

**SARAH CATHERINE MCDONALD**  
[smcdonald@g.harvard.edu](mailto:smcdonald@g.harvard.edu) | [GitHub](#) | [LinkedIn](#)

## EDUCATION

- 2023 - PhD Student, Statistics, Harvard University
- 2021 - 2023 Predoctoral Research Fellow, Graduate School of Business, Stanford University
- 2016 - 2020 BSc, Mathematics and BA, Economics, Marymount University  
Honors Program, *Summa cum laude*
- 2017 Oxford Study Abroad Programme  
Participated in a guided reading tutorial with Dr. Gabrielle Watson (Faculty of Law) exploring contemporary perspectives of the United Kingdom penal system, including capital punishment, life imprisonment, and restorative justice.

## RESEARCH EXPERIENCE

- 2021- *Predoectional Research Fellow*, Graduate School of Business, Stanford University
- Modeling Network Interference with Real Data | Dr. Mohsen Bayati | Operation, Information, and Technology
- Conducted a literature review on penalized multidimensional autoregressive models, and the intersection of network interference and experimental design.
  - Used a variety of high-dimensional vector autoregression packages to model real data (such as daily covid cases) and simulated data on a high-performance computing (HPC) cluster. (R)
- An Experiment on Individual Customization of a Digital Advisory Service | Dr. Susan Athey | Economics
- Used Wasserstein Generative Adversarial Networks (GAN) to simulate a replication dataset using the joint and conditional relationships from our real dataset on an HPC cluster. (Python)
- Chatbots to Address Vaccine Hesitancy: Segmentation and Targeted Interventions | Dr. Susan Athey | Economics
- Cleaned and summarized experimental data, prepared chatbot attrition statistics, treatment effect analysis, and user segmentation, and applied the Romano-Wolf correction for multiple hypotheses testing. (R)
  - Conducted complex power calculations using simulated data, families of hypotheses, and multiple hypothesis correction.
- Facebook Misinformation | Dr. Susan Athey | Economics
- Cleaned and summarized experimental data, prepared chatbot attrition statistics, and treatment effect analysis. (R)
  - Conducted complex power calculations using simulated data, families of hypotheses, and multiple hypothesis correction. Visualized power for different sample sizes and treatment probabilities for multiple outcomes.
- [Do Jobseekers Value Diversity Information? Evidence from a Field Experiment](#) | Dr. Jung Ho Choi | Accounting
- Cleaned experimental data, assessed covariate balance, and analyzed data with ANOVA and probit models. (R)
- 2019-21 *Research Assistant, Fellow, and Intern*, Data Science for the Public Good Program  
Biocomplexity Institute & Initiative, Social & Decision Analytics Division, University of Virginia
- Developed a mathematical formalization and proofs of common knowledge on directed Facebook and Twitter-type communication networks.
  - Conducted network and node-level analysis of labor force workflows in Fairfax County, Virginia at the census tract level using LEHD Origin-Destination Employment Statistics data. (R)
  - Structured and scripted a *shiny* dashboard (<https://dspgtools.shinyapps.io/VDOE-App/>) with interactive visualizations for locating potentially vulnerable Virginia high school seniors impacted by COVID-19. (R)
  - Applied hierarchical clustering to job ad data to examine skills required for specific jobs; used sequence analysis to analyze resume data; explored issues with standard skilled technical workforce (STW) criteria; specified changes in STW designations from Occupational Information Network-Standard

**SARAH CATHERINE MCDONALD**  
[smcdonald@g.harvard.edu](mailto:smcdonald@g.harvard.edu) | [GitHub](#) | [LinkedIn](#)

Occupation Classification (O\*NET-SOC) 2010 to the new O\*NET-SOC 2019 taxonomy; identified and organized an inventory of federal, private, and not-for-profit data for describing the national STW on an interactive dashboard (<https://dspgtools.shinyapps.io/Measuring-STW-App/>); visualized educational pathways, career pathways, and skill supply and demand for STW in Virginia metropolitan areas. (R)

- Structured and scripted a *shiny* dashboard (<https://dspgtools.shinyapps.io/EM-Data-Infrastructure/>) with *leaflet* and *plotly* interactive visualizations that extends the Community Capitals Framework to provide data insights about community strengths at the county level in Virginia, Oregon, and Iowa.
- Investigated new techniques for visualizing National Center for Science and Engineering Statistics data with variability, specifically lollipop charts, slope graphs, and “statebins” charts. (R)
- Compared and mapped discrepancies between provider-reported Federal Communications Commission broadband coverage estimates with relevant household-reported American Community Survey and Microsoft Airband usage data at census tract, county, and state levels. (R)
- Led two undergraduate interns through a three-month project benchmarking Burning Glass Technologies data with the Job Openings and Labor Turnover Survey and Occupational Employment Statistics data; produced benchmarking *shiny* dashboard (<https://dspgtools.shinyapps.io/stw2/>). (R)
- Communicated with program sponsors through oral presentations, poster presentations, and written weekly updates; collaborated with multidisciplinary, vertically integrated teams of students and faculty.
- Used git and Github for version control, and R packages (e.g., tidyverse, data.table) to acquire, profile, clean, transform, analyze, and visualize data.

2019-20 *Honors Scholar*, Undergraduate Honors Thesis, Marymount University Honors Program

- Independently designed and executed a research project examining the effect of Catholic sex abuse scandals on Catholic higher education in the United States.
- Retrieved multiple datasets through web scraping; processed and merged complex data, including the Integrated Postsecondary Education Data System.
- Used multiple linear regression to explore the influence of negative publicity due to scandals in the Catholic Church on enrollment at Catholic colleges, universities, and seminaries. (R)

**PUBLICATIONS (Peer-reviewed)**

**McDonald, S.**, Korkmaz, G. (2021). Common Knowledge on Facebook Communication Networks: Models and Experimental Findings.” In Proceedings of the 12th Conference on Complex Networks (CompleNet 2021), Springer Proceedings in Complexity, [https://doi.org/10.1007/978-3-030-81854-8\\_3](https://doi.org/10.1007/978-3-030-81854-8_3).

**PUBLICATIONS (Technical Report)**

Lancaster V.A., **McDonald S.**, Montalvo, C., Siwe, L. (2021). Designating the Skilled Technical Workforce Using O\*NET-SOC (2019). Proceedings of the Biocomplexity Institute, Technical Report TR# 2021-087, <https://doi.org/10.18130/142t-4976>.

**PUBLICATIONS (Broad Audience)**

Lancaster, V., Cohen, S., Fowers, A., Isch, C., & **McDonald, S.** (2020). Using Resume Data to Identify Skilled Technical Candidates. SAGE MethodSpace, <https://www.methodspace.com/blog/using-resume-data-to-identify-skilled-technical-candidates>.  
Pristavec, T., Goldstein, J., Savchyn, K., **McDonald, S.**, & Sawhney, R. (2020). Broadband coverage in the US: When sources disagree. SAGE MethodSpace, <https://www.methodspace.com/blog/broadband-coverage-in-the-us-when-sources-disagree>.

## CONFERENCE PRESENTATIONS

- McDonald, S.,** Korkmaz, G. (2022). “Common Knowledge and Collective Action on Directed Communication Networks: Models and Experimental Findings.” Poster presented at the 2022 American Economic Association/Allied Social Science Associations Annual Meeting, Virtual.
- McDonald, S.,** Korkmaz, G. (2021). “Common Knowledge on Facebook Communication Networks: Models and Experimental Findings.” Presented at the 5th European Conference on Social Networks, Virtual.
- McDonald, S.,** Korkmaz, G. (2021). “Modeling Common Knowledge and Collective Action on Directed Communication Networks.” Presented at Networks 2021, Virtual.
- McDonald, S.,** Korkmaz, G. (2021). “Common Knowledge on Facebook Communication Networks: Models and Experimental Findings.” Presented at the International Conference on Complex Networks (CompleNet), Virtual.
- Ray, S., **McDonald, S.,** Hart, O., Pietrowicz, S., Pristavec, T., Kramer, B., Linehan, K., Tobin, J., Shipp, S., Keller, S. (2020). “Fairfax County Labor Markets: Characterizing Local Workforce and Employment Network.” Presented at the Knowledge Discovery and Data Mining Conference, Association for Computing Machinery, Virtual.
- Ray, S., **McDonald, S.,** Hart, O., Pietrowicz, S., Pristavec, T., Kramer, B., Linehan, K., Tobin, J., Shipp, S., Keller, S. (2020). “Fairfax County Labor Markets: Characterizing Local Workforce and Employment Network.” Presented at the Data Science for Public Good Symposium, University of Virginia Biocomplexity Institute and Initiative, Virtual.
- McDonald, S.** (2020). “The Effect of Catholic Sex Abuse Scandals on Catholic Higher Education in the United States.” Presentation accepted at Virginias Collegiate Honors Council 2020 Honors Conference. Canceled due to COVID-19.
- McDonald, S.** (2020). “The Effect of Catholic Sex Abuse Scandals on Catholic Higher Education in the United States.” Presented at Marymount University Student Research Conference, Virtual.
- McDonald, S.** (2020). “The Effect of Catholic Sex Abuse Scandals on Catholic Higher Education in the United States.” Poster accepted at National Conference on Undergraduate Research. Canceled due to COVID-19.
- Savchyn, K., **McDonald, S.,** Sawhney, R., Pristavec, T., Goldstein J., Shipp, S. (2019). “Broadband Data Validation: Comparing U.S. Broadband Coverage”. Poster presented at the GEOINTegration Summit, United States Geospatial Intelligence Foundation, Herndon, VA.
- Savchyn, K., **McDonald, S.,** Sawhney, R., Pristavec, T., Goldstein J., Shipp, S. (2019). “Broadband Data Validation: Comparing U.S. Broadband Coverage”. Poster presented at the Data Science for Public Good Symposium, University of Virginia Biocomplexity Institute and Initiative, Arlington, VA.
- McDonald, S.** (2018). “An Economic Inquiry into Voting Preferences in the 2012 and 2016 Presidential Elections.” Presented at Marymount University Student Research Conference, Arlington, VA.

## FUNDING

- 2023- *NSF Graduate Research Fellowship*, National Science Foundation
- 2022 *Selected and funded Fall Preview Weekend Attendee*, University of Michigan, Department of Statistics.
- 2021- *Selected and funded Predoctoral Research Fellow*, Stanford Graduate School of Business.
- 2020 *Selected and funded Fellow/Intern, Data Science for the Public Good Young Scholars Program*, University of Virginia Biocomplexity Institute.
- 2020 *Selected and funded Topology for Data Science (T4DS) attendee*, Montana State University Data Science. Canceled due to COVID-19.
- 2019 *Selected and funded Undergraduate Intern, Data Science for the Public Good Young Scholars Program*, University of Virginia Biocomplexity Institute.
- 2018-19 *Pablo Coto Memorial Scholarship*, Marymount University. Awarded for exemplifying academic excellence and leadership, \$1,350.
- 2016-20 *Honors Academic Scholarship*, Marymount University, \$16,000.

**SARAH CATHERINE MCDONALD**  
[smcdonald@g.harvard.edu](mailto:smcdonald@g.harvard.edu) | [GitHub](#) | [LinkedIn](#)

2016-20 *President's Academic Scholarship*, Marymount University, \$60,000.

### **HONORS & AWARDS**

- 2020 *Academic Excellence Award (Economics)*, Marymount University. Awarded for superior academic performance in major and contributions to the department through curricular and co-curricular activities.
- 2020 *Academic Excellence Award (Mathematics)*, Marymount University. Awarded for superior academic performance in major and contributions to the department through curricular and co-curricular activities.
- 2020 *Senior Leadership Award*, Marymount University. One of ten awardees for outstanding contributions the Marymount Community and the greater community as student leaders.
- 2020 *Semi-Finalist*, Fulbright Budapest Semesters In Mathematics-Rényi Institute.
- 2019 *Senior Resident Assistant of the Year*, Marymount University.
- 2019 *Finalist, Behavior-Centered Design Challenge*, Rare Center for Behavior and the Environment. Worked in a multidisciplinary team to create behavioral ridesharing incentives to reduce greenhouse gas emissions for INOVA, one of Northern Virginia's largest employers.
- 2018 *Second Place Undergraduate Presentation*, Marymount University Student Research Conference.

### **SKILLS**

*Statistical programming environments:* R, RStudio: tidyverse, data.table, glm, igraph, stringr, plotly, leaflet, sf, shiny, knitr, rvest, statebins

*Other:* SQL, Stata, Java, HTML, CSS, LaTeX, GitHub, git

### **UNIVERSITY SERVICE**

- 2021 *Alumni Panelist*, Marymount University Mathematics Department Seminar on Post-Graduation Opportunities in Industry and Academia.
- 2019-20 *President*, Student Government Association, Marymount University. Worked with students and administrators to develop programs including a no-questions-asked, discreet, free meal voucher program; a grant for students studying abroad who had financial need or it was their first time abroad; the university's first student emergency fund which made an impact during the onset of COVID-19; and free menstrual products in bathrooms.
- 2018-19 *Peer Tutor*, Center for Teaching and Learning, Marymount University. Conducted one-on-one tutoring sessions for college students in mathematics courses at all levels, from remedial mathematics to graduate-level courses.

### **PROFESSIONAL MEMBERSHIPS**

- 2021-22 *Member*, American Economic Association.
- 2019-20 *Member*, Association for Women in Mathematics, Marymount University.
- 2018- *Member*, American Statistical Association. *President and Chapter Founder*, Marymount University.
- 2018-20 *Member*, Kappa Mu Epsilon, Marymount University.