

ADAM KAPELNER

516-435-6795 ◊ kapelner@qc.cuny.edu ◊ (updated: October 21, 2021)

RESEARCH INTERESTS

Data Science, Machine Learning, Experimental Design, Statistical Software, Crowdsourced Social Science Experiments, Biomedical Applications, Educational Technology

ACADEMIC EMPLOYMENT

Queens College

Associate Professor of Mathematics

Nov 2020 - present

Director, Undergraduate Data Science and Statistics Program

Aug 2019 - present

Assistant Professor of Mathematics

Aug 2014 - Nov 2020

The Technion, Israel Institute of Technology

Jul 2018 - present

Visiting Scholar, Faculty of Industrial Engineering & Management

EDUCATION

Wharton School of the University of Pennsylvania

May 2014

Ph.D. in Statistics advised by Abba Krieger and Edward George

Wharton School of the University of Pennsylvania

May 2012

A.M. in Statistics advised by Dean Foster

Stanford University

June 2006

B.S. in Mathematical & Computational Science
(minors in Physics and Economics)

PUBLICATIONS

citations: 1953, h-index: 14 and i10-index: 16

Methodology

- Kapelner, A. & Krieger, A. (2021) A Matching Procedure for Sequential Experiments that Iteratively Learns which Covariates Improve Power *in press at Biometrics*
- Krieger, A., Azriel, D. & Kapelner, A. (2021) Better Experimental Design by Hybridizing Binary Matching with Imbalance Optimization *in press at the Canadian Journal of Statistics*
- Kapelner, A., Shalit, U., Krieger, A., Sklar, M. & Azriel, D. (2020) Harmonizing Optimized Designs with Classic Randomization in Experiments. *The American Statistician* 75 (2), 195–206
- Krieger, A., Azriel, D. & Kapelner, A. (2019) Nearly Random Designs with Greatly Improved Balance. *Biometrika* 106 (3), 695–70
- Kapelner, A., Krieger, A. & Blanford, W. J. (2016). Optimal Experimental Designs for Estimating Henry's Law Constants via the Method of Phase Ratio Variation. *Journal of Chromatography A* 1468, 183–191

- Kapelner, A. & Krieger, A. (2014). Matching on-the-fly in Sequential Experiments for Higher Power and Efficiency. *Biometrics* 70 (2), 378–388
- Blanford, W. J., Jofat, D. & Kapelner, A. (2020). Solution Density Models as Functions of Sodium Chloride, Hydroxypropyl- β -cyclodextrin, and Temperature (278.15–333.15 K) via Progressive Linear and Stepwise Regression. *Journal of Chemical & Engineering Data* 65 (10), 4735–4750
- Mehlretter, J., Fratila, R., Benrimoh, D.A., Kapelner, A., Perlman, K., Snook, E., Israel, S., Miresco, M. & Turecki, G. (2020) Differential Treatment Benefit Prediction For Treatment Selection in Depression: A Deep Learning Analysis of STAR*D and CO-MED Data *Computational Psychiatry* 4, 61–75
- Kapelner, A., Bleich, J., Levine, A., Cohen, Z. D., DeRubeis, R. J. & Berk, R. A. (2021) Evaluating the Effectiveness of Personalized Medicine with Software. *Frontiers in Big Data — Medicine and Public Health* 4 (8), 1–19
- Kapelner, A., Sklar M., Krieger, A., & Azriel, D. Optimal Rerandomization via a Criterion that Provides Insurance Against Failed Experiments. *in second review at Journal of Statistical Planning and Inference*
- Kleinerman, A., Rosenfeld, A., Benrimoh, D., Fratila, R., Armstrong, C., Mehlretter, J., Yaniv-Rosenfeld, A., Turecki, G. & Kapelner, A. Automated Treatment Selection Latent-space Prototyping Depression Treatment *in review at Artificial Intelligence*
- Krieger, A., Azriel, D., Sklar, M. & Kapelner, A. Improving the Power of the Randomization Test *in review at Communications in Statistics — Theory and Methods*
- Kleinerman, A., Rosenfeld, A., Benrimoh, D., Fratila, R., Armstrong, C., Mehlretter, J., Shneider, E., Yaniv-Rosenfeld, A., Karp, J., Reynolds, C.F., Turecki, G & Kapelner, A. Treatment selection using prototyping in latent-space with application to depression treatment *in review at PLOS ONE*
- Azriel, D., Kallus, N. & Kapelner, A. Optimal Regret Designs in a Sequential Trial. *in preparation for Biometrics*
- Kapelner, A. & Krieger, A. Robust Experimental Designs via Simultaneous Minimization of Multiple Kernel Distances *in preparation for Journal of Computational and Graphical Statistics*
- Krieger, A., Azriel, D. & Kapelner, A. Optimal Experimental Design for Incidence Endpoints *in preparation for JRSS-B*

Machine Learning and Data Science

- Kapelner, A. & Bleich, J. (2016). `bartMachine`: Machine Learning with Bayesian Additive Regression Trees. *Journal of Statistical Software* 70 (4)
- Kapelner, A. & Bleich, J. (2014). Prediction with Missing Data via Bayesian Additive Regression Trees. *Canadian Journal of Statistics* 43 (2) 224–239
- Bleich, J., Kapelner, A., George, E. I. & Jensen, S. T. (2014). Variable Selection Inference for Bayesian Additive Regression Trees. *Annals of Applied Statistics* 8 (3) 1750–1781
- Goldstein, A., Kapelner, A., Bleich, J. & Pitkin, E. (2014). Peeking Inside the Black Box: Visualizing Statistical Learning with Plots of Individual Conditional Expectation. *Journal of Computational & Graphical Statistics* 24(1), 44–65

- Bleich, J & Kapelner, A. Bayesian Additive Regression Trees With Parametric Models of Heteroskedasticity. *Arxiv*
- Berk, R., Bleich, J., Kapelner, A., Henderson, J., Barnes, G., Kurtz, E. Using Regression Kernels to Forecast A Failure to Appear in Court *Arxiv*
- Kapelner, A., Bleich, J. & Berk, R. A. An R Package for Kernel Regression with PCA, *in preparation for Statistical Surveys*
- Kapelner, A. YARF: A Fully-Flexible Non-Parametric Learning Suite *in preparation for Journal of Statistical Software*

Crowdsourcing and Social Science

- Kapelner, A. & Weinberg. Do Readers Care about Author Race and Gender? Results from a Randomized Experiment across Genres. *in preparation for PLoS One*
- Kapelner, A. & Weinberg. (2019) Do Readers Judge Books by Author Gender? Results from a Randomized Experiment. *Socius* 5
- Weinberg, D. B. & Kapelner, A. (2018) Comparing gender discrimination and inequality in indie and traditional publishing. *PLoS One* 13 (4) e0195298
- Schwartz, H. A., Eichstaedt, J., Blanco, E., Agrawal, M., Dziurzyński, L., Kern, M. L., Kapelner, A., Park, G., Jha, S., Stillwell, D., Kosinski, M. & Ungar, L. H. (2016) Predicting individual well-being through the language of social media. *Biocomputing: Proceedings of the Pacific Symposium* 516–527
- Chandler, D. & Kapelner, A. (2013) Breaking Monotony with Meaning: Motivation in Crowdsourcing Markets. *Journal of Economic Behavior & Organization*, 90: 123-133
- Kapelner, A., Kaliannan, K., Schwartz, H. A., Ungar, L. H. & Foster, D. P. (2012) New Insights from Coarse Word Sense Disambiguation in the Crowd. *CoLING*
- Kapelner, A. & Chandler, D. (2010) Preventing Satisficing in Online Surveys. *Proceedings of CrowdConf*

Education Applications

- Adlof, A., Baron, L. S., Scoggins, J., Kapelner, A., McKeown, M. G., Perfetti, C., Miller, E., Soterwood, J. & Petscher, Y. (2019) Accelerating Adolescent Vocabulary Growth: Development of an Individualized, Web-based, Vocabulary Instruction Program. Language, Speech, and Hearing Services in Schools. *Language, Speech, and Hearing Services in Schools* 50 (4): 579-595
- Kapelner, A., Nessaiver, S., Soterwood, J. & Adlof, A. (2018) Predicting Contextual Informativeness for Vocabulary Learning. *IEEE Transactions on Learning Technologies* 11(1) 13-26
- Adlof, A., McKeown, M., Perfetti, C., Petscher Y. & Kapelner, A. Demonstrating the Effectiveness of a New Vocabulary Teaching Method with an RCT. *in preparation for Language, Speech, and Hearing Services in the Schools*

Biomedical Applications

- Benrimoh, D., Israel, S., Fratila, R., Armstrong, C., Perlman, K., Rosenfeld, A. & Kapelner, A (2021) Editorial: ML and AI Safety, Effectiveness and Explainability in Healthcare. *Frontiers in Big Data* 4, 1–54
- Bleich, J., Cole, B., Kapelner, A., Baillie, C. A., Gupta, R., Hanish, A., Calgua, E., Umscheid, C. & Berk, R. (2021) Using Random Forests with Asymmetric Costs to Predict Hospital Readmissions *medrxiv*
- Clarke, G. P. & Kapelner, A. (2020) The BART Formula for Safe Machine-Learning Based IOL Predictions. *Frontiers in Medicine* 3 (46) 1–10
- Schoeler, N., Bell, G., Yuen, A., Kapelner, A., Heales, S. J. R., Cross, J. H. & Sisodiya, S. (2017) Acetyl carnitine and association with response to ketogenic dietary therapies. *Epilepsia* 58 (5), 893-900
- Kapelner, A. & Vorsanger, M. (2015) Starvation of Cancer via Induced Ketogenesis and Severe Hypoglycemia. *Medical Hypotheses*, 84(3): 162–168
- Chang, A. Y., Bhattacharya, N., Mu, J., Setiadi, A. F., Carcamo-Cavazos, V., Lee, G. H.; Simons, D. L., Yadegarynia, S., Hemati, K., Kapelner, A., Zheng, M., Krag, D. N., Schwartz, E. J., Chen, D. Z. & Lee, P. P. (2013) Spatial organization of dendritic cells within tumor draining lymph nodes impacts clinical outcome in breast cancer patients. *Journal of translational medicine*, 11(1): 242
- Setiadi, A. F.; Ray, N. C., Kohrt, H. E., Kapelner, A., Carcamo-Cavazos, V., Levic, E. B., Yadegarynia, S., van der Loos, C. M., Schwartz, E. J., Holmes, S. & Lee, P. P. (2010) Quantitative, architectural analysis of immune cell subsets in tumor-draining lymph nodes from breast cancer patients and healthy lymph nodes. *PloS one*, 5(8): e12420
- Holmes, S., Kapelner, A. & Lee, P. P. (2009) An interactive java statistical image segmentation system: Gemident. *Journal of Statistical Software*, 30(10): 1–20
- Kapelner, A., Lee, P. P. & Holmes, S. (2007) An interactive statistical image segmentation and visualization system. *in proceedings of IEEE, Medivis*

GRANTS AWARDED

- | | |
|--|---------------------|
| Israel-USA Binational Science Foundation (BSF) | 2019-2023 |
| More Powerful Experiments via Harmonizing Classic Randomization with Modern Optimization, 2018112, \$167,000 Co-PI | |
| PSC CUNY | 2019-2020 |
| A Natural Field Experiment on Race and Gender Discrimination in the Gig Economy, TRADB-50-65, \$6,000 | |
| MQ Foundation | May 2017 - Jul 2018 |
| The Stratified Medicine Approaches for Treatment Selection Mental Health Prediction Tournament, \$3,200 • Winner of tournament | |
| PSC CUNY | 2017-2018 |
| Optimal Experimentation: Trading Randomization for Balance, TRADA-48-469, \$3,500 | |
| MQ's Psy-IMPACT | Nov 2014 - Dec 2018 |
| • Serve as an expert consultant and investigator to the grant | |

U.S Dept. of Education, Inst. of Educational Science Sept 2013 - Aug 2018
Grant #R305A130467, \$1,500,000, PI: Suzanne Adlof, University of South Carolina. I
cowrote and served as the expert consultant to the grant.

GRANTS IN REVIEW

ONR N00014-20-S-F003, \$3.97M Co-PI 2022 - 2026
Malign Information Operations, Narrative Resonance, and the Fracturing of American Identity
NSF Environmental Engineering program PD 20-1440, \$422K 2022 - 2026
Impact, Analysis, and Advanced Modeling of Miscible Displacement Transport Studies of
Monoaromatic Compounds in Sandstone Cores

GRANTS APPLIED

ONR N00014-20-S-F005 (applied 2020), \$750K Co-PI 2021 - 2024
Data Science as Pathway to Naval Careers for Diverse Students at CUNY Queens College
NSF PD18-1269 (applied 2018), \$500K Co-PI 2019 - 2023
More Powerful Experiments via Harmonizing Classic Randomization with Optimization
PSC CUNY (applied 2015) TRADA-47-330, \$2500 2016
Demonstrations of Inference for Personalized Medicine Treatment Models

HONORS AND AWARDS

Certificate for Highly Cited Research in the Journal
of Economic Behavior and Organization Jan 2017
National Science Foundation Graduate Research Fellowship May 2010 - April 2013
J. Parker Bursk Memorial Award for Excellence in Research Dec 2013
Donald S. Murray Award for Excellence in Teaching Dec 2012
Intel Science Talent Search Semifinalist Jan 2002

FORMAL PRESENTATIONS AND TALKS

Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Harvard University, Applied Statistics Lecture Series Feb 2020
Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Design & Analysis of Experiments Conference Oct 2019
Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Atlantic Causal Inference Conference May 2019
Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Wharton Statistics Faculty Seminar Mar 2019
Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Economics Dept Seminar, Queens College Feb 2019
Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, *EMR-IBS, '18, Jerusalem, Israel* Dec 2018

Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Haifa University Statistics Faculty Seminar Oct 2018

Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Tel Aviv University Statistics Faculty Seminar Oct 2018

Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Hebrew University of Jerusalem Statistics Faculty Seminar Oct 2018

Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Weizmann Institute of Science Statistics Faculty Seminar Oct 2018

Kapelner A., Personalized Medicine Inference via the R Package PTE, Industrial Engineering and Management Seminar, the Technion July 2018

Kapelner A., Personalized Medicine Modeling with Survival and Incidence Endpoints *TSIL, '18, London, England* Jun, 2018

Kapelner A., Personalized Medicine Inference and Machine Learning via the R Packages PTE & YARF *TSIL 1/2 Day Workshop, '18, London, England* Jun, 2018

Kapelner, A., Weighted Matching on-the-fly: Improved Sequential Allocation with Higher Power and Efficiency *SAE '18 Shanghai, China* Jun, 2018

Kapelner A., Weighted Matching on-the-fly: Improved Sequential Allocation with Higher Power and Efficiency, Technical Seminar, Amazon Inc. June 2018

Kapelner, A., Starving Cancer through Induced Ketogenesis *QC Biology Symposium* Jan, 2018

Kapelner, A., YARF: A Fully Customizable Non-Parametric Regression Toolbox *The Technion, seminar series* Jul, 2017

Kapelner, A., Weighted Matching on-the-fly: Improved Sequential Allocation with Higher Power and Efficiency *The 6th International Workshop in Sequential Methodologies (IWSM '17 Rouen, France)* Jun, 2017

Kapelner, A., Predicting Contextual Informativeness for Vocabulary Learning *Kasisto, Inc. seminar series* Nov, 2016

Kapelner, A., Optimal experimental designs for estimating Henry's law constants via the phase ratio method *ACS National Meeting* Aug, 2016

Clarke, G., Hill, W., Kapelner, A. Data-Driven IOL Calculations *Amer. Society of Cataract and Refractive Surgeons Annual Meeting* May, 2016

Kapelner, A., Inference for Personalized Medicine Models *UPenn Treatment Lab* June, 2016

Kapelner, A., Better Experiments on MTurk? NYU Statistics Seminar May, 2016

Jensen, S., Kapelner, A., Variable Selection with Bayesian Additive Regression Trees. ENAR Mar, 2015

Kapelner, A., Experiments via Crowdsourcing: A New Platform for Social Science Research? Economics Dept Seminar, Queens College Feb 2015

Kapelner, A., Better Randomization via Greedy Pair Switching *IMS China* Jul, 2015

Kapelner, A., Ungar L. Crowdsourcing for Statisticians. *JSM*, Continuing Education Course Aug 2013

Kapelner, A., Chandler D. Preventing Satisficing in Online Surveys. *CrowdConf* Oct 2010

Wieland, K., Fitzgerald, J., Kapelner, A. Contextual Vocabulary Analysis Processes. *National Reading Conference* Dec 2009

INFORMAL PRESENTATIONS AND TALKS

Episode 109 - Experimental Design *Local Max (Data Science Podcast)* Mar 9, 2020
The Data Science & Statistics Mathematics Major Option at QC *NYC Tech Talent Pipeline Conference* May 14, 2018
Predicting Informativeness from Context *QC Math Club* Feb 5, 2018
How to Teach an Effective Summer Course (for Wharton graduate students) Wharton, May, 2012

TEACHING EXPERIENCE

Queens College, City University of New York

Math 369/690.3 (Statistical Inference) Aug 2020 - present
Math 342W/650.4 (Data Science via Machine Learning and Statistical Modeling with the R Language) Jan 2018 - present
Math 368/621 (Advanced Probability) Aug 2017 - present
Math 341/650.3 (Bayesian Modeling) Feb 2016 - present
Math 241 (Probability and Statistics) Aug 2014 - present

Wharton School of the University of Pennsylvania

Stat 422/722 (Predictive Analytics) Jan 2017 - Feb 2017
Stat 101 (Probability and Statistics) May 2011 - July 2011
Teaching Assistant for Stat 101 (Probability and Statistics) and Statistics 102 (Linear Regression) Sept 2009 - June 2010
Teaching Assistant for Stat 613 (Required Statistics course for MBA students) Sept 2013 - Dec 2013
Teaching Assistant for Stat 112 (Statistical Inference) Jan 2014 - Jun 2014

MENTORING EXPERIENCE & STUDENT INDEPENDENT STUDIES

Studies in Predicting Monotonic Breakout Curves using Machine Learning
Kennly Weerasinghe Spring, 2021 - present

Studies in Missing Data in Random Forests
Abhinav Patil Summer, 2020 - Fall, 2021

An Algorithm for Automatic Convergence of Random Forests
Rebecca Horowitz Summer, 2020

Studies in Optimal Experimental Design
Abhinav Patil Spring, 2020

A Nonparametric Bayesian Model for Extreme Events
Bracha Blau Fall, 2018

Studies in Fully Customizable Tree Models
Ashok Rao Spring, 2018

Optimal Experimental Design

Bracha Blau

Spring, 2018

A Nonparametric Bayesian Model for Extreme Events

Evangeline Spzylka

Summer, 2017 - Spring, 2018

Personalized Medicine Models for Survival

Alina Levine

Summer, 2016 - Spring, 2018

and Xin Ling Luang (local HS student)

Summer, 2016

Deep Learning for Image Segmentation

Christian Colon and Stefan Hernandez

Spring, 2016 - Spring, 2017

Predicting Congressmens' Party Affiliation

Savvas Tjortjoglou

Spring, 2016

Web Application Engineering for Social Science Experimentation

Rikki Katz

Spring, 2015

Studies in Real Analysis

Elliot Gangaram

Spring, 2015

NEW COURSES DEVELOPED

Queens College, City University of New York

Math 369 (Statistical Inference)

Sept 2020 - present

Math 368 (Advanced Probability)

Sept 2019 - present

Math 342W (Data Science Basics: Machine Learning and Statistical Modeling with the R Language)

Jan 2018 - present

Math 341 (Bayesian Modeling)

Feb 2016 - present

UNDERGRADUATE PROGRAMS DEVELOPED

Queens College, City University of New York

Data Science & Statistics Option for the BA in Mathematics

(with Chris Hanusa and Alan Sultan)

Sep 2018 - present

Actuarial Studies Track in the MA in Risk Management

(with Cara Marshall)

Jan 2018 - present

Data Science & Statistics Speaker Series

Feb 2019 - present

DEPARTMENT SERVICE

Position**dates**

Data Science & Statistics Option Advisor

9/19-present

Coaching and Interviewing Adjunct Instructors

9/19-present

Hiring Committee

9/17-present

Mathematics Minor Advisor

1/17-present

Mathematics Curriculum Committee

6/16-present

Learning Outcomes Committee	1/16-present
Faculty Development Committee	1/16-present

COLLEGE SERVICE

Position	dates
Academic Senate	6/16-present

PROFESSIONAL SERVICE (REVIEW WORK)

Journal (or Organization)	Year
Computational Statistics & Data Analysis	2021
Journal of Agricultural, Biological, and Environmental Statistics	2021
Open Statistics	2020
INFORMS Journal on Computing	2020
International Statistical Review	2020
Biometrika	2019
Journal of the American Statistical Association	2019
PSC CUNY Grant for Statistics Research	2019
Annals of Applied Statistics	2019
PLoS One	2019
Journal of Educational & Behavioral Statistics	2018
Neural Networks	2018
IEEE Access	2018
PLoS One	2018
The R Journal	2018
PSC CUNY Grant for Statistics Research	2018
Journal of the American Statistical Association	2017
Statistical Analysis and Data Mining	2017
PSC CUNY Grant for Statistics Research	2017
PLoS One	2016
The R Journal	2016
Statistical Analysis and Data Mining	2016
International Journal of Approximate Reasoning	2016
Journal of the Royal Statistical Society (Series B)	2015
Statistics in Medicine	2015
Bayesian Analysis	2014
Medical Hypothesis	2014
Journal of Cancer Research & Therapy	2014
Annals of Applied Statistics	2013
Transactions of the Association for Computational Linguistics	2013
National Science Foundation	2010

INDUSTRY EXPERIENCE

Data Science Private Consulting

June 2014 - Present

- Prediction modeling, data mining, statistical testing for a variety of clients from tech to biomedical to finance

DictionarySquared, Inc.

Founder & CTO

April 2010 - Dec 2018

San Francisco, CA

- Conceived and engineered a web application that teaches vocabulary via reading contextual snippets
- Wharton Business Plan Semifinalist Winner
- Applied and received federal grant money for research (see Grants section)

Eventbrite, Inc.

Software Engineer

April 2007- Aug 2007

San Francisco, CA

- First engineer. Designed and engineered portions of their web platform.

Stanford University, Lab of Peter Lee

Staff Scientist

June 2005 - Mar 2007

Stanford, CA

- Conceived and engineered software that finds objects in images, used to find cells in microscopic images. Uses Java-R programming, Random Forests and image processing

SOFTWARE

R Package: CovBalAndRandExpDes (in dev.)	Optimal Designs for Experiments
R Package: YARF (in dev.)	A highly customizable predictive modeling suite
R Package: SampleRepresentativenessChecker	Assesses sample representativeness
R Package: GreedyExperimentalDesign	A suite of better experimental designs
R Package: optDesignSlopeInt	Optimal designs for slope-to-intercept ratios
R Package: ICEbox	Statistical learning visualization suite
R Package: bartMachine	A flexible statistical learning suite
R Package: PTE	Inference for personalized medicine models
GemIdent	Finds objects of interest in images
GemVident	Finds objects of interest in videos
dictionaryquared	Teaches High School Students Vocabulary
gradesly.com	gives students grade transparency and helps professors administer grades

TECHNICAL STRENGTHS

Software Languages	R, Java, Ruby on Rails, C++, Python, HTML/CSS/JS
R Skills	ggplot2, dplyr, data.table, mlr3, rJava, Rcpp, most ML packages
Other Skills	git, postgresSQL, MySQL, Linux, Gurobi, grid computing, simulation, AWS