Colin Lewis-Beck

Contact Information

2823 18th Street
Apartment 634
San Francisco, CA 94110

E-mail: clewisbeck@gmail.com
Phone: (617) 784-3596
Web: Personal Website

Research Interests

Causal inference, machine learning, forecasting, reliability, Bayesian methods

Degrees Held

Ph.D., Statistics 2018
Iowa State University Ames, IA

Dual M.A./M.P.P., Applied Statistics and Public Policy 2010
University of Michigan Ann Arbor, MI

B.A., French with honors 2005
Middlebury College Middlebury, VT

Computer Skills

Scientific Programming: Python, PySpark, `R`, `Rcpp`, `RStan`, UNIX
Markup Languages: LaTeX, Markdown
Software Development: GitHub

Employment

Data Scientist, Research 2024 - present
Google Mountain View, CA

Research Scientist II 2022 - 2024
Amazon.com, Inc. Sunnyvale, CA

- Designed and coded a data quality monitoring package in Python that was added into a production pipeline.
- Deployed, summarized, and interpreted statistical/ML models to measure the business impact of cross-channel marketing investments.
- Explained causal inference DNN models, and interpreted model output, to business stakeholders during weekly office hours and through in-depth written analyses.
- Researched and implemented methods to calibrate causal models with estimates from randomized control trials.

Research Scientist, Statistics 2021 - 2022
Eli Lilly and Company Indianapolis, IN

- Assessed the impact of decentralized clinical trial study design on data quality using Cox proportional hazard model simulations. Results were published in an internal white paper to provide guidance for future studies.
- Lead the statistical programming team to deliver analysis reports for medical review meetings, and senior management review.
- Collaborated cross-functionally to monitor and clean the data of a large-scale clinical trial, leading to a successful interim analysis showing superiority of company molecule.

Visiting Assistant Professor 2018 - 2021
University of Iowa, Department of Statistics & Actuarial Science Iowa City, IA

- Developed Bayesian forecasting methodology to predict hard drive failures. Wrote and published a first author peer-reviewed paper in a top-tier statistical journal.
- Built Bayesian nonlinear hierarchical model to predict time of peak crop growth to inform agricultural strategy. Evaluated model fit using a combination of cross-validation and comparison with USDA survey data.

- Collaborated with researchers in engineering, statistics, and political science on applied statistical research resulting in ten peer-reviewed publications.

Software Development Intern Summer 2017
Google Summer of Code, NIMBLE Berkeley, CA
- Productionized `R` package, `nimbleEcology`, to dynamically generate model code allowing non-technical users to fit Bayesian ecological statistical models.
- Validated software using unit testing via the `testthat` package.
- Collaborated with statisticians and software developers on source code using GitHub.

Lead Computing Consultant Summer 2013, 2014
University of Michigan Ann Arbor, MI
- Managed a staff of eight Computer Consultants as part of the Interuniversity Consortium for Political and Social Research (ICPSR) Summer Program.
- Assisted program participants and faculty with a wide range of statistical and programming backgrounds with questions related to statistical software including `R`, Stata, SPSS, and SAS.

Statistical Analyst 2011 - 2012
STATinMED Research Ann Arbor, MI
- Analyzed large claim databases using advanced statistical techniques (e.g., GLMs, SEMs, propensity score matching, and meta-analysis).
- Drafted statistical protocols and wrote final manuscripts analyzing and interpreting study results for client, as well as academic, publication.
- Worked directly with clients, senior researchers, and programmers to ensure projects were completed correctly and on schedule.

Teaching Experience

Instructor Spring 2021
Stonehill College, Meehan School of Business Easton, MA
- *Quantitative Analysis* (*Online*)

Instructor 2018 - 2021
University of Iowa, Department of Statistics & Actuarial Science Iowa City, IA
- *Econometric Analysis* (S19)
- *Elementary Statistics* (S21)
- *Mathematical Statistics I* (F18, F19, F20)
- *Mathematical Statistics II* (S19, S20)
- *Statistics & Society* (F18, S19, F19, S20, F20, S21)

Instructor 2018 - 2020
University of Michigan, ICPSR Summer Program Ann Arbor, MI
- *Introduction to Meta-Analysis* (Su18)
- *Introduction to Regression Analysis* (Su19, Su20)

Instructor 2017 - 2018
Iowa State University, Department of Statistics Ames, IA
- *Statistical Methods for Research Workers (Graduate Course)* (Su18)
- *Probability and Statistical Inference for Engineers* (S18)
- *Engineering Statistics* (S17, F17)

Teaching Assistant 2014 - 2015
Iowa State University, Department of Statistics Ames, IA
- *Applied Statistical Modeling (Graduate Course)* (F15)

- *Introduction to Statistics* (F14, S15)

Graduate Student Instructor 2007 - 2010
University of Michigan, Department of Statistics Ann Arbor, MI
- *Introduction to Statistical Reasoning* (F09, S10)
- *Introduction to Statistics and Data Analysis* (S09)
- *Statistics for Public Policy (Graduate Course)* (S08, F08)
- *Introduction to Microeconomics for Public Policy* (F07)

Awards and Grants

University of Iowa Public Policy Center, Summer Scholars Grant, $3,000 (with Tom Rice) (2020)

Teaching Excellence Award, Iowa State University, Dept. of Statistics (2018)

SAGE Cornerstone Author Award for publication (with Michael Lewis-Beck) of Applied Regression: An Introduction, Second Edition (2015)

Outstanding Teaching Award, University of Michigan, Dept. of Statistics, $500 (2010)

Popular Press Interviews

Lynch, James Q. (2020, November 23). Record voter turnout masks Iowa schism. *The Gazette*.

Publications

Tian Q., **Lewis-Beck C.**, Niemi J., & Meeker W.Q. (2024). Rejoinder to Specifying Prior Distributions in Reliability Applications. *Applied Stochastic Models in Business and Industry*, 40(1), 130-143.

Tian Q., **Lewis-Beck C.**, Niemi J., & Meeker W.Q. (2024). Specifying Prior Distributions in Reliability Applications (with discussion). *Applied Stochastic Models in Business and Industry*, 40(1), 5-62.

Lewis-Beck C., Tian Q., & Meeker W.Q. (2022). Prediction of Future Failures for Heterogeneous Reliability Field Data. *Technometrics*, 64(1), 125-138.

Shiraef, Mary A, Hirst, Cora., [and 17 others, including **Lewis-Beck C.**] (2021). Border Accountability Project, a hand-coded global database of border closures introduced during 2020. *Scientific data*, 8(1), 1-11.

Berg E., Im J., Zhu Z., **Lewis-Beck C.**, & Li, J. (2021). Integration of Statistical and Administrative Agricultural Data from Namibia. *Statistical Journal of the IAOS*, 37(2), 557-578.

Lewis-Beck C., & Martini, N.F. (2020). Economic Perceptions and Voting Behavior in U.S. Presidential Elections. *Research and Politics*, 7(4), 1-6.

Togliatti, K., **Lewis-Beck C.**, Walker, V.A., Hartman, T., VanLoocke, A., Cosh, M.H., & Hornbuckle, B.K. (2020). Quantitative Assessment of Satellite L-Band Vegetation Optical Depth in the U.S. Corn Belt. *IEEE Geoscience and Remote Sensing Letters.*

Lewis-Beck C., Zhu Z., Walker V.A., & Hornbuckle B.K. (2020). Modeling Crop Phenology in the U.S. Corn Belt using Spatially Referenced SMOS Satellite Data. *Journal of Agricultural and Biological Statistics*, 25(4), 657-675.

Lewis-Beck, C., & Lewis-Beck, M.S. (2020). U.S. Presidential Election Forecasting:

The Economist Model. *Foresight: The International Journal of Applied Forecasting*, 59, 38-44.

Lewis-Beck, C., Walker, V.A., Niemi, J., Caragea, P., & Hornbuckle, B.K. (2020). Extracting Agronomic Information from SMOS Vegetation Optical Depth in the U.S. Corn Belt Using a Nonlinear Hierarchical Model. *Remote Sensing*, 12(5), 827.

Lewis-Beck C., Zhu Z., Mondal A., Jin Song J., Hobbs, J., Hornbuckle B.K, & Patton J. (2019). A Parametric Approach to Unmixing Remote Sensing Crop Growth Signatures. *Journal of Agricultural and Biological Statistics*, 24(3), 502-516.

Mittman, E., **Lewis-Beck, C.**, & Meeker, W.Q. (2019). A Hierarchical Model for Heterogeneous Reliability Field Data. *Technometrics*, 61(3), 354-368.

A. Alhasan, A. Ali, D. Offenbacker, O. Smadi, & **Lewis-Beck C.** (2018). Incorporating Spatial Variability of Pavement Foundation Layers Stiffness in Reliability-Based Mechanistic-Empirical Pavement Performance Prediction. *Transportation Geotechnics*, 17, 1-13.

Lewis-Beck, C., & Lewis-Beck, M.S. (2015). Applied Regression: An Introduction, Second Edition. SAGE Publications.

Lewis-Beck, C., Abouzaid, S., Xie, L., Baser, O., & Kim, E. (2013). Analysis of the relationship between psoriasis symptom severity and quality of life, work productivity, and activity impairment among patients with moderate-to-severe psoriasis using structural equation modeling. *Patient Preference and Adherence*, 7, 199-205.

Wang, L., **Lewis-Beck, C.**, Baser, E., Fritschel, E., & Baser, O. (2013). Applied Comparison of Meta-Analysis Techniques. *Value in Health*, 16(7), 14-22.

Presentations and Workshops

"Specifying Prior Distributions in Reliability Applications." Talk presented at the Joint Statistical Meetings, Portland, OR, August 2024.

"The MiM Machine Learning Attribution Model: A Case Study." Talk presented at the Amazon Prime & Marketing Science (PriMa) Conference, Amsterdam, NL, July 2023.

"Prediction of Future Failures for Heterogeneous Reliability Field Data." Invited seminar presented at the Bayesian Seminar Series, Eli Lilly, Virtual, September 2021.

"Social Capital and Shared Leadership in Small Iowa Communities." Invited talk (with Tom Rice) at the University of Iowa Public Policy Center, Virtual, February 2021.

"Forecasting the 2020 U.S. Elections." Invited talk (with Michael Lewis-Beck) at the Data Analytics Colloquium, University of Texas at Dallas and the National Chung Hsing University, Virtual, November 2020.

"Prediction of Future Failures for Heterogeneous Reliability Field Data." Poster presentation at the Joint Statistical Meetings, Virtual, July 2020.

"Using the M-RA Approximation to Integrate Multiple Data Sources on Temperature." Talk presented at the Joint Statistical Meetings, Denver, CO, July 2019.

"A Hierarchical Model for Heterogeneous Reliability Field Data." Invited seminar presented at the Department of Statistics and Actuarial Science, University of Iowa,

Iowa City, IA, March 2019.

"A Parametric Approach to Unmixing Remote Sensing Crop Growth Signatures." Talk presented at the Joint Statistical Meetings, Vancouver, B.C., August 2018.

"A Nonlinear Hierarchical Approach for Modeling Crop Growth in the US Corn Belt." Talk presented at the Kansas State University Conference on Applied Statistics in Agriculture, Manhattan, KS, May 2018.

"A Hierarchical Model for Heterogeneous Reliability Field Data." Poster presentation at the Joint Statistical Meetings, Baltimore, MD, August 2017.

Graduate Workshop on Environmental Data Analytics, National Center for Atmospheric Research, Boulder, CO, June 2017.

"An Introduction to Statistical Thinking for Forensic Practitioners." Invited talk (with Hal Stern) at the Center for Statistics and Applications in Forensic Evidence, Palm Beach County Sheriff's Office, Palm Beach, FL, March 2016.

"Regression Questions You Always Wanted to Ask." Invited Blalock lecturer (with Michael Lewis-Beck) at the Interuniversity Consortium for Political and Social Research, University of Michigan, July 2015.

"Analysis of relationship between psoriasis severity and quality of life, work productivity, and activity impairment among patients with moderate to severe psoriasis using structural equation modeling." Poster presented at International Society for Pharmacoeconomics and Outcomes Research, Washington, D.C., June 2012.

Editorial Service

Reviewer for the following journals: SAGE Publications (2017), Annals of Applied Statistics (2021), Reliability Engineering & System Safety (2023), Statistical Analysis and Data Mining (2024)

Professional Memberships

American Statistical Association