

COHEN R. SIMPSON

c.r.simpson@lse.ac.uk | <https://www.cohensimpson.com> | U.S. Citizen | U.K. Visa (Tier 1 Talent) Already In Hand

CORE SKILLS

Programming: R • Python • Spark SQL • LaTeX [Technical Docs] • Replicable Workflow w/ Micromamba Environments • GitHub
Causal Inference: Potential Outcomes Framework • Experiments (i.e., fully-, block-, and cluster-randomised designs w/ two arms [A/B Tests] or more) • Instrumental Variables • Diff-in-Differences • Regression Discontinuity • Synthetic Control
Regression: Ordinary-Least Squares • (Bayesian) Generalised Linear Models • Multilevel Models • Survival Analysis

EDUCATION

PhD: Social Research Methods (i.e., Applied Statistics) [Full Scholarship], **London School of Economics**, 2012 – 2016

MSc: Social Science of the Internet [Clarendon Fund Full Scholarship], **University of Oxford**, 2011 – 2012

BA: Communication Studies [Summa Cum Laude; GPA = 3.97/4.00], **Clemson University**, 2007 – 2011

EXPERIENCE

Data Scientist (Applied Econometrics and Experimentation) **Octopus Energy Group • July 2023 – Present**
Experiment to deliver scientific and business impact whilst preserving the customer experience and interfacing w/ major (non-)technical stakeholders (i.e., division CEO, div. COO, chief economist, external affairs director, commercial product owners, and product marketers).

- Lead 12-mth. project on the adoption and behavioural impact of a keystone product for British electric-vehicle drivers.
 - Designed 2024 field experiment (i.e., blocked-cluster randomised encouragement design) on 13,233 customers to obtain the causal effect of cash incentives on product adoption + the causal effect of incentivised product uptake on electricity use.
 - Managed £50K incentive budget with savvy trial design (e.g., wave-based enrolment, re-randomisation contingency plans).
 - Planned quasi-experiment (i.e., diff-in-differences for staggered treatments) using historical data on 101,328 customers with electric cars to obtain the causal effect of non-incentivised product adoption on electricity use throughout all of 2023.
- Lead 3-mth. project on atmospheric temperature and customer rejection of remote control of their home heating/cooling system.
 - Used data from 1,581 smart thermostats (1,000 American households) + Bayesian survival analysis to model the time until customers override remote control of their home climate system during 206K energy-savings events (1-150-minutes).
 - Munged high-resolution (hourly) data with shapfiles to create geospatial weather statistics for customers' residence areas.
 - Authored report advising U.S. product owner on remote-control strategies for maximising customer comfort given weather.
- Used data from 600K British customers and regression discontinuity + instrumental variables to obtain the causal impact of the timing (evening vs. morning) and type (email vs. SMS text) of digital requests to save energy during promotional events in 2023.
 - Performed two-stage least-squares estimation w/ Python (linearmodels) — munging data + output with numpy + pandas.
- **General Responsibilities:**
 - Determine best approach to solve problems — communicating clearly with technical and non-technical decision-makers.
 - Autonomously execute tasks with large elements of ambiguity (e.g., entire research reports, compelling visuals w/ ggplot2).
 - Simulate fake data w/ R (i.e., declaredesign + simstudy + dplyr) to plan and diagnose experiment designs (e.g., power, bias).
 - Write performant queries of DBT tables with Spark SQL + Databricks to pull/munge data for simulation studies and scientific analyses (i.e., filter, join, group, explode, and pivot 2D arrays w/ millions of rows such as tables of half-hourly electricity use).
 - Report directly to head of randomised trials to influence decisions on timelines, data, analyses, reporting, and management.

Fellow in Quantitative Research Methods (Academic Faculty) **London School of Economics • Jan 2022 – Dec 2022**

- Secured \$447,116 in project funding from the U.S. National Science Foundation by co-designing 3-year project on inter-group cooperation in Ethiopia — leading hypothesisation + design of inferential network analysis (400 adults) and project roadmap.
- Taught practical, 1-hour classes (3/week) for 60 students (MSc & PhD) on highly interpretable forms of supervised learning (i.e., linear, binary logistic, multinomial logistic, ordinal logistic, Poisson, and negative binomial regression) using R + Markdown.

British Academy Postdoctoral Research Fellow (Academic Faculty) **University of Oxford • Jan 2018 – Dec 2020**

- Drove 4 projects (3 solo) on the genesis of social relationships (friendship, altruism) in 165 villages (China, India, Nicaragua) end-to-end — moving from broad aims to testable hypotheses, analysing data with R, and writing all parts of published papers.
- Collaborated with colleagues to discuss, prioritise, and complete tasks related to teaching, research, ethics, and hiring.
- Co-led course on basic statistical inference (e.g., t-tests; p-values), delivering weekly lectures + practicals to 50 pupils (MSc & PhD).
- Evaluated work (i.e., exams, R code, theses) — tactfully giving feedback to students with high anxiety about maths and statistics.

Postdoctoral Researcher **University of Cambridge • Feb 2016 – Dec 2017**

- Secured £337,789 in funding from the British Academy (U.K.'s National Academy of Social Sciences) for a solely-designed project.