina.github.io

---

Education

The Wharton School, University of Pennsylvania **2026**
Ph.D. in Marketing (expected)
Committee: Ron Berman (chair), Ryan Dew, Raghuram Iyengar

New Economic School, Moscow, Russia **2021**
B.A. in Economics, *Summa Cum Laude*

Research Interests

Probabilistic Machine Learning and Causal Inference applied to Advertising, Targeting and Personalization, Privacy, and Experimentation

Job Market Paper

"Blind Targeting: Personalization under Third-Party Privacy Constraints" Anya Shchetkina

Working Papers

"When Is Heterogeneity Actionable for Personalization?" Anya Shchetkina and Ron Berman. Accepted for publication as an extended abstract at ACM Conference on Economics and Computation (EC'24). Major revision at *Management Science*.

"Your MMM is Broken: Identification of Nonlinear and Time-Varying Effects in Marketing Mix Models" Ryan Dew, Nicolas Padilla, and Anya Shchetkina. *Authors contributed equally.* Risky revision at *Journal of Marketing Research*.

"Increasing Interest in Claiming a Tax Credit: Evidence from Two Large-Scale A/B/n Field Experiments Among Lower Income People" Wendy De La Rosa et al. Reject and Resubmit at *Marketing Science*.

Honors and Awards

American Marketing Association - Sheth Foundation Doctoral Fellow **2025**

INFORMS Society for Marketing Science Doctoral Fellow **2025**

The Winkelman Fellowship **2023-2025**
Awarded to one rising 3rd year PhD student annually who has shown the greatest academic job potential across all departments at Wharton

Andrei Bremzen Award for Contribution to the Dissemination of Economic and Financial Knowledge to a Wide Audience **2023**
New Economic School

Bachelor of Excellence Award **2021**
New Economic School

Best Undergraduate Thesis Award **2021**
New Economic School, for "Why Do Customers Return More Over Time?"

Increased Academic Scholarship **2019**

Alfa-Bank Scholarship **2017-2019**

First female winner of Russian National Olympiad for High School Students in Economics **2017**

Research Talks

"Blind Targeting: Personalization under Third-Party Privacy Constraints"
Marketing Science Conference (Washington DC) **2025**

"Your MMM is Broken: Identification of Nonlinear and Time-Varying Effects in Marketing Mix Models"
Conference on Digital Experimentation (CODE) at MIT (poster session) **2024**

"When Is Heterogeneity Actionable for Personalization?"
ACM Conference on Economics and Computation (EC'24) **2024**
Causal Data Science Meeting **2023**
Conference on Digital Experimentation (CODE) at MIT (poster session) **2023**

"Autoregressive Difference-in-differences"
Marketing Science Conference (Miami) **2023**
11th Wharton-INSEAD Doctoral Consortium **2022**

Teaching Experience

The Wharton School, University of Pennsylvania
Teaching Assistant, Pricing Policy **2023-2025**
New Economic School, Moscow, Russia
Teaching Assistant, Introduction to Economics **2019**
Teaching Assistant, Microeconomics I **2018**
Teaching Assistant, Calculus I **2018**

Work Experience

Jellyfish (advertising and marketing analytics agency) **2023**
Data Science Intern
Columbia Business School **2019-2021**
Research Assistant
New Economic School **2019**
Research Assistant

Service

Organizer of the 13th Wharton-INSEAD Doctoral Consortium
Ad-hoc reviewer for ACM EC'24

References

Ron Berman
Associate Professor of Marketing
The Wharton School
ronber@wharton.upenn.edu

Ryan Dew
Assistant Professor of Marketing,
Govil Family Faculty Scholar
The Wharton School
ryandew@wharton.upenn.edu

Raghuram Iyengar
Miers-Busch, W'1885 Professor of Marketing
The Wharton School
riyengar@wharton.upenn.edu

Research abstracts

“Blind Targeting: Personalization under Third-Party Privacy Constraints” Anya Shchetkina (Job market paper).

Major advertising platforms recently increased privacy protections by limiting advertisers’ access to individual-level data. Instead of providing access to granular raw data, the platforms only allow a limited number of aggregate queries to a dataset, which is further protected by adding differentially private noise. This paper studies whether and how advertisers can design effective targeting policies within these restrictive privacy preserving data environments. To achieve this, I develop a probabilistic machine learning method based on Bayesian optimization, which facilitates dynamic data exploration. Since Bayesian optimization was designed to sample points from a function to find its maximum, it is not applicable to aggregate queries and to targeting. Therefore, I introduce two innovations: (i) integral updating of posteriors which allows to select the best regions of the data to query rather than individual points and (ii) a targeting-aware acquisition function that dynamically selects the most informative regions for the targeting task. I identify the conditions of the dataset and privacy environment that necessitate the use of such a “smart” querying strategy. I apply the strategic querying method to the Criteo AI Labs dataset for uplift modeling (Diemert et al., 2018) that contains visit and conversion data from 14M users. I show that an intuitive benchmark strategy only achieves 33% of the non-privacy-preserving targeting potential in some cases, while my strategic querying method achieves 97-101% of that potential, and is statistically indistinguishable from Causal Forest (Athey et al., 2019): a state-of-the-art non-privacy-preserving machine learning targeting method.

“When Is Heterogeneity Actionable for Personalization?” Anya Shchetkina and Ron Berman.

Targeting and personalization policies can be used to improve outcomes beyond the uniform policy that assigns the best performing treatment in an A/B test to everyone. Personalization relies on the presence of heterogeneity of treatment effects, yet, as we show in this paper, heterogeneity alone is not sufficient for personalization to be successful. We develop a statistical model to quantify ``actionable heterogeneity," or the conditions when personalization is likely to outperform the best uniform policy. We show that actionable heterogeneity can be visualized as crossover interactions in outcomes across treatments and depends on three population-level parameters: within-treatment heterogeneity, cross-treatment correlation, and the variation in average responses. Our model can be used to predict the expected gain from personalization prior to running an experiment and also allows for sensitivity analysis, providing guidance on how changing treatments can affect the personalization gain. To validate our model, we apply five common personalization approaches to two large-scale field experiments with many interventions that encouraged flu vaccination. We find an 18% gain from personalization in one and a more modest 4% gain in the other, which is consistent with our model. Counterfactual analysis shows that this difference in the gains from personalization is driven by a drastic difference in within-treatment heterogeneity. However, reducing cross-treatment correlation holds a larger potential to further increase personalization gains. Our findings provide a framework for assessing the potential from personalization and offer practical recommendations for improving gains from targeting in multi-intervention settings.

“Your MMM is Broken: Identification of Nonlinear and Time-Varying Effects in Marketing Mix Models” Ryan Dew, Nicolas Padilla, and Anya Shchetkina. *Authors contributed equally.*

Recent years have seen a resurgence in interest in marketing mix models (MMMs), which are aggregate-level models of marketing effectiveness. Often these models incorporate nonlinear effects, and either implicitly or explicitly assume that marketing effectiveness varies over time. In this paper, we show that nonlinear and time-varying effects are often not separately identifiable: while certain data patterns may be suggestive of nonlinear effects, such patterns may also emerge under simpler models with time-varying effects. Moreover, problematically, these two types of effects may suggest fundamentally different optimal marketing allocations. We examine this identification issue through theory and simulations, describing the conditions under which conflation between the two types of models is likely to occur. We show that conflating the two types of effects is especially likely in the presence of autocorrelated marketing variables, which are common in practice, especially given the widespread use of stock variables to capture long-run effects of advertising. We illustrate these ideas through numerous empirical applications to real-world marketing mix data, showing the prevalence of the conflation issue in practice. Finally, we show how marketers can avoid this conflation, by designing experiments that strategically manipulate spending in ways that pin down model form.