Karim Naguib

Economist/Data Scientist | Bayesian Methods & Causal Inference

karimn2.0@pm.me karimn.co github.com/karimn

Summary

Economist and applied Bayesian statistician with 10+ years building **hierarchical models** for **longitudinal and survival data**. Most recently at AstraZeneca, building **patient-level digital-twin models** of oncology tumor dynamics with real-world evidence integration; earlier work spans field experiments in East Africa and South Asia, real-estate market design, and health-behavior interventions.

Skills

Methods: Bayesian hierarchical modeling, state-space and longitudinal models, multistate survival and competing risks, joint longitudinal–survival ("digital twin") models, cross-validation and predictive assessment (LFO-CV, PSIS-LOO), MCMC diagnostics and prior sensitivity, causal inference and randomized experiment design, Gaussian processes

Languages: R (tidyverse), Python, Stan, Julia, SQL, C/C++

Agentic Workflows: Built and maintain an agentic-workflow layer on Claude Code; at AstraZeneca it was shared across the modeling team. Components include collaborative mathematical derivation, Stan diagnostic sub-agents, `targets` pipeline explorer, MCP-integrated skills for Domino, and reusable plugin templates codifying team Bayesian workflow conventions.

Infrastructure: CmdStan / cmdstanr, `targets` pipelines, Domino, Docker, AWS, Databricks, Snowflake, PostgreSQL, git, SLURM

Education

Ph.D. in Economics | *Boston University* Boston, MA | 2014
Fields: development, health, and experimental economics

B.S. in Computer Science, Minor: Mathematics | *The American University in Cairo* Cairo, Egypt | 1999

Experience

Independent Researcher / Consultant | *Self-employed* Jul 2026 – present

- Independent Bayesian modeling and causal-inference work; continued development of agentic tooling for statistical modeling

Senior Data Scientist | *AstraZeneca* Aug 2023 – Jul 2026

- Bayesian digital-twin modeling of oncology tumor dynamics to forecast Phase 3 outcomes from Phase 2 trials, historical data, and real-world evidence, supporting Phase 3 initiation decisions
- Hierarchical state-space tumor-dynamics models and multistate survival models, jointly fit for competing-risk decomposition of progression-free survival; validated with leave-future-out cross-validation
- Modular Stan architecture and `targets`-based pipeline for reproducible trial analyses
- Agentic-workflow layer for collaborative mathematical derivation and MCMC diagnostics, adopted across the modeling team

Independent Researcher | *Self-employed* Nov 2022 – Aug 2023

- Built a Julia package for a partially observable MDP solution using Monte Carlo tree search to optimize intervention implementation and evaluation decisions
- Bayesian hierarchical models for future-state prediction inside the POMDP; particle filters for online planning

Senior Data Scientist | *Opendoor* May 2021 – May 2022

- Designed randomized experiments to estimate demand and supply elasticities for pricing policy across regional real-estate markets
- Hierarchical Bayesian and Gaussian-process models pooling across markets while preserving local heterogeneity

Independent Researcher / Consultant | *Self-employed* May 2019 – May 2021

- Structural Bayesian model of social-experiment outcomes using data from a large-scale RCT in Kenya; hierarchical intervention-effect models
- Hierarchical model of COVID-19 transmission heterogeneity; Gaussian processes for spatial correlation in intervention effects
- R package for bounded identification of counterfactuals in DAGs under imperfect compliance

Economist | *Evidence Action* May 2014 – May 2019

- Designed and analyzed large-scale randomized evaluations in Kenya and Bangladesh for flagship programs: deworming across East Africa and the at-scale replication of the No Lean Season seasonal-migration subsidy
- Contributed to the replication analysis of No Lean Season; the failure to replicate the pilot effect informed Evidence Action's decision to close the program
- Study design, power analysis, field data-collection supervision, and Bayesian and frequentist analysis for operational and policy audiences

Software Design Engineer in Test | *Microsoft*, Windows Debugging Tools Apr 1999 – Mar 2007

- Test framework and automation for Windows debugging tools; C/C++, systems-level testing

Publications

Barker, Davis, López-Peña, Mitchell, Mobarak, Naguib, Reimão, Shenoy, and Vernot, "Migration and Resilience during a Global Crisis," *European Economic Review* (accepted); two additional working papers on migration at scale (Bangladesh) and optimal incentives under social norms (Kenya); Naguib et al., "PIONEER: Bayesian Joint Modelling of Mechanistic Tumour Growth and Time-to-Event Endpoints for Dynamic Prediction of Ongoing Oncology Trials" (arXiv:2607.17908).