HASAN TOSUN

1257 Gibbs Avenue, Falcon Heights, MN 55108 • hktosun@gmail.com • (612) 458-4140 • hasantosun.ca

Data-oriented economist with a strong background in statistics, causal inference, industrial organization, and machine learning. Experienced in handling larger than memory datasets. Worked within team environments, fulfilling data-related tasks, including creating ETL pipelines and conducting econometric analyses. Collaborated with people from different organizations and fields. Skilled at communicating technical concepts to non-technical people, with over three years of award-winning teaching experience.

PROJECTS

Using Machine Learning to Predict Wages *(with Lejvi Dalani)*

- Goal: Predicting individual wage earnings using survey data.
  - Used Current Population Survey dataset from 2012.
- Achieved 34% improvement compared to the baseline model used in the economics literature.
- Data Preprocessing:
  - Used LASSO for feature selection. Tuned penalty parameter using cross-validation.
  - Used Entity Embedding for dimension reduction and transformed a categorical variable with many distinct values to a small number of continuous variables.
  - Created new variables that contain information on family members using individual IDs and family interrelations.
  - Imputed missing values using KNN.
- Prediction:
  - Trained tree-based models (Random Forest, XGBoost and LightGBM).
  - Used Randomized Search Cross Validation to tune hyperparameters.
  - Built a stacking model that combines predictions from multiple models.
  - Trained a CatBoost model using data and the predictions from the models above.

Measuring the Impact of Lyft's Acquisition of a Bike Sharing Company on the Competition in Ride Sharing *(with Suleyman Gozen)*

- Based on a dataset with ~1 billion observations.
  - Raw data was split into dozens of flat files with different schemas.
  - Used Spark on R to handle the data.
- Summary:
  - It focuses on the role of synergies generated by M&A transactions.
  - Used Lyft's acquisition of Motivate, the biggest bike sharing company in the US at the time, to measure the impact of synergies on Lyft and Uber's ridership.
  - Estimated the causal effect of the acquisition on Lyft's ridership by using difference-in-differences-in-differences model, and the trip-level ride sharing data from New York.
- Finding: The causal impact of the acquisition is a 6% increase in Lyft ridership.
- Paper presented at the International Industrial Organization Conference 2021, one of the top conferences in the field.

ETL Pipeline and R Package Development for Minnesota Child Care Research Team **Jan 2020 – present**

- Designed and implemented an ETL process to import, clean and combine hundreds of flat data files from different sources, with a group of PhD students.
- Wrote an R package *(minnccaccess)* to improve collaboration within the team.
- Created a Shiny app to calculate and visualize key policy metrics.
- Documented the data wrangling processes to ensure reproducibility.
- Worked with a group of research scientists from a child care research institution in the writing of several policy reports.

Estimating the Causal Impact of Public Pre-Kindergarten on the Private Child Care Sector

- Gathered data from various sources; cleaned and merged the datasets.
- Used a difference-in-differences model to estimate the effect of the expansion in public pre-kindergarten programs on private child care supply.
- Found that the expansion caused more than 10% decline in the number of family-based child care providers.

EDUCATION

PhD in Economics *University of Minnesota* **2022 (expected)**

Econometrics, Industrial Organization, Machine Learning, Statistics, Causal Inference, Difference-in-Differences

MA in Economics *Bogazici University, Istanbul, Turkey* **2016**

- Ranked 50th out of more than 250,000 students in National Graduate Record Examination.
- Best performing student in the cohort.

BSc in Industrial Engineering *Bogazici University, Istanbul, Turkey* **2011**

- Ranked in top 150 out of more than 1.6 million students in Turkish University Entrance Examination.

PROGRAMMING SKILLS

R *(tidyverse, sf, devtools, shiny, sparklyr)***, Python** *(numpy, pandas, scikit-learn)***, SQL, Spark**