

# MaryClare Roche

POLITICAL SCIENTIST · DATA SCIENTIST

✉ macroche@princeton.edu | 🏠 maryclareroche.com | 📷 MCRoche | 📄 maryclare-roche

## Skills

---

### Programming

R: tidyverse, topicmodels, tm, quanteda, zoo, arabicStemR, ggplot2;

python: numpy, pandas, scipy, spaCy, gensim, matplotlib, multithreading, multiprocessing, selenium;

### Software

Matlab, Mathematica, ArcGIS, Docker, Stata, SPSS, LaTeX, Excel

### Technical

Econometrics (Advanced), Time Series Analysis, Structural Estimation, Causal Inference, Machine Learning, Imputation, Web Scraping, Natural Language Processing (NLP) in English and Arabic

### Languages

Spanish (Proficient), French (elementary), Arabic (Modern Standard, elementary)

## Work Experience

---

### Arab Barometer, Princeton University

Washington, D.C.

SENIOR RESEARCH SPECIALIST

2020-present

- Oversaw and aided in the development of a custom data verification software used in real time to evaluate the quality of incoming surveys
- Built several custom graphing and statistical analysis R packages exclusively for Arab Barometer that work with both English and Arabic
- Developed an original SMS messaging program in python to routinely deliver surveys across the Middle East and North Africa
- Researched and evaluated the effectiveness of various SMS and survey settings to ensure the optimal response rate
- Built topic models and performed sentiment analysis with machine learning for Arabic survey responses
- Designed and oversaw data processing pipeline for incoming data from 10 countries across three platforms with the use of Stata, R, and SPSS
- Created econometric models and data visualization to be presented at the US State Department, several US Embassies, the United Nations, and various international non-profits.
- Trained colleagues on the use of R, git/GitHub, ODK
- Oversaw the adaptation of any and all new technologies used by Arab Barometer

### University of Rochester

Rochester, New York

THESIS WORK

2013-2020

- Scraped original data via multi-threading over a generated cluster of temporary, localized IP addresses with Microsoft Azure and Docker
- Created topic models using both LDA and GloVe on Arabic text in both R and python
- Utilized Optical Character Recognition (OCR) API on official government documents in Arabic to supplement missing information
- Built an Arabic-to-English dictionary using Google and Bing translator APIs
- Analyzed Irregular Time Series data to uncover temporal speaking patterns
- Analyzed internal dynamics of pro- and anti-regime parties in authoritarian regimes
- Uncovered unexpected gender dynamics in Middle Eastern semi-democratic parliaments
- Created game theoretic models to explain the conditions that allow democratic institutions in authoritarian regimes to simultaneously create coordination problems for regime opposition and solve them for regime allies

TEACHING ASSISTANT

- Taught R for statistical analysis and data presentation
- Lead recitations and mentored students at both the undergraduate and graduate level; classes included Data Analysis, Game Theory, Introduction to Comparative Politics, and Political Philosophy

### International Monetary Fund

Washington, D.C.

RESEARCH ASSISTANT

Fall 2015

- Created and solved a game theoretic model describing the relation between government regulation and banks' investment portfolios
- Taught advanced game theoretic concepts

## Education

---

### University of Rochester

Rochester, NY

PH.D., M.A. IN POLITICAL SCIENCE

2013 - 2020

### University of Maryland, College Park

College Park, MD

B.S. IN PHYSICS (PRIMARY) AND GOVERNMENT & POLITICS (SECONDARY)

2009 - 2013