

MASc Thesis

andres.castillo

December 2023

Abstract

Acknowledgements

Contents

1	Introduction	5
1.1	Outline	6
2	Literature review	7
3	Data sources	8
3.1	Survey of Labour and Income Dynamics - SLID	8
3.2	Hierarchical structure: Industry and Occupation	8
3.3	Exploratory Data Analysis	8
4	Salary model estimation	9
4.1	Data structure: from single to multilevel structure	9
4.2	Model variables	9
4.3	Model specification	9
4.4	Model interpretability	9
4.5	Sensitivity analysis	9
4.6	Longitudinal analysis of the estimated parameters	9
4.7	Model validation	9
5	Model integration	10
5.1	Integration with the existing implementation	10
5.2	Model comparison: proposed vs. existent	10
6	Computational cost of Bayesian inference	11
7	Conclusions	12
8	Future work	13

Chapter 1

Introduction

Transportation models have grown in complexity by including more details on travel behaviour. However, some demographic and socioeconomic variables are still considered exogenous factors, introducing uncertainty and reducing the model's effectiveness in forecasting future scenarios. Since work trips are the most frequent trips in urban areas, understanding work-related attributes is relevant for travel demand modelling. Some attributes such as the place of residence, place of work, household income, auto ownership, and mode choice can directly or indirectly relate to the labour market.

According to Harmon and Miller (2020), the inclusion of labour market interactions within urban simulations has had little development despite the critical relationship between work and transportation systems in urban areas. In their paper, Harmon and Miller (2020) proposed an Agent-Based framework for simulating the demand and supply interactions of the labour market. This framework provided the first approach to a fully endogenous labour market simulation within a transportation-related model. Nevertheless, as discussed by Harmon (2013), much research still needs to be done to ensure the validity of results as the model evolves in time, to understand the interactions between agents, and to investigate the factors influencing the recruitment process within firms. In particular, He provides some evidence that the existent wage model could be underestimating salaries in the long run, which can be critical given the role of wages in the labour market.

Therefore, this thesis presents a model that predicts salary based on individual attributes of a worker using the principles of Bayesian inference. This model improves the prediction accuracy by accounting for the hierarchical structure of the data, which better simulates the variability in the salaries at both an aggregated and disaggregated level.

1.1 Outline

Although all sections in this document are structured sequentially, some can be optional according to the reader's knowledge of Bayesian inference. After this introduction, Chapter 2 presents an overview of economic theory and discusses the role of salaries and wages in the labour market interactions. Chapter 3 briefly introduces Bayesian inference, the framework for estimating the proposed model. Chapter 4 discusses the data sources and the hierarchy of labour data. Additionally, it presents the main results of the exploratory data analysis that guides the model specification. Then, Chapter 5 presents the details of the proposed salary model and the validation results with new data, followed by Chapter 6, which discusses the integration of this proposed model into the existing ILUTE framework. Finally, Chapter 7 compiles the principal results, and Chapter 8 discusses the future work.

Chapter 2

Literature review

This chapter provides a systemic perspective on the labour markets supported by the literature. It starts with a brief discussion of the economic theory that governs labour markets, followed by an overview of different implementations of microsimulation models with an emphasis on the ILUTE framework. Finally, this chapter closes with a discussion of the gaps that guide the work in the subsequent chapters. Although the economic literature on this subject is extensive, this section is mainly based on the concepts and ideas stated by

Chapter 3

Data sources

- 3.1 Survey of Labour and Income Dynamics
- SLID
- 3.2 Hierarchical structure: Industry and Occupation
- 3.3 Exploratory Data Analysis

Chapter 4

Salary model estimation

- 4.1 Data structure: from single to multilevel structure
- 4.2 Model variables
- 4.3 Model specification
- 4.4 Model interpretability
- 4.5 Sensitivity analysis
- 4.6 Longitudinal analysis of the estimated parameters
- 4.7 Model validation

Chapter 5

Model integration

5.1 Integration with the existing implementation

5.2 Model comparison: proposed vs. existent

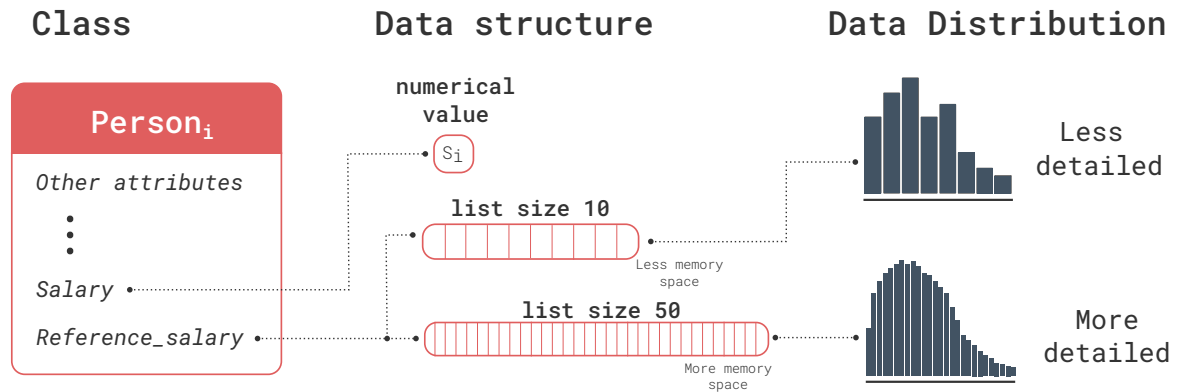


Figure 5.1: Representation of salary within the existent framework.

Chapter 6

Computational cost of Bayesian inference

Chapter 7

Conclusions

Chapter 8

Future work

Bibliography

- Harmon, A. (2013). A microsimulated industrial and occupation-based labour market model for use in the integrated land use, transportation, environment (ilute) modelling system.
- Harmon, A. and Miller, E. J. (2020). Microsimulating labour market job-worker matching. *Journal of Ambient Intelligence and Humanized Computing*, 11:993–1006.