Predicting salaries in the GTA

A Bayesian approach

Andrés Castillo

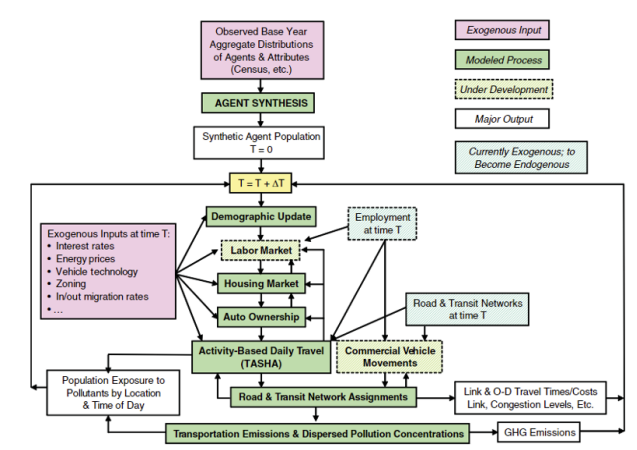
Supervisor: Prof. Eric Miller
Department of Civil and Mineral Engineering
University of Toronto
2023

Outline

1. Research motivation
2. Labour Market in ILUTE
3. Predicting salaries in ILUTE
4. Conclusions and future work

1. Research motivation

- Transportation models have grown in complexity, but some inputs are still considered **exogenous**.
- Work (HBW) is the **second most frequent** trip purpose in the GTA (TTS, 2016).
- Place of residence, place of work, household income, and auto ownership are **directly or indirectly related** to the labour market.

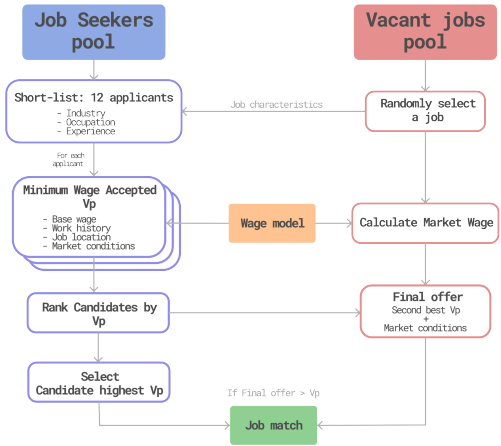


ILUTE flowchart (Miller et al, 2021).

2. Labour Market in ILUTE

Job matching process

- Wages **facilitate the interaction** between agents in the labour market
- Wages **allocate labour** to the most efficient use
  - Industries
  - Occupations
  - Regions

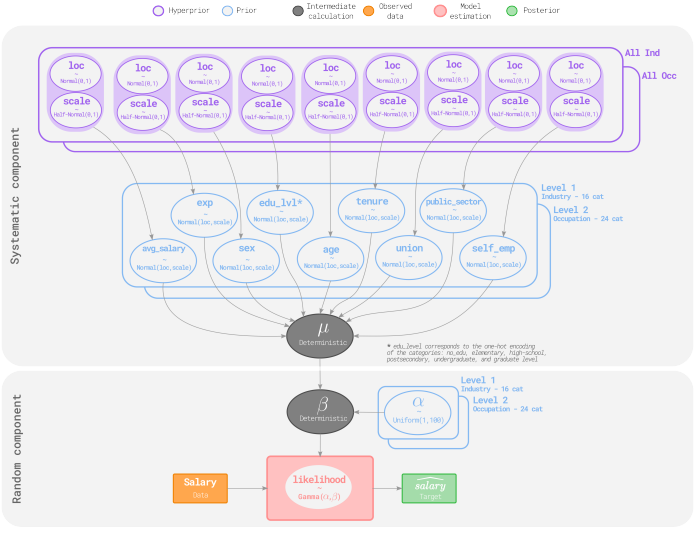


Job search and matching process in ILUTE (Adapted from Harmon, 2013).

3. Predicting salaries in ILUTE

Model structure

- Labour markets are organized in a **hierarchical** structure
  - Industries
  - Occupations
- The proposed model uses a **Hierarchical Gamma GLM** and **Bayesian inference**

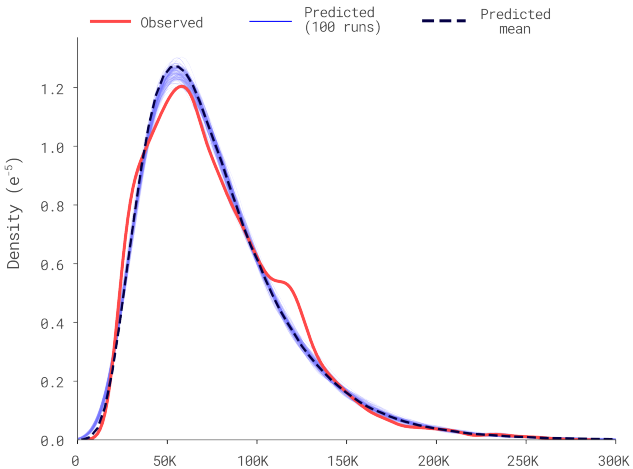


Proposed salary model

3. Predicting salaries in ILUTE

Results - Aggregated level (GTA)

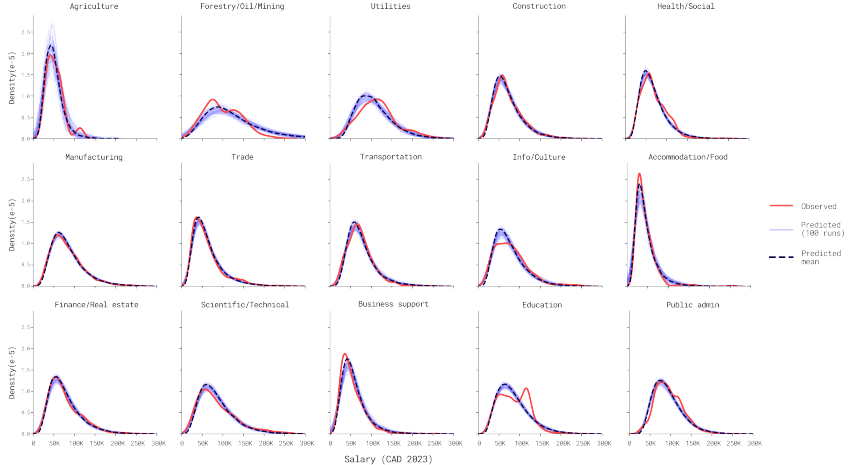
- Annual salaries are **positive** and **right skewed**.
- The proposed model resembles the observed distribution by using the **Gamma random component**.



Observed and predicted annual salary distribution for the GTA

3. Predicting salaries in ILUTE

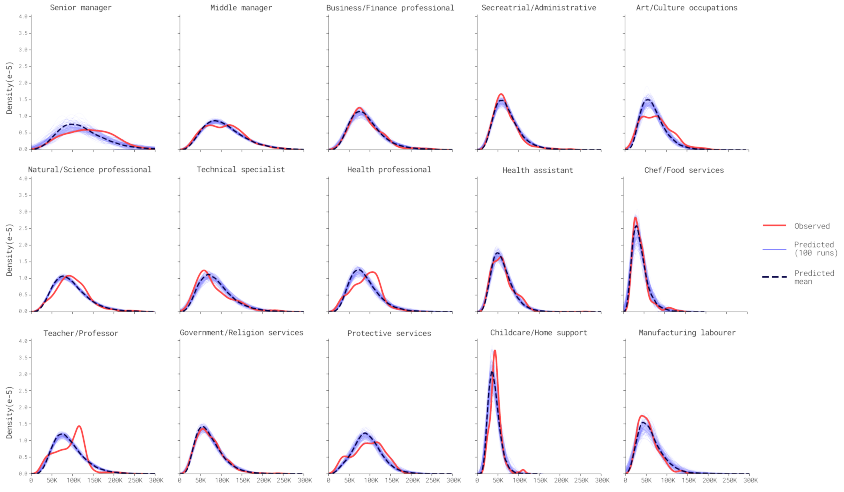
Results - Aggregated level (Industry)



Observed and predicted annual salary distribution by industry for the GTA

3. Predicting salaries in ILUTE

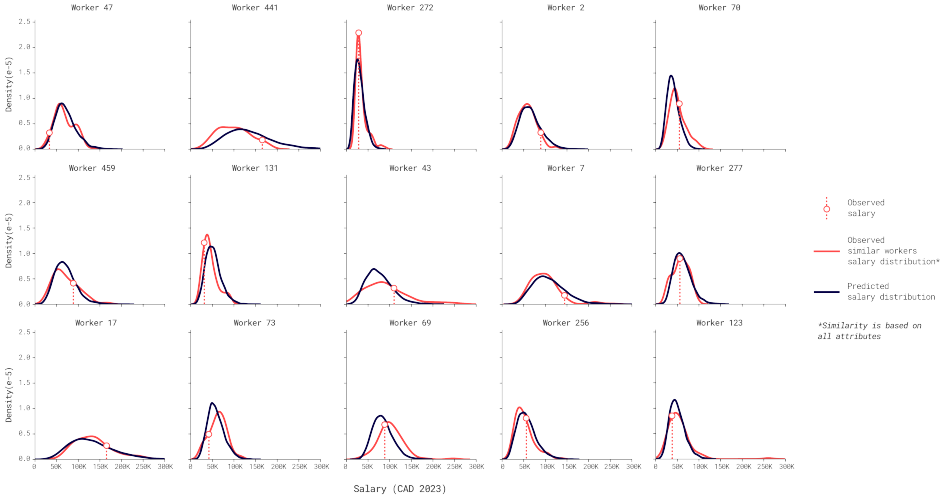
Results - Aggregated level (Occupation)



Observed and predicted annual salary distribution by occupation for the GTA

3. Predicting salaries in ILUTE

Results - Disaggregated level (Worker)



Observed and predicted salaries for randomly selected workers

4. Conclusions

- Considering the hierarchical structure of the labour market has several advantages in the proposed model:
  - Improves the **prediction accuracy** of salaries by sharing information across industries and occupations.
  - Enhances the **robustness** by reducing the effect of outliers *(Shrinkage effect)*.
  - Both **information sharing** and **shrinkage effect** reduces the risk of **overfitting**.
- The **Gamma random component** allows the model to capture the **right skewness** of the observed salary distribution.
- Replacing point estimates with **probability distributions** might improve the interactions within the ILUTE labour market module.

Questions?

Thank you for your attention!