8-28-2014 12:00 AM

Role of Search, Human Capital and Learning in Occupational Mobility and Immigrant Assimilation

Masashi Miyairi, *The University of Western Ontario*

Supervisor: Audra J. Bowlus, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Economics

ROLE OF SEARCH, HUMAN CAPITAL AND LEARNING IN OCCUPATIONAL MOBILITY AND IMMIGRANT ASSIMILATION
(Thesis format: Integrated Article)

by

Masashi Miyairi

Graduate Program in Economics

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

Abstract

This thesis contains three studies of job and occupational mobility, and their implications for earnings. The second chapter of the thesis develops and estimates a model of job and occupational search to examine how and how much learning influences young workers' job search and transition patterns. The model incorporates uncertainty regarding the accumulation processes of workers' different skills, and features directed search whereby workers choose search effort intensities for different occupations. The model is estimated using U.S. data, with individuals' occupational affiliations grouped into skilled white-collar, skilled blue-collar, and non-skilled occupations. The estimates show large differences in search frictions, skill acquisition rates, and learning opportunities across occupations. Simulation exercises show that learning can have a sizeable effect on young workers' job search. However, because of job search frictions, changes in job search effort due to learning do not result in a comparable effect in occupational transition outcomes. Search frictions have a particularly large consequence for those directing their search effort to the white-collar occupation.

Building on the search and matching model of Albrecht and Vroman (2002), the third chapter develops a dynamic model of employment transitions among full-time work, part-time work and nonemployment, and offers an explanation based on human capital depreciation for British women's life-cycle employment transition patterns. Numerical examples of the model indicate that the model can capture their stylized life-cycle transition patterns through their endogenous decision making under reasonable parameter values.

The fourth chapter develops and estimates an equilibrium search model of immigrants operating in the same labour market as natives, where newly arrived immigrants have lower job offer arrival rates than natives but can acquire the same arrival rates according to a stochastic process. Using Canadian panel data, substantial differences in job offer arrival and destruction rates are found between natives and immigrants that are able to account for three quarters of the observed earnings gap. The estimates imply that immigrants take, on average, 13 years to acquire the native search parameters. Counterfactual exercises show that the vast majority of earnings growth immigrants experience after migration is due to the job search assimilation process.

Keywords: Job search, human capital, learning, occupation, part-time employment, immigrant assimilation

Co-Authorship Statement

This thesis contains material co-authored with Audra Bowlus and Chris Robinson. All the authors are equally responsible for the work which appears in Chapter 4 of this thesis.

Acknowlegements

I would like to express my sincere gratitude to Audra Bowlus for her tireless guidance and generous support. I also would like to thank Chris Robinson and Lance Lochner very much for their advice and comments.

Warm thanks to the faculty and staff members of economics department at Western for their help, to friends I shared good times with, and to my family.

Parts of this thesis were made possible with facilities provided at the Statistic Canada Research Data Centre at the University of Western Ontario, and the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET), as well as Chris Robinson's permission to use his data on the skill/task factor scores of the 1970 Census 3-digit occupation codes.

Masashi Miyairi

London, Canada
July, 2014

To my parents.

Contents

Abstract i

Co-Authorship Statement ii

Acknowlegements iii

Table of Contents v

List of Figures viii

List of Tables ix

List of Appendices xi

1 Introduction 1
- Bibliography . . . 5

2 Job and Occupational Search and Learning of Comparative Advantages 7
- 2.1 Introduction . . . 7
- 2.2 The Model . . . 12
  - 2.2.1 Setup . . . 12
  - 2.2.2 Job Search . . . 13
  - 2.2.3 Skill Acquisition Process . . . 14
  - 2.2.4 Learning . . . 16
  - 2.2.5 Timing within a Period . . . 17
  - 2.2.6 Dynamic Programming Formulation of Worker's Problem . . . 18
  - 2.2.7 Analysis of Worker's Problem . . . 22
  - 2.2.8 Numerical Issues . . . 23
- 2.3 Estimation and Identification Strategy . . . 24
  - 2.3.1 Estimation Procedure . . . 24
  - 2.3.2 Identification Discussion . . . 27
- 2.4 Data . . . 28
  - 2.4.1 Job History . . . 29
  - 2.4.2 Occupation Classification . . . 31

2.4.3 Analysis of Estimation Data 34
2.5 Estimation Results 39
2.5.1 Parameter Estimates 39
2.5.2 Model Fit 43
2.5.3 Augmented Model 46
2.6 The Effect of Learning on Job Search Effort Decisions and Occupational Affiliation Outcomes 53
2.6.1 The Effect of Learning on Job Search Effort Decisions 54
2.6.2 The Effect of Learning on Transition Outcomes 58
2.7 Concluding Remarks 63
Bibliography 63

3 Motherhood, Part-time Work, and Skill Dynamics of British Women 67
3.1 Introduction 67
3.2 Model 70
3.3 Individual Worker's Problem 73
3.4 Numerical Illustrations 80
3.5 Concluding Remarks 90
Bibliography 91

4 Immigrant Job Search Assimilation in Canada 94
4.1 Introduction 94
4.2 Equilibrium Search Model 99
4.2.1 Environment 99
4.2.2 Workers' Problem 100
4.2.3 Firms' Problem and Equilibrium Wage Offer Distribution 104
4.2.4 Implications for Native-Immigrant Wage Gap 107
4.3 Estimation 109
4.3.1 Estimation Procedure 109
4.3.2 Identification 111
4.3.3 Data 111
4.4 Estimation Results 116
4.4.1 Parameter Estimates 116
4.4.2 Model Fit 121
4.4.3 Implications for Earnings Assimilation 123
4.5 Concluding Remarks 127
Bibliography 129

5 Conclusion 133

A A Case Allowing for Correlation between the Skill Acquisition Probabilities and the Initial Skill Levels 135
Bibliography 138

B Occupation Classification **140**
Bibliography 143

C Supplementary Results on Predicted Effects of Learning on Search Effort **145**

D Mathematical Proofs for Chapter 4 **150**

E Likelihood Contributions for Chapter 4 **164**

Curriculum Vitae **171**

List of Figures

2.1 Fraction of Individuals Working in Given Occupations 37
2.2 Observed Log Wage 38
2.3 Observed and Predicted Occupation Affiliations 46
2.4 Observed and Predicted Log Wages 46
2.5 Observed and Predicted Occupational Choice – Augmented Model 50
2.6 Predicted Occupational Choice by Diiferent Utility Type 52
2.7 Predicted Log Wage 53
2.8 Joint Probability Density of θ_1 and θ_2 54
2.9 Predicted Search Effort Allocations in the 41st Quarter 57
2.10 Predicted Occupational Affiliations in the 41st Quarter 60
2.11 Predicted Occupational Affiliations in the 41st Quarter with the Faster Learning Process 62

4.1 Observed and Predicted Earnings Distributions for Natives and Immigrants 123
4.2 Predicted Immigrants' Earnings Growth 124
4.3 Predicted Immigrant Earnings Relative to Native Earnings 125
4.4 Counterfactual Immigrant Earnings Relative to Native Earnings 126
4.5 Predicted Native Earnings Distributions 127

B.1 Factor Scores of the 3-digit Occupations by Aggregated Occupation Group . 142

C.1 Predicted Search Effort Allocations in the 41st Quarter – Type A Workers . 147
C.2 Predicted Search Effort Allocations in the 41st Quarter – Type B Workers . 148
C.3 Predicted Search Effort Allocations in the 41st Quarter – Type B Workers (Continued) 149

List of Tables

2.1 Summary of Aggregated Occupation Groups 33
2.2 Average Factor Scores of Aggregated Occupation Groups 33
2.3 Job Spells by Aggregated Occupation 34
2.4 Probability of Transitions from Jobs per Quarter 35
2.5 Transition Profiles between Occupations 36
2.6 Parameter Estimates 40
2.7 Predicted Probability of Transitions from Jobs per Quarter 43
2.8 Predicted Quarterly Occupational Transition Matrices 44
2.9 Parameter Estimates – Augmented Model 49
2.10 Predicted Probability of Transitions from Jobs per Quarter 50
2.11 Predicted Quarterly Occupational Transition Matrices 51
2.12 Predicted Job Search Effort Allocations and Distributions of Job Offered in the 2nd Quarter 55
2.13 Predicted Changes in Job Search Effort Allocations and Distribution of Offered Jobs in the 3rd Quarter 56
2.14 Predited Occupational Choice Distributions in the 41st Period 59

3.1 Baseline Parameter Values 80
3.2 Cutoff Values - High-Skilled Workers 83
3.3 Cutoff Values - Low-Skilled Workers 84
3.4 Ranking of Value Function in the Benchmark Parameter Configuration . . . 85

4.1 Summary Statistics from the Estimation Sample 116
4.2 Parameter Estimates: Event Arrival Rates 117
4.3 Parameter Estimates: Wage Cuts 118
4.4 Parameter Estimates: Firm Productivity Levels and Distribution 119
4.5 Predicted Moments from the Estimation Result with Separate Job Destruction Rates for Natives and Immigrants 122

A.1 Parameter Estimates - Skill Acquisition, Initial Skill Level 139

B.1 Parameter Estimates from Cluster Analysis 141
B.2 List of the 1970 Census Occupation Codes in Each Aggregated Occupation Group 144

C.1 Predicted Job Search Effort Allocations in the 2nd Quarter – Type A Workers 145
C.2 Predicted Job Search Effort Allocations in the 2nd Quarter – Type B Workers 146

List of Appendices

Appendix A A Case Allowing for Correlation between the Skill Acquisition Probabilities and the Initial Skill Levels 135
Appendix B Occupation Classification 140
Appendix C Supplementary Results on Predicted Effects of Learning on Search Effort 145
Appendix D Mathematical Proofs for Chapter 4 150
Appendix E Likelihood Contributions for Chapter 4 164

Chapter 1

Introduction

Various movements that individuals make in labour markets have been major subjects of labour economics. Examples include transitions between unemployment and employment or between jobs (Mortensen, 1986; Topel and Ward, 1992), occupational changes (Neal, 1999; Kambourov and Manovskii, 2008), and migration between regions or between countries (Greenwood, 1997; Borjas, 1999). As well as documenting and analyzing mobility patterns of different groups of individuals, the literature has been investigating their implications for a variety of issues ranging from life-cycle wage growth, unemployment, to inequalities. Aiming to contribute to this literature, my thesis consists of three studies of job and occupational transition patterns, the potential processes behind them, and their implications for workers' earnings.

Young workers change jobs and occupations more frequently than more experienced workers, with their mobility declining as they age. The second chapter aims to measure the extent to which learning behaviour accounts for their mobility patterns. With a view that human capital acquisition, search frictions, and learning all influence workers' job transition patterns, it develops a dynamic Roy (1951) model of comparative advantages incorporating these three processes. Specifically, the model incorporates uncertainty regarding the accumulation processes of workers' two skills, and features directed search whereby workers choose job search effort intensities for different occupations. Not fully informed of their own skill accumulation processes upon their labour market entry, workers make job search decisions relying on their beliefs about them, and resolve this uncertainty by observing their skill acquisition outcomes over time.

The model is estimated by simulated maximum likelihood using job history data of white male high school graduates from the 1979 cohort of the U.S. National Longitudinal Survey of Youth (NLSY79), with individuals' occupational affiliations grouped into skilled white-collar, skilled blue-collar, and non-skilled occupations. The main findings from the estimation results are as follows. First, the estimates reveal large differences among occupations in search frictions. There are substantial search frictions hindering transitions to the white-collar occupation while the non-skilled occupation is associated with the least search frictions.

Second, the estimated parameters governing skill growth suggest that the two skills in the model are primarily related to the white-collar and blue-collar occupations, respectively. Specifically, one of them is enhanced mostly in the white-collar occupation, while the other grows the fastest with work experience in the blue-collar occupation. The non-skilled occupation provides a limited skill growth environment.

Third, simulation exercises show that learning can have a sizeable effect on job search effort over time. Workers who learn that they have high skill acquisition probabilities for the white-collar related skill choose to focus more on the white-collar occupation, and those who learn the opposite instead choose to allocate less job search effort on this occupation. A signal regarding the blue-collar skill, however, has only a small effect on workers' job search effort choice.

Fourth, because of job search frictions, changes in job search effort due to learning do not result in a comparable effect in occupational transition outcomes. Job search frictions have a particularly large consequence for those who direct their job search effort to the white-collar occupation because this occupation presents high search frictions. Moreover, because signals about the white-collar skill are a predominant factor driving workers' learning process but hard to come by without employment in the white-collar occupation, the high search frictions associated with the white-collar occupation slow down the learning process, reducing the effect of learning as a result.

In some cases, low mobility warrants scrutiny. It is common for British women to switch from full-time to part-time employment when they become mothers, even though such moves are often accompanied with occupational downgrading. Interestingly, not many women reverse these changes later in their careers, with part-time employment becoming prevalent among mothers. The third chapter offers an explanation emphasizing the role of skill depreciation resulting from leaving their pre-motherhood occupations for this pattern by developing a dynamic model of employment transitions among full-time work, part-time work and non-participation. Building on Albrecht and Vroman (2002), the model incorporates differences in skill levels among workers and skill requirements among jobs. In addition, I add to the worker's problem two more dynamic aspects. The first comes from changes in family circumstances due to the arrival of a child, which changes the pecuniary returns from work in the labour market and potentially changes workers' preferred hours of work. The second is a skill depreciation process, which changes workers' ability to meet the skill requirements of different jobs.

After formulating the workers' problem and describing their transition behaviour, numerical examples of the model are given to examine whether the model can capture the stylized facts of the life-cycle employment transition patterns of British mothers. The model's ability to produce the patterns depends on two decisions made by workers. First, they are willing to downgrade their occupations in order to switch from full-time work to part-time work when childcare needs arise. Second, they prefer to stay employed part time should skill depreciation prevent them from reversing their earlier downward occupational changes. The numerical exercises presented in the chapter find reasonable parameter values satisfying these two conditions. In particular, these parameter values envision a scenario where the median length of the childcare period is seven years, and occupational downgrading is associated with a 20% wage reduction. They also imply that skill depreciation occurs, on average, with five years of employment in the low skill occupation, or with two and a half years of nonparticipation.

Departing from examining processes behind mobility patterns, the fourth chapter studies the role of job mobility in immigrant assimilation. Immigration has always played a major role in Canada, and successful integration of immigrants in the labour market has important implications for the Canadian economy. The vast majority of research in immigrant assimilation has been based on the standard human capital model, while differences in job search behaviours or job search environments between native-born individuals and immigrants have received less attention. However, there are various reasons to expect that new immigrants face a different search environment than natives, including their lower knowledge level of the host-country labour market, qualification recognition problems, and social networks. Taking this idea of differences in job search behaviour one step further, this chapter uses search theory to examine the role of job search in immigrant assimilation. We develop and estimate a Burdett and Mortensen (1998) style equilibrium search model of immigrants operating in the same labour market as natives, with newly arrived immigrants searching for jobs at lower job offer arrival rates than natives but allowed to acquire the same arrival rates according to a stochastic process.

The model is estimated using duration and earnings data from the Canadian Survey of Labour and Income Dynamics (SLID) to measure the difference in job offer arrival rates between new immigrants and natives, estimate how long it takes immigrants to acquire the same job search parameters as natives, and study the implications of this assimilation process for the native-immgirant earnings gap and immigrants' earnings growth. Our estimation results indicate that there are substantial differences in job offer arrival and job destruction rates between natives and newly arrived immigrants. Job offer arrival rates for immigrants are 36% lower while unemployed and 93% lower while employed. These differences are able to account for three quarters of the observed earnings differential between natives and immigrants. Our results also indicate that it takes immigrants 13 years, on average, to assimilate and acquire the same search parameters as natives, with counterfactual exercises indicating that job search assimilation accounts for the vast majority of life-cycle earnings growth of immigrants.

Bibliography

Albrecht, James, and Susan Vroman. 2002. A matching model with endogenous skill requirements. *International Economic Review* 43(1):283–305.

Borjas, Geroge J. 1999. The economic analysis of immigration. In *Handbook of Labor Economics*, eds. Orley C. Ashenfelter and David Card, vol. 3A, 1697 – 1760. Elsevier.

Burdett, Kenneth, and Dale T. Mortensen. 1998. Wage differentials, employer size, and unemployment. *International Economic Review* 39(2):257 – 73.

Greenwood, Michael J. 1997. Internal migration in developed countries. In *Handbook of Population and Family Economics*, eds. Mark R. Rosenzweig and Oded Stark, vol. 1B, 647–720. Elsevier.

Kambourov, Gueorgui, and Iourii Manovskii. 2008. Rising occupational and industry mobility in the United States: 1968–97. *International Economic Review* 49(1):41–79.

Mortensen, Dale T. 1986. Job search and labor market analysis. In *Handbook of Labor Economics*, eds. Orley C. Ashenfelter and Richard Layard, vol. 2B, 849–919. Elsevier.

Neal, Derek. 1999. The complexity of job mobility among young men. *Journal of Labor Economics* 17(2):237–261.

Roy, A. D. 1951. Some thoughts on the distributions of earnings. *Oxford Economic Papers* 3(2):135–146.

Topel, Robert H., and Michael P. Ward. 1992. Job mobility and the careers of young men. *Quarterly Journal of Economics* 107(2):439–479.

Chapter 2

Job and Occupational Search and Learning of Comparative Advantages

2.1 Introduction

Young workers change jobs and occupations more frequently than more experienced workers, with these transitions declining as they age.[1] Human capital accumulation and job search are two processes broadly consistent with these observed mobility patterns, with high job mobility of young workers being viewed as a reflection of their career progression.[2] While these theories view job and occupational transitions as upward mobility and therefore beneficial to workers' labour market outcomes, not all observed transitions are consistent with this idea. It is not uncommon for young workers to experience occupational downgrading, or transitions between seemingly unrelated jobs.[3]

[1]Topel and Ward (1992); Light and McGarry (1998); Neal (1999)

[2]According to the human capital theory, acquisition of specific human capital increases the cost of moving to jobs that do not reward such specific human capital, generating an inverse relationship between age and the propensity of job change. In contrast, the job search theory regards a series of job-to-job transitions as worker's upward mobility and predicts a decline of job changes .

[3]Sullivan (2010) reports a nontrivial likelihood of occupational transition from craftman occupations to labouror and operative occupations (0.13). Sanders (2012) presents an example of an interesting occupational

Transition patterns at odds with either the human capital or job search model may reflect information frictions that individuals face when they enter the labour market. A number of studies argue that young workers' tendency to change jobs reflects their job-shopping behaviour.[4] Through experimentation with different jobs, they gradually sort themselves into better fitting jobs and occupations. In the job-shopping model, therefore, transitions between unrelated jobs are outcomes of a learning process, and an integral part of workers' career progression.

This chapter aims to measure the extent to which learning behaviour affects workers' job search behaviours and transition patterns. Although workers may benefit from job-shopping in the long run, how and how much they are affected by it depends on the natures of initial uncertainty and job search environment. Uncovering the role of them in workers' early career transition behaviour is therefore an important empirical question.

With a view that human capital acquisition, search frictions, and learning all influence workers' transition patterns, it is important to consider them together in a unified framework to understand how these different incentives compete or interact with each other. As an instance in which such a possibility is important, consider occupational choice models in Keane and Wolpin (1997)'s framework, which has been widely used to describe worker's dynamic labour market outcomes. In this framework, workers move to different occupations if the gain from moving to their preferred occupations exceeds the mobility cost. It does not account for possible differences in search frictions that workers experience in different occupations. However, since unlikely jobs are likely to give workers weak incentives to search for, workers may not necessarily transition to occupations that they find the most attractive. In addition, if different occupations provide different learning environments by providing signals for different unknowns or because of differences in the frequency of sig-

transition profile featuring a sequence of unrelated occupations made by a NLSY79 respondent.

[4] See Jovanovic (1979), Miller (1984), and Neal (1999), for example.

nal arrivals, how learning progresses depends on work history.[5]

Based on the dynamic Roy (1951) framework of comparative advantages, this chapter first develops a model of workers' job search behaviour, learning, and human capital acquisition. In the model, workers are endowed with two types of skill, and their skills are enhanced by stochastic learning-by-doing in the labour market. They pursue different occupations because of differences among occupations in returns to skills and skill acquisition opportunities.

Different workers possess different learning abilities for each skill, which are embodied by the probabilities of acquiring higher skill levels in a period workers are employed. These differences lead forward-looking workers to follow different occupational choice paths. For example, a good learner in fine motor skills has an incentive to pursue occupations that enhance fine motor skills. In contrast, for workers who have low learning abilities for any skills, their occupational concerns mainly depend on how their current skill levels are rewarded.

When workers begin their labour market careers, they are not fully informed about their own learning abilities. Instead of making their transition decisions based on the true state of their skill acquisition prospects, they rely on their beliefs about them. Workers resolve this uncertainty by observing their skill acquisition outcomes. During this learning process, workers may pursue different occupations because their beliefs evolve.

Workers face search frictions in the labour market, and obtain jobs through job search. Instead of receiving a job offer at constant offer arrival rates, workers can vary job search effort exerted for different occupations to optimize the likelihood of making their preferred transitions. This captures the non-random transition behaviour across occupations docu-

[5] Additional insights for possible interactions between human capital acquisition and job search processes are obtained from job search models in a single labour market with the inclusion of human capital, such as Rubinstein and Weiss (2006), Burdett et al. (2011) and Bowlus and Liu (2011).

mented in the literature.[6] Gathmann and Schönberg (2010) found that workers were more likely to move between occupations that require similar tasks. Poletaev and Robinson (2008) found that a distant occupational change was associated with a larger wage loss than a close change for displaced workers. This finding also suggests that workers are less likely to make a distant change in occupation unless they receive information favouring such a move.

The model is estimated by simulated maximum likelihood (SML) using job history data of white male high school graduates from the 1979 cohort of the U.S. National Longitudinal Survey of Youth (NLSY79), with individuals' occupational affiliations grouped into skilled white-collar, skilled blue-collar, and non-skilled occupations by cluster analysis based on the 1970 census occupation codes and their skill/task factor scores extracted in Robinson (2011). In sum, the white-collar occupation is associated with high intelligence levels, while the blue-collar occupation is associated with high fine motor skill levels.

The estimates reveal large differences among occupations in search frictions. There are substantial search frictions hindering transitions to the white-collar occupation while the non-skilled occupation is associated with the least search frictions. If an unemployed worker allocates his job search effort equally among the three occupations, the likelihood that a job offer comes from the non-skilled occupation is nearly 50% as opposed to a 10% chance that a job offer comes from the white collar occupation.

The estimated parameters governing skill growth suggest that the two skills in the model are primarily related to the white-collar and blue-collar occupations, respectively. Specifically, one of them is enhanced mostly in the white-collar occupation, while the other grows the fastest with work experience in the blue-collar occupation, with the non-skilled occupation providing a limited skill growth environment. Moreover, the white-collar occupation

[6] Occupational matching models, such as Miller (1984), McCall (1990), Neal (1999), and Pavan (2010, 2011), produce random occupational mobility, since every time workers make turnover decisions, a new occupation is randomly drawn from a given match quality distribution.

does not provide a favourable skill growth environment for the blue-collar related skill, while the same can be said about the blue-collar occupation and the skills primarily related to the white-collar occupation.

Simulation exercises show that while workers initially spread out their job search effort across occupations, learning can have a sizeable effect on their job search effort allocation over time. Workers who learn that they have high skill acquisition probabilities for white-collar related skills choose to focus more on the white-collar occupation, and those who learn the opposite instead choose to allocate less job search effort on this occupation. A signal regarding blue-collar skills, however, has a very small effect on workers' job search effort choice. As a result, the learning process affects workers mostly through signals regarding white-collar skills.

Because of job search frictions, changes in job search effort due to learning do not result in a comparable effect on occupational transition outcomes. Job search frictions have a particularly large consequence for those who direct their job search effort to the white-collar occupation because this occupation presents high frictions. Moreover, the same frictions slow down learning because signals about white-collar skills, the predominant factor driving workers' learning process, are hard to receive without employment in the white-collar occupation. As a result, they also reduce the effect of learning early in the life cycle for all workers.

Two recent papers also study the effect of uncertain comparative advantages in occupational mobility in dynamic frameworks. In Papageorgiou (2014), the economy is populated with workers of different productivity types, which determine their comparative advantages over different occupations. The true state of their comparative advantages, however, are not revealed initially, and they go through a learning process to settle down in the occupations which they are suited for. While his model incorporates directed search to consider the effect of search frictions on the speed of learning, workers are assumed to search only in a single occupation at any given time. However, my estimates indicate that workers' search efforts are spread out over different occupations, especially while unemployed. In addition, his model abstracts from the effect of human capital in mobility decisions, and therefore, attributes any job or occupational change to learning.

Building on Yamaguchi (2012) and Antonovics and Golan (2012), Sanders (2012) developed a learning model in which workers are uncertain about their initial skill endowments. His counterfactual simulation finds a sizeable role for initial uncertainty on workers' occupational mobility. However, search frictions are omitted from Sanders' model. This means that workers' mobility decisions are not influenced by differences in search frictions, and they transition to their preferred occupations every period. This paper shows that search frictions are significant and affect the learning process considerably.

The remainder of the paper is organized as follows. The next section presents the model. Sections 2.3 and 2.4 discuss the estimation strategy and the data used for estimation, respectively. Section 2.5 presents the estimation results. Section 2.6 analyzes how learning affects workers' search effort allocation and occupational transition outcomes based on the parameter estimates. Section 2.7 concludes.

2.2 The Model

2.2.1 Setup

The model starts with the dynamic Roy (1951) framework of comparative advantages. There are $K > 0$ different occupations in the labour market. An occupation is a group of jobs that require workers to perform a common set of tasks.[7] Workers' productivity at a given occupation is determined by their skill levels. Assume that there are two kinds of

[7] Heckman and Sedlacek (1985)

skills and let (s_{1it}, s_{2it}) represent worker i's skill levels in period t. The wage paid to worker i working at job j in occupation k in period t is given by

$$w_{ijkt} = R_j \exp(y_k(s_{1it}, s_{2it})),$$

where function $y_k(s_1, s_2)$ is increasing in both s_1 and s_2. R_j represents the employer specific portion of the wage, and varies across employers.

A worker enters the labour market initially unemployed, with his initial period given by $t = 1$, and stays in the market until an exogenously given terminal date $t = T$. The objective of a worker is to maximize his expected discounted lifetime utility. Assume that a worker's instantaneous utility is given by the logarithm of the wage if employed, or the value of non-market time b if unemployed. Thus, worker i's utility flow in period t is given by

$$u_{it} = \ln w_{ijkt} = r_j + y_k(s_{1it}, s_{2it}) \tag{2.1}$$

where $r_j = \ln(R_j)$. The distribution of r_j is given by c.d.f. $F_k(\cdot)$. I also assume that all workers have the time discount rate of $\frac{1}{1+\rho}$.

The workers' problem has three main components: job search, human capital acquisition, and learning. These processes are discussed in turn in the next three subsections. In the remainder of this section, the individual and time subscripts are subsumed to simplify the notation.

2.2.2 Job Search

Workers find employment through job search. The arrival of a job offer is uncertain, but workers can increase the likelihood of receiving an offer by raising the intensities of their search efforts. I construct the following job search process with multiple occupations by modifying the specification of the endogenous search technology in Christensen et al. (2005). Each period, a worker exerts search efforts $(e_1, e_2, \ldots, e_K)$ where each e_k represents the search effort intensity exerted to find a job in occupation k. Unemployed and employed workers receive a job offer in occupation k with probabilities

$$\lambda_k^0(e_k) = \frac{\lambda_k^0}{1-\alpha} e_k^{1-\alpha} \tag{2.2}$$

and

$$\lambda_k^1(e_k) = \frac{\lambda_k^1}{1-\alpha} e_k^{1-\alpha}, \tag{2.3}$$

respectively, where $0 < \alpha < 1$, and $\lambda_k^0 > 0$, $\lambda_k^1 > 0$ for all $k = 1, 2, \ldots, K$. The assumption $\alpha > 0$ means that search efforts have diminishing returns. If $\alpha = 0$, then the job offer arrival probabilities are proportional to search effort intensities, and the optimal job search strategy dictates that workers search for only one occupation in any given period. The cost function of search efforts $C(e_1, e_2, \cdots, e_K)$ is assumed to be the total search effort, and thus $C(e_1, e_2, \cdots, e_K) = \sum_{k=1}^K e_k$.

It is assumed that a worker receives at most one job offer within a period. Each job j specifies R_j, which is drawn from a known occupation-specific offer distribution with c.d.f. $F_k(\cdot)$.

2.2.3 Skill Acquisition Process

Workers start their labour market careers with their initial skill levels given by (s_1^0, s_2^0), and accumulate their skills through stochastic learning-by-doing on the job.

Each period a workers is employed, he encounters an opportunity to increase at most one of his skills with a certain probability. If he is employed in occupation k, a skill acquisition opportunity arrives with probability $\omega_{k\ell}$ for skill $\ell = 1, 2$. With probability $1-\omega_{k1}-\omega_{k2}$, the worker's skill levels are unchanged in that period. If a worker gains such an opportunity, s_ℓ changes to s'_ℓ according to the following stochastic transition process:

$$s'_\ell = \begin{cases} s_\ell + \gamma_{k\ell} & \text{with probability } \theta_\ell \\ s_\ell & \text{with probability } 1 - \theta_\ell \end{cases} \tag{2.4}$$

where $\gamma_{k\ell} > 0$ for all $\ell = 1, 2$, and $k = 1, 2, \ldots, K$. As shown in equation (2.4), skill ℓ grows by $\gamma_{k\ell}$ with probability θ_ℓ in a given period. $\gamma_{k\ell}$ can vary across occupations so that different occupations provide different skill acquisition opportunities.

It is assumed that $\gamma_{k\ell}$ and $\omega_{k\ell}$ are common to all workers, but the values of θ_1 and θ_2 differ among them. θ_ℓ can be thought of as "the ability to learn" of workers for skill ℓ, as the higher the value of θ_ℓ, the more likely workers are to increase the corresponding skill.

The initial skill levels may be possibly related with the skill acquisition probabilities. For example, a high initial skill level may signal a high skill acquisition probability. Although it is desirable to accommodate this possibility, doing so in general would add considerable complexity to analysis of the model. Therefore, I first abstract from this possible relationship and study a case where the initial skill level and skill acquisition probability are independently drawn. Specifically, I assume that the initial skill level takes either $\underline{s}_\ell$ or $\overline{s}_\ell$ ($\underline{s}_\ell < \overline{s}_\ell$), and $m_\ell \in (0, 1)$ denotes the fraction of workers starting with $\overline{s}_\ell$ for their initial skill ℓ level. It is assumed that for each skill $\ell \in \{1, 2\}$, the population distribution θ_ℓ is given by the Beta distribution with parameters p_ℓ and q_ℓ, and these distributions are known to the workers. In Appendix A, I analyze a case in which a certain form of correlation between θ_ℓ and s^0_ℓ is allowed. Interestingly, estimation results for this specification indicate that the correlation is very small.

2.2.4 Learning

When workers enter the labour market, the true values of their skill acquisition probabilities are unknown. Instead, they form their beliefs about these probabilties from the population distributions of θ_1 and θ_2, and update them according to observed signals. The signals are skill acquisition outcomes, which are Bernoulli random variables. Thus, the model is a Bernoulli learning model.

The learning process is greatly simplified because the Beta distribution is a natural conjugate of the Bernoulli distribution, with the workers' belief distributions remaining the Beta distribution.[8] Specifically, in each skill acquisition opportunity, the belief distribution is updated to the Beta distribution with parameter $p_\ell + 1$ and q_ℓ if a skill growth happens, or updated to the Beta distribution with parameter p_ℓ and $q_\ell + 1$ otherwise.

The true acquisition probability for each of skills 1 and 2 will be revealed to workers if they encounter a sufficiently large number of learning opportunities, because the process leads to a degenerate distribution with a point mass at θ_ℓ.[9] However, this may not occur before the workers leave the labour market in the terminal period, if the workers' occupational choice and the probability of signal arrival at particular occupations results in infrequent signals arrival.

Finally, instead of using the updating rule based on p_ℓ and q_ℓ, it is more convenient to reparameterize the distribution by defining η_ℓ and υ_ℓ by

$$\eta_\ell = \frac{p_\ell}{p_\ell + q_\ell},$$
$$\upsilon_\ell = \frac{1}{p_\ell + q_\ell + 1}.$$

Note that η_ℓ is the mean of the Beta distribution with parameters p_ℓ and q_ℓ, and υ_ℓ can be

[8] See, for example, p.160 in DeGroot (1970).
[9] See DeGroot (1970) p.225

expressed in terms of the mean and variance of the same distribution.[10] There is a one-to-one transformation between (p_ℓ, q_ℓ) and $(\eta_\ell, \upsilon_\ell)$, which allows me to use $(\eta_\ell, \upsilon_\ell)$ as sufficient statistics to capture the learning process. The updating rule of the belief system is given as follows. Let o_ℓ be an indicator variable taking 1 if skill ℓ increase in a given period, or 0 otherwise. Then Bayes' rule yields the following updating rule for η_ℓ:

$$\eta'_\ell = \eta_\ell(1 - \upsilon_\ell) + o_\ell \, \upsilon_\ell, \tag{2.5}$$

and υ_ℓ is updated according to

$$\upsilon'_\ell = \frac{\upsilon_\ell}{1 + \upsilon_\ell}. \tag{2.6}$$

2.2.5 Timing within a Period

The timing of labour market activities within a period is modelled by dividing them into two sub periods. Workers spend the first sub period on job search and transition activities, while they engage in work and may encounter a skill acquisition opportunity the second sub period.

More specifically, in the first subperiod, workers engage in job search, exerting search efforts $e = (e_1, e_2, \ldots, e_K)$. Unemployed workers receive a job offer from occupation k with probability $\lambda^0_k(e_k)$. If they receive a job offer, workers decide whether to accept the offer or not. Employed workers receive a job offer from occupation k with probability $\lambda^1_k(e_k)$, and decide whether to accept the offer or not. Employed workers also face a risk of separating from their current job and transition to unemployment with probability δ_k. I assume that a job offer and job separation are mutually exclusive events.

[10] Specifically, if a random variable X follows the beta distribution with parameters p_ℓ and q_ℓ, then υ_ℓ is given by

$$\upsilon_\ell = \frac{Var(X)}{E(X)(1 - E(X))}.$$

Unemployed workers receive b, the value of non-market time in the second sub period. Employed workers are paid based on their skill levels at the beginning of the second sub period. Then they may encounter a skill acquisition opportunity. Workers update their beliefs about (θ_1, θ_2) based on skill acquisition outcomes according to Bayes' rule. It is possible that employed workers find it preferable to quit the current job given the skill acquisition outcomes. In that case, they are able to leave their jobs without incurring any costs and become unemployed. However, it is assumed that they do not engage in job search activities immediately after job separations and spend one period unemployed.[11]

2.2.6 Dynamic Programming Formulation of Worker's Problem

The state variables used to describe the workers' dynamic problem are (t, k, S, r), where t indicates the current period, and k gives their employment and occupation states with $k = 0$ indicating they are unemployed, and $k = 1, 2, \ldots, K$ indicating that they are employed in occupation k. S is a set of state variables pertaining to their skills, consisting of $(s_1, \eta_1, \upsilon_1, s_2, \eta_2, \upsilon_2)$. Recall that s_1 and s_2 represent their skill levels, and η_ℓ and υ_ℓ for $\ell = 1, 2$ are the parameters of their belief distributions about their skill acquisition probabilities. r is the employer specific component in the log wage equation (2.1), and only relevant when the workers are employed.

Each period, workers engage in job search in the first sub-period .Then employed workers engage in work, and may experience human capital acquisition and learning in the second sub-period. The value to workers carrying state variables (t, k, S, r) in the first sub-period is given by $W_t^k(S, r)$, if they are employed in occupation k, or $W_t^0(S)$, if they are unemployed. The second sub-period features work, learning, and job separation, and the

[11]This assumption is made to simplify the estimation process implemented in this chapter. The estimation data constructed for this paper do not contain information on whether workers ended jobs voluntarily or involuntarily. Without means to distinguish direct job-to-job transitions from job-quit-to-job transitions, likelihood-based estimation involves substantial complication.

value of carrying state variables (t, k, S, r) in the second sub-period is given by $V_t^k(S, r)$, if they are employed in occupation k, or $V_t^0(S)$, if they are unemployed.

The value to employed workers in the first sub-period, $W_t^k(S, r)$, is given by

$$W_t^k(S, r) = \max_{(e_1,\ldots,e_K)} \Big[\sum_{k'=1}^{K} \Big(\lambda_{k'}^1(e_{k'}) \int_{-\infty}^{\infty} \max[V_t^{k'}(S, r') - V_t^k(S, r),\ 0\,] dF_{k'}(r') \Big) + \delta_k V_t^0(S) + (1-\delta_k) V_t^k(S, r) - C(e_1, \ldots, e_K) \Big]. \quad (2.7)$$

Equation (2.7) reflects the workers' job search behaviour. The offer is accepted if its value to a worker exceeds the value of staying in the present state. The first term of the right-hand side of the equation represents the expected gain due to job search, which accrues from a transition to a new job whose value exceeds the current job. The last term is the cost associated with the search efforts. The equation states that workers optimally choose the search efforts to maximize the net benefit of job search each period.

Because $\lambda_k^1(e_k)$ is strictly concave in e_k and the marginal value of search effort is decreasing but tends to infinity at effort levels sufficiently close to 0, the optimal effort intensity to search for a job in occupation k' is characterized by the first-order condition

$$\frac{\partial \lambda_{k'}^1(e_{k'})}{\partial e_{k'}} \int_{-\infty}^{\infty} \max[V_t^{k'}(S, r') - V_t^k(S, r), 0\,] dF_{k'}(r') = 1.$$

Let $e_t^*(k, S, r, k')$ denote the optimal search effort intensity function, where k and k', respectively, indicate the current occupation and a prospective occupation. Then solving the first-order condition for $e_{k'}$ with $\lambda_{k'}^1(e_{k'})$ specified by equation (2.3) gives $e_t^*(k, S, r, k')$ as

$$e_t^*(k, S, r, k') = \left(\lambda_{k'}^1 \int_{-\infty}^{\infty} \max \left[V_t^{k'}(S, r') - V_t^k(S, r), 0 \right] dF_{k'}(r') \right)^{\frac{1}{\alpha}}. \quad (2.8)$$

Substituting equation (2.8) into equation (2.7) yields

$$W_t^k(S,r) = \frac{\alpha}{1-\alpha}\sum_{k'=1}^{K}\left(\lambda_{k'}^{1}\int_{-\infty}^{\infty}\max[V_t^{k'}(S,r') - V_t^k(S,r),0]dF_{k'}(r')\right)^{\frac{1}{\alpha}} \\ + \delta_k V_t^0(S) + (1-\delta_k)V_t^k(S,r).$$

Similarly, $W_t^0(S)$, the value of being unemployed at the beginning of period t, is given by

$$W_t^0(S) = \max_{(e_1,\ldots,e_K)}\Bigg[\sum_{k'=1}^{K}\left(\lambda_{k'}^{0}(e_{k'})\int_{-\infty}^{\infty}\max[V_t^{k'}(S,r') - V_t^0(S),0]dF_{k'}(r')\right) \\ + V_t^0(S) - C(e_1,\ldots,e_K)\Bigg]. \quad (2.9)$$

Equation (2.9) has the same interpretation as equation (2.7) but the job offer arrival function is given by $\lambda_k^0(\cdot)$ instead of $\lambda_k^1(\cdot)$. Solving the first-order condition for the maximization problem above, the optimal search effort function for unemployed workers is given by

$$e_t^*(0,S,k') = \left(\lambda_{k'}^{0}\int_{-\infty}^{\infty}\max\left[V_t^{k'}(S,r') - V_t^0(S),0\right]dF_{k'}(r')\right)^{\frac{1}{\alpha}}. \quad (2.10)$$

Then the value function can be rewritten as

$$W_t^0(S) = \frac{\alpha}{1-\alpha}\sum_{k'=1}^{K}\left(\lambda_{k'}^{0}\int_{-\infty}^{\infty}\max[V_t^{k'}(S,r') - V_t^0(S),0]dF_{k'}(r')\right)^{\frac{1}{\alpha}} + V_t^0(S).$$

Now the second sub-period is considered in detail. Before turning to a characterization of $V_t^k(S,r)$ and $V_t^0(S)$, however, I introduce the following notation to facilitate the description of the state variables' laws of motion related to this sub-period.

S evolves in response to skill acquisition outcomes on the job. Let $S_{k\ell}(a)$ denote the values of the state variables when a worker encounters an opportunity to increase his skill

ℓ level while working in occupation $k = 1, 2, \ldots, K$. Variable a takes 1 if the worker's skill ℓ has increased in the current period or 0 otherwise. For example, $S_{31}(1)$ represents the state variables after a worker gets to accumulate skill 1 while working in occupation 3, and $S_{21}(0)$ represents the state variables after a worker encounters a chance to accumulate skill 1 but fails to do so. Then based on equations (2.4), (2.5) and (2.6), $S_{k\ell}(a)$ is given by

$$S_{k1}(a) = \Big(s_1 + a\gamma_{k1},\ \eta_1(1 - \upsilon_1) + a\upsilon_1,\ \frac{\upsilon_1}{1 + \upsilon_1},\ s_2, \eta_2,\ \upsilon_2\Big),$$

$$S_{k2}(a) = \Big(s_1, \eta_1, \upsilon_1,\ s_2 + a\gamma_{k2},\ \eta_2(1 - \upsilon_2) + a\upsilon_2,\ \frac{\upsilon_2}{1 + \upsilon_2}\Big).$$

Now I characterize $V_t^k(S, r)$ and $V_t^0(S)$ to complete the description of the workers' problem. First, after completing job search in the terminal period, $t = T$, a worker only cares about the current utility flow, therefore $V_T^k(S, r)$ and $V_T^0(S)$ are given, respectively, as

$$V_T^k(S, r) = r + y_k(s_1, s_2)$$

and

$$V_T^0(S) = b.$$

For a non-terminal period $t = 1, 2, \ldots, T - 1$, $V_t^k(S, r)$ can be given by

$$\begin{aligned} V_t^k(S, r) = r + y_k(s_1, s_2) + \frac{1}{1 + \rho}\Big[&(1 - \omega_{k1} - \omega_{k2}) \max[W_{t+1}^k(S, r), V_{t+1}^0(S)] \\ + \sum_{\ell=1}^{2} \omega_{k\ell} \int_0^1 \sum_{a=0}^{1} \tilde{\theta}_\ell^a (1 - \tilde{\theta}_\ell)^{1-a} &\max[W_{t+1}^k(S_{k\ell}(a), r), V_{t+1}^0(S_{k\ell}(a))] f(\tilde{\theta}_\ell; \eta_\ell, \upsilon_\ell) d\tilde{\theta}_\ell\Big]. \end{aligned} \quad (2.11)$$

The first term on the right-hand side is the current utility flow, and the remainder of the right-hand side of equation (2.11) expresses the continuation value of occupying state (t, k, S, r). With probability $1 - \omega_{k1} - \omega_{k2}$, a worker does not encounter a skill acquisition opportunity and thus his state variables remain unchanged. Then he makes a decision whether to quit or not the current job.[12] With probability $\omega_{k\ell}$, the worker encounters a chance to increase his skill ℓ level, and the state variables are updated according to the skill acquisition outcome. Since the worker does not know his skill acquisition probability for certain, the continuation value is based on his current belief distribution. By calculating the integrals in equation (2.11), the above equation can be simplified to

$$\begin{aligned} V_t^k(S,r) = r + y_k(s_1,s_2) + \frac{1}{1+\rho}\Big[&(1-\omega_{k1}-\omega_{k2})\max[W_{t+1}^k(S,r), V_{t+1}^0(S)] \\ &+ \sum_{\ell=1}^{2}\omega_{k\ell}\sum_{a=0}^{1}\eta_\ell^a(1-\eta_\ell)^{1-a}\max[W_{t+1}^k(S_{k\ell}(a),r), V_{t+1}^0(S_{k\ell}(a))]\Big]. \end{aligned} \tag{2.12}$$

While in unemployment, workers neither accumulate skills nor acquire any new information about their skill acquisition probabilities. Thus $V_t^0(S)$ is given simply by

$$V_t^0(S) = b + \frac{1}{1+\rho}W_{t+1}^0(S),$$

for $t = 1, 2, \ldots, T-1$.

2.2.7 Analysis of Worker's Problem

The worker's problem involves several decisions and a large number of state variables, making it intractable to fully analyze. Yet, it is straightforward to show that $V_t^k(S,r)$ and $W_t^k(S,r)$ are increasing in r for all $k = 1, 2, \ldots, K$. This establishes that workers employ reservation wage strategies for accepting and quitting jobs. Specifically, let $\bar{r}_t(k, S, r, k')$ denote the reservation wage for occupation k' set by workers working in occupation k in

[12] Notice that the value after separating from the current job is given by $V_{t+1}^0(S)$, not by $W_{t+1}^0(S)$, reflecting the modelling assumption that workers do not engage in job search immediately after a job quit.

period t. $\bar{r}_t(k, S, r, k')$ is implicitly defined by equation

$$V_t^{k'}(S,\ \bar{r}_t(k, S, r, k')) = V_t^k(S, r).$$

Clearly, the reservation wage for jobs in the current occupation coincides with the current r, i.e., $\bar{r}_t(k, S, r, k') = r$ if $k' = k$. Similarly, the reservation wages set by unemployed workers are defined by

$$V_t^{k'}(S,\ \bar{r}_t(0, S, r, k')) = V_t^0(S).$$

As for the voluntary job quit decision, let $\bar{q}_t(k, S)$ denote the reservation wage governing the workers' quitting decision in period t. This is implicitly defined by equation

$$W_{t+1}^k(S, \bar{q}_t(k, S)) = V_{t+1}^0(S).$$

The total search intensity while employed is calculated as

$$\sum_{k'=1}^{K} e_t^*(k, S, r, k') = \sum_{k'=1}^{K} \left(\lambda_{k'}^1 \int \max(V_t^{k'}(S, r') - V_t^k(S, r), 0) dF_k(r') \right)^{\frac{1}{\alpha}}$$

from equation (2.8). The monotonicity of $V_t^k(S, r)$ with respect to r establishes that the total search intensity is non-increasing in r while everything else is held constant.

2.2.8 Numerical Issues

The value functions, $V_t^0(S)$, $W_t^0(S)$, $V_t^k(S, r)$ and $W_t^k(S, r)$ for $k \in \{1, 2, \ldots, K\}$, are calculated backwards starting at the terminal period $t = T$. Since they involve a large number of the state variables and face a severe problem of the curse of dimensionality, I rely on an approximation method similar to that in Keane and Wolpin (1994, 2001). Specifically, I randomly draw a large number of the state vectors, and evaluate the value functions at these vectors using equation (2.12) for each $k \in \{1, 2, \ldots, K\}$. Then I fit approximating polynomials for $V_t^0(S)$ and $V_t^k(S, r)$ by regression. The estimated regression coefficients are stored to be used subsequently.

2.3 Estimation and Identification Strategy

2.3.1 Estimation Procedure

The model is estimated by simulated maximum likelihood (SML). I construct the individual contribution to the likelihood from observed transitions and wage data. Individual i's labour market history can be expressed by $O_{it} = \{(occ_{it}, spl_{it}, w_{it})\}_{t=1}^{T_i}$ where T_i represents the last quarter of the observed history of individual i. Variable $occ_{it} \in \{0, 1, \ldots, K\}$ represents individual i's employment state and occupation in period t, with $occ_{it} = 0$ indicating that individual i is unemployed in period t. Variable spl_{it} keeps track of individual i's position within the current employment cycle.[13] Specifically $spl_{it} = 0$ if individual i is unemployed in period t, and $spl_{it} = j$ if the worker is in the jth job after the most recent unemployment spell. w_{it} is the wage observation for individual i in period t.[14]

I add the following features to the model to make it estimable. The model cannot account for wage decreases in several instances, including wage decreases on the same job, and those occurring when individuals move to a new job in the same occupation. If they are observed in the data, zero likelihood would be assigned to these events by the model. In order to deal with these probability-zero events, I assume that wages are observed with multiplicative measurement error with a lognormal distribution. In the log wage term, wage

[13] As defined in Wolpin (1992), an employment cycle consists of an unemployment spell and continuous job spells. Each cycle begins with a new unemployment spell, and ends with a transition to another unemployment spell.

[14] Note wage information is not available for all periods workers were employed. Therefore, although subsumed in the discussion, variables indicating that wage information was available for given periods are included in the estimation data.

observation w_{it} can be given by

$$\ln w_{it} = \ln w^*_{ijkt} + \varepsilon_{it},$$

where w^*_{ijkt} is the true wage value earned by individual i working at job j in occupation k in period t, and ε_{it} is the measurement error term, which follows the *i.i.d.* normal distribution with mean 0 and standard deviation σ_ε.

The true log wage value is determined by equation (2.1). I specify this equation as

$$\ln w^*_{ijkt} = r_j + \beta_{k1} s_{1it} + \beta_{k2} s_{2it}$$

for $k = 1, 2, \ldots, K$, with β_{K1} and β_{K2}, the coefficients for occupation K, normalized to 1.[15] In addition, I assume that r_j, the employer specific component in the log wage function, is drawn from the normal distribution with mean μ_k and standard deviation σ_k, with μ_K normalized to 0.

Furthermore, the following modification to the model is made to reconcile the observed data and model solution. As shown later, a large fraction of individuals in my estimation sample were employed in the first post-schooling quarter. Attributing this high initial employment rate to search efforts in the first quarter would likely be at odds with the more moderate transition rates observed in the subsequent quarters. The initial high employment rate likely reflects search activities in school as well as those done after transition, and this calls for additional structure to the model to properly account for it.

Rather than explicitly modelling search activities in school, I take the following approach to deal with this issue. First, the probability of being employed in the first quarter

[15] I adopted this linear specification after more general specifications were tried, for example, with quadratic terms in skill levels. The quadratic terms were not jointly significant at the 5% significance level in the likelihood ratio test.

is included in the set of parameters to be estimated. This probability is denoted by π. By directly estimating the fraction of the sample initially employed, it is possible to reconcile the discrepancy between the employment rate in the first quarter and transition rates in the subsequent quarters.

This does not solve the problem entirely, however, for I also need to account for individuals' occupations to calculate the likelihood function. To this end, I specify the probability of working in a particular occupation in the first period as follows:

$$Pr(occ_{i1} = k) = Pr(occ_{i1} \in \{1, 2, \ldots, K\}) Pr(occ_{i1} = k | occ_{i1} \in \{1, 2, \ldots, K\}).$$

This equation states that the relevant probability is divided into two components: the probability of being employed and the conditional probability that worker i works in occupation k in the first period. The first component is equal to π, which is to be estimated directly, and the second component is specified as

$$Pr(occ_{i1} = k | occ_{i1} \in \{1, 2, \ldots, K\}) = \frac{\lambda_k^0(e_t^*(0, S, k))(1 - F(\bar{r}_t(0, S, r, k))}{\sum_{h=1}^{K} \lambda_h^0(e_t^*(0, S, h))(1 - F(\bar{r}_t(0, S, h)))}. \tag{2.13}$$

The numerator on the right-hand side of equation (2.13) is the product of the job offer arrival probability from occupation k and the probability of accepting an offer given their reservation wage strategy. Therefore, it is the probability of exiting from unemployment to occupation k given workers' optimal job search strategies. The denominator is the sum of these probabilities for all the occupations.

Accounting for individuals' wage and transition profiles using the model requires me to account for the several unobservable state variables: workers' skill levels, skill acquisition probabilities, beliefs and employer-specific wage components. These pieces of information are given by latent variables (θ_1, θ_2) and the state variables in the model

$$V_{it} = (s_{1it}, \eta_{1it}, \upsilon_{1it}, s_{2it}, \eta_{2it}, \upsilon_{2it}, r_j).$$

Define $\tilde{O}_{it}$ by

$$\tilde{O}_{it} = (O_{it}, V_{it}).$$

The likelihood contribution of individual i is obtained by integrating out the unobservable variables from the probability of realizing a history given by $\{\tilde{O}_t\}_{t=1}^{T_i}$, θ_1 and θ_2:

$$L_i = Pr(\{O_t\}_{t=1}^{T_i}) = \int \cdots \int Pr(\{\tilde{O}_t\}_{t=1}^{T_i} \theta_1, \theta_2) dV_{T_i} \cdots dV_1 d\theta_1 d\theta_2.$$

Although the closed-form expression for the integral is unavailable, it can be calculated by Monte Carlo simulation because of ease of simulating profiles of V_{it} consistent with the model. The SML estimator is a set of structural parameters that maximizes the log-likelihood function defined by

$$\ell(\Theta) = \sum_{i=1}^{N} \ln L_i$$

where Θ represents the parameters to be estimated.

2.3.2 Identification Discussion

The model parameters are identified through differences in observed wages and quarterly transition data with respect to a variety of dimensions. Formally establishing identification of the model parameters is difficult due to the complex nature of the model. Instead, I attempt to provide the intuition for identification.

The parameters governing the job search process, i.e., λ_k^0 and λ_k^1 for $k = 1, 2, \ldots, K$, and α, are identified from the observed quarterly transitions. In particular, α, which enters into the job offer arrival probability functions in equations (2.2) and (2.3), governs the extent of diminishing returns to scale in search effort for a single occupation. Therefore wide differences in transition destination indicates strong diminishing returns to focusing on a single occupation, while concentrated transition destinations indicates weak diminishing returns.

The estimation data set provides panel data on wage observations. The parameters in the log wage equations, in the wage offer distributions, and those governing the skill acquisition process are identified from differences in wage observation across various dimensions. Wage growth rates observed in different occupations are used to identify $\beta_{k\ell}$ and $\gamma_{k\ell}$. Differences in wage growth among individuals are used to identify the distribution of skill acquisition probabilities, i.e., p_ℓ and q_ℓ. Differences in individuals' first wage observations are used to identify the distributions of initial skill levels. Wages out of unemployment and wage growth observed at job-to-job transitions are used to identify μ_k and σ_k. Note that the identification relies on the assumptions $\beta_{K1} = \beta_{K2} = 1$ and $\mu_K = 0$.

The time discount rate ρ is set to 0.0192. This value is equivalent to an annual discount rate of 5 percent.

2.4 Data

The estimation strategy employed in the paper requires panel data that follow individual labour market outcomes since the school-to-work transition. To meet this data requirement, I use the U.S. National Longitudinal Survey of Youth 1979 (NLSY79) to construct quarterly post-schooling labour market histories.

NLSY79 is a longitudinal study of a nationally representative cohort of individuals who were between 14 and 21 years old when they were interviewed for the first time. The survey conducted annual interviews from 1979 to 1994 and then the interview frequency was changed to bi-annual from then on. It collects information on respondents' work and schooling activities since the last interview, enabling the construction of detailed post-schooling labour market histories including wages, occupations, and spells of jobs and unemployment.

The estimation sample is constructed from the white male cross section sample, and uses the interview responses from 1979 to 2000. The sample is restricted to high school graduates to make the sample population as homogeneous as possible. It also excludes respondents who were in the armed forces or self-employed. Moreover, it excludes respondents who graduated from high school before 1978 because it is not possible to construct complete post-schooling job histories before 1978.

2.4.1 Job History

In the estimation data set, a job is defined as a continuous employment relationship with a particular employer which lasted at least 13 weeks of full time work.[16] In each interview, NLSY79 records the starting and stopping weeks of up to 5 jobs held by respondents since the last interview. Using this information, I first construct their weekly job histories consisting all job spells. A job for which relevant job characteristics are missing or inconsistent with each other (e.g., information on occupation or hours of work is unavailable, or the stopping week precedes the starting week), results in the job history being censored immediately before the problematic job. Thus the likelihood calculation is based only on the information prior to this spell.

Although my model assumes that workers hold only one job in a given quarter, it is quite common to find cases in which respondents worked for more than one employer in the same week. To deal with these cases, I employ a commonly used treatment for this in

[16]Full-time work is defined as 35 weekly hours or more.

the literature. Specifically, when encountering instances of multiple job holding, I assume that the one which started later did not start until the one that started earlier ended.

The model is estimated in quarterly terms (13 weeks), and therefore the weekly job histories constructed according to the above steps are converted into quarterly histories as follows. Respondents are considered employed in a given quarter if they had job spells whose total length within the quarter was no less than 7 weeks. Otherwise, they are considered unemployed in the quarter. If more than one job spell cover a particular quarter, then the one that the respondent worked for more weeks during that quarter is considered as the job for that quarter.

As for earnings information, NLSY79 provides at most one earnings observation for each reported job during each interview. This information is assigned to particular quarters according to the following rules. If a job was ongoing during the interview week, I treat the reported pay as the one earned in the interview week and the corresponding quarter. If the job spell ended earlier, the reported pay is treated as the pay earned in the quarter when the job ended. The earnings information is converted into quarterly terms based on the reported time unit of pay, converted into the real terms using 1993 dollars.

Finally, the calendar quarter of labour market entry is determined for each respondent by locating the last time he was enrolled in high school. If it was in February, March or April, then I determine the second quarter of the year as the entry quarter to the labour market. Similarly, if it is in May, June or July, I determine the third quarter of the year as the entry quarter to the labour market. If it is August, September or October, I determine the fourth quarter of the year as the entry to the labour market. If it is November, December or January, then I determine the nearest first calendar quarter as the initial quarter.

2.4.2 Occupation Classification

The estimation strategy requires that the data provide information on occupations held by individuals. To meet this requirement, the 1970 census 3-digit occupation codes assigned to reported jobs are used to determine individuals' occupations during given job spells.[17] One limitation of using the Roy model is that estimation becomes unmanageable with a large number of occupations. Thus aggregation of occupations is made to make the estimation feasible. This treatment is also justified because of concerns of misclassification of occupation codes in the NLSY79 as pointed out by Neal (1999). Furthermore, it may be desirable even in the absence of classification error because changes in occupation codes at such detailed level may not involve material changes in task performed on the job, as Robinson (2011) argues.

The model defines occupations as sets of jobs requiring workers to perform similar tasks, and it is important to reflect this when assigning occupations into a small number of groups. While previous studies on occupational choice, such as Keane and Wolpin (1997) and Sullivan (2010), aggregated 3-digit occupations based on which major occupation category they belong to (e.g., professional, managerial, etc.) I turn to the growing literature on task and skill content of occupations to guide my classification scheme.

This literature takes advantage of surveys such as the U.S. Dictionary of Occupational Titles, or the German Qualification and Career Survey, which collect detailed information on tasks performed and skills required for a wide range of occupations. Poletaev and Robinson (2008) and Robinson (2011) are examples of these studies for the United States. Using factor analysis, they converted the rich array of characteristics of occupations from the Dictionary of Occupational Titles to a small number of numerical scores. Each extracted factor captures a certain type of skill such as intelligence or fine motor skill, and its value

[17] This assumes that individuals' occupations do not change within jobs.

indicates the extent to which the skill is utilized. Thus each score serves as the coordinate on a space where occupations are positioned with respect to their skill or task content. This allows us to discuss the distance or closeness of occupations.[18] I use these factor scores as summary measures of task content of occupations and use them to construct aggregated occupational groups used in the present analysis.

To group occupations, I use cluster analysis to aid the aggregation process. Cluster analysis is a statistical method used to divide a sample into a number of subsets based on their observed set of characteristics. Of various clustering methods developed and implemented in the literature, I adopt the finite mixture model of clustering.[19] Detail of the cluster analysis performed for this paper is given in Appendix B.

The census occupation codes are grouped into three sets of aggregated occupations. Table 2.1 shows the summary of the constructed occupation groups. With a few exceptions, the professional occupations are evenly divided into occupations 1 and 2. The professional occupations assigned to occupation 2 are in general either technical (e.g. 153: electorical and electronic engineering technicians) or in natural or medical science (e.g. 80: clinical laboratory technologists and technicians). Managerial and sales occupations are mostly included in occupation 1, and most craftsmen occupations are grouped into occupation 2. Occupation 3 represents clerical, operatives, labourer and service occupations. Overall, occupations 1 and 2 represent respectively white- and blue- collar occupations that require high skill levels.

The cluster analysis employed three factors associated with intelligence (factor 1), fine motor skills (factor 2), and physical strength (factor 3).[20] The average factor scores in each aggregated occupation groups are presented in Table 2.2. These constructed occupation

[18]To make the extracted information available to an analysis with census occupations, the crosswalk between the DOT occupations and census occupations was used. and the skill or task connects are calculated.

[19]See Fraley and Raftery (2002) for the methodology.

[20]While 4 factors are retained in Robinson (2011)'s factor analysis, one of these factors, which is associated with visual skills, is omitted in my cluster analysis because use of this factor produced a less intuitive result.

Table 2.1: Summary of Aggregated Occupation Groups

1970 Census	Occupation		
Occupation Category	1	2	3
Professional	57	58	5
Managers	21	2	0
Sales workers	10	0	4
Clerical workers	4	0	40
Craftmen	1	78	13
Operatives	0	7	44
Transport equipment operatives	0	6	5
Laborers	0	2	13
Farmers	0	3	3
Service workers	5	7	27

Table 2.2: Average Factor Scores of Aggregated Occupation Groups

Group	Factor 1	Factor 2	Factor 3
1	1.280	-0.300	0.593
2	-0.276	1.366	0.739
3	-0.738	-0.378	-0.522

groups have the following skill/task characteristics. On average, occupation group 1 scores high on the factor associated with intelligence, while occupation group 2 scores high on the factor associated with fine motor skills. Occupation 3 scores low on all the three factors.

Finally, the occupation classification constructed so far is used to assign each job spell to an aggregated occupation group. While the model assumes that a worker's occupation remains the same within a job, I observe cases in the data where different 3 digit occupation codes were assigned to the same job spell over years, and these reported 3 digit codes are mapped to different groups in my aggregated occupation classification. These discrepancies need to be resolved. To this end, I institute a set of rules handling these problems, which is described in Appendix B.

Table 2.3 presents the results of the assignment. Most of the jobs are assigned either

Table 2.3: Job Spells by Aggregated Occupation

Group	Frequency	Proportion
1	279	15.15
2	810	44.00
3	752	40.85

occupation 2 or 3. This is understandable given the sample population is high school graduates and occupation 1 mainly consists of professional occupations, which are associated with higher educational qualifications.

2.4.3 Analysis of Estimation Data

The final sample contains 413 individuals with an average length of post-schooling history of 61 quarters. The average number of jobs spells held during the observed history is 4.5, with an average duration of 12.4 quarters. The average number of unemployment spells experienced during the observed history is 1.8, with an average duration of 3.4 quarters.

Table 2.4 presents the probabilities of making different kinds of transitions while employed each quarter. In the table, the sample period was divided into 4 distinct periods as follows to show their patterns over time: (1) the first 10 quarters, (2) the second 10 quarters, (3) the following 20 quarters, and (4) the last 20 quarters. Three types of transitions are given here: job-to-job transitions, job-to-job transitions involving a change in occupation, and job-to-unemployment transitions. All of them show declining profiles, indicating that individuals' employment relationships become more stable as they age.

Table 2.5 presents the quarterly and yearly occupation transition matrices. Again the transition rates are reported separately for the 4 periods defined above. As discussed previously, the probability of making an occupational change declines with age, and this can be seen in the occupation transition matrices. The diagonal elements of the transition matrices

Table 2.4: Probability of Transitions from Jobs per Quarter

	1-10Q	11-20Q	21-40Q	41-60Q
Probability of making a job-to-job transition	0.063	0.044	0.043	0.031
Probability of making an occupation change	0.026	0.021	0.015	0.011
Probability of making a job-to-unemployment transition	0.053	0.042	0.022	0.017

for occupations 2 and 3 increase with age. Interestingly, there is no discernible age pattern for occupation 1.

The quarterly occupational transition matrix also shows that occupation 3 initially captures the largest share of the destination of unemployment-to-job transitions. This rate remains stable for the first 10 years. Meanwhile, transitions from unemployment to occupations 1 and 2 increase over time, and after the first 10 years, occupation 2 overtakes occupation 3 as the top destination of transitions out of unemployment.

Quarterly transitions between different occupations are small and it is difficult to observe any patterns in them. More clear patterns emerge in the yearly transition matrix. In particular, transitions between occupation 2 and 3 are more frequent in the initial 10 quarters than in later quarters. Transitions to occupation 2 from occupation 3 are the most common type of occupational change. This is likely due to occupational upgrading. Occupation 2 mainly consists of professional and craftsmen occupations and occupations 3 includes labourers and operatives. Sullivan (2010) also reports high transitions from the latter occupation groups to the craftsmen occupations in his data from NLSY 79.

Figure 2.1 shows the fraction of the sample working in each occupation over time. Slightly above two-fifths of the sample remains unemployed in their first quarter in the labour market. The fraction unemployed rapidly drops in the first 4 years to around 0.1, and then gradually declines to 0.07.

Table 2.5: Transition Profiles between Occupations

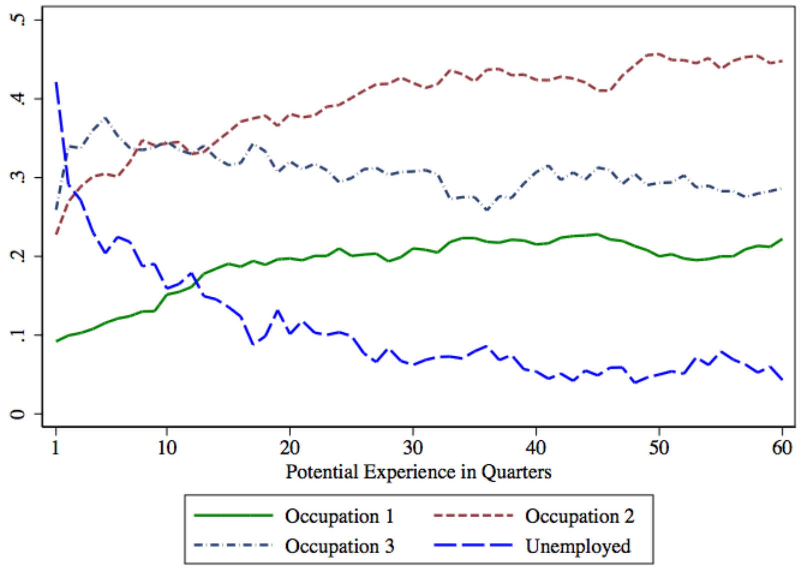
Quarterly Transition

Occupation (Origin)	Period	UE	Occ 1	Occ 2	Occ 3
		Occupation (Destination)			
Unemployed (UE)	1-10Q	0.721	0.025	0.107	0.145
	11-20Q	0.685	0.040	0.124	0.144
	21-40Q	0.689	0.043	0.110	0.146
	41-60Q	0.716	0.040	0.138	0.098
Occupation 1 (Occ 1)	1-10Q	0.015	0.969	0.008	0.008
	11-20Q	0.019	0.967	0.012	0.006
	21-40Q	0.011	0.973	0.010	0.004
	41-60Q	0.014	0.976	0.005	0.006
Occupation 2 (Occ 2)	1-10Q	0.051	0.006	0.925	0.017
	11-20Q	0.045	0.007	0.938	0.008
	21-40Q	0.021	0.005	0.969	0.006
	41-60Q	0.014	0.003	0.977	0.005
Occupation 3 (Occ 3)	1-10Q	0.068	0.007	0.023	0.901
	11-20Q	0.050	0.006	0.020	0.922
	21-40Q	0.032	0.005	0.015	0.950
	41-60Q	0.022	0.006	0.009	0.965

Yearly Transition

Occupation Origin	Period	UE	Occ 1	Occ 2	Occ 3
		Destination			
Unemployed (UE)	1-10Q	0.456	0.048	0.215	0.279
	11-20Q	0.391	0.082	0.244	0.281
	21-40Q	0.364	0.097	0.210	0.311
	41-60Q	0.393	0.113	0.264	0.217
Occupation 1 (Occ 1)	1-10Q	0.022	0.935	0.028	0.026
	11-20Q	0.025	0.895	0.055	0.024
	21-40Q	0.026	0.915	0.036	0.019
	41-60Q	0.033	0.909	0.029	0.032
Occupation 2 (Occ 2)	1-10Q	0.097	0.035	0.805	0.063
	11-20Q	0.080	0.034	0.862	0.028
	21-40Q	0.037	0.018	0.914	0.033
	41-60Q	0.028	0.015	0.933	0.018
Occupation 3 (Occ 3)	1-10Q	0.122	0.038	0.092	0.747
	11-20Q	0.079	0.028	0.079	0.813
	21-40Q	0.061	0.023	0.068	0.854
	41-60Q	0.046	0.020	0.039	0.905

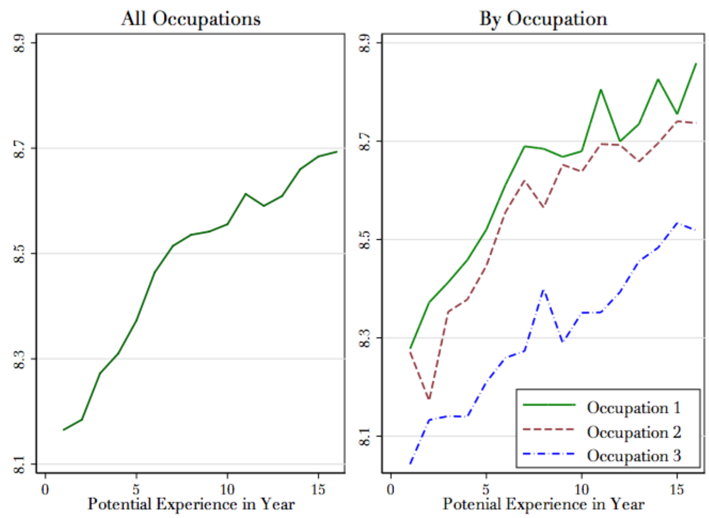
Figure 2.1: Fraction of Individuals Working in Given Occupations



A large fraction of initial employment is in either occupations 2 or 3, with occupation 3 accounting for the largest share. The fractions of the sample working in occupations 2 and 3 initially increase rapidly. However, while the fraction for occupation 2 steadily rises and reaches about 0.45, the fraction for occupation 3 experiences a steady decline after the first year and then stabilizes. Occupation 2 accounts for the largest fraction of employed high school graduates' occupations after about 3 or 4 years. The fraction working in occupation 1 remains the lowest among the three occupations, but the share of the sample working in occupation 1 gradually increases from about 0.1 to slightly above 0.2 over the 15 year period.

Figure 2.2 presents the age profiles of mean log earnings. As discussed previously, earnings information is not available for every quarter worked, and this appears to induce fluctuations if the average is calculated by quarter. Thus to smooth the earnings profile, the average was calculated annually. The left panel of the figure shows the values corresponding to all employed workers. The graph shows the concave profiles of log earnings, which

Figure 2.2: Observed Log Wage



is widely documented in the literature.[21] It shows that log earnings increases by nearly 0.5 in the first 15 years in the labor market. The right-hand panel of the figure shows the average log earnings among workers in different occupations. These graphs do not account for self-selection of workers, and therefore, do not reflect the occupation specific age profiles of earnings for the whole sample population. Those working in occupation 1 earn the most throughout the sample period, while those working in occupation 3 earn the least on average. Earnings for those working in occupation 2 are closer to that for occupation 1. As for the wage growth profiles in different occupations, occupations 1 and 2 appear to have higher growth rates than occupation 3.

[21] See Rubinstein and Weiss (2007) for example.

2.5 Estimation Results

2.5.1 Parameter Estimates

The upper panel of Table 2.6 presents the parameter estimates governing the skill acquisition process and initial skill level distributions. Given the parameter estimates, workers gain a skill acquisition opportunity each period they work in occupations 1, 2, or 3 with probability 0.87, 0.88, or 0.90, respectively.[22] 84% of skill acquisition opportunities are for skill 1 in occupation 1. The corresponding values for skill 1 in occupations 2 and 3 are 8% and 35%, respectively.

Skill acquisition opportunities also represent opportunities to receive signals about workers' own skill acquisition probabilities. Given the estimates of $\omega_{k\ell}$, a signal about θ_1 arrives for skill 1 every 1.4, 14.1, and 3.2 quarters in occupations 1, 2, and 3, respectively. The corresponding frequencies for θ_2 are 7.2, 1.3, and 1.7 quarters. Thus, occupation 1 provides the fastest learning opportunity for skill 1 and occupation 2 for skill 2, while occupation 3 provides learning opportunities for both skills.

Recall from equation (2.4) that $\gamma_{k\ell}$ represents the amount by which skill ℓ increases in occupation k. The estimated values of $\gamma_{k\ell}$ for $k = 1, 2, 3$ and $\ell = 1, 2$ also reveal differences in skill acquisition processes across occupations. Since the estimates of γ_{31} and γ_{12} are negligibly small, skills 1 and 2 hardly grow in occupation 3 and 1, respectively. Occupation 1 provides the best skill acquisition opportunities for skill 1. Occupation 2 provides the counterpart for skill 2, followed by occupation 3.

The right-hand side of the first and second rows of the same panel of Table 2.6 present the estimated parameters of the population distributions of skill acquisition probabilities. These estimates imply that the average values of skill acquisition probabilities are 0.18 and

[22]These values are equal to $\omega_{k1} + \omega_{k2}$ for $k = 1, 2, 3$.

Table 2.6: Parameter Estimates

	Skill (ℓ)			Skill (ℓ)	
Parameter	1	2	Parameter	1	2
$\gamma_{1\ell} \times 100$	8.7477	0.0000	p_ℓ	0.7605	9.3642
	(1.3834)	(0.8966)		(0.1720)	(3.1471)
$\gamma_{2\ell} \times 100$	0.4771	1.3016	q_ℓ	3.4022	9.4328
	(5.7166)	(0.1393)		(1.0956)	(3.2968)
$\gamma_{3\ell} \times 100$	0.0527	1.0482	$\underline{s}_\ell$	4.0105	2.7031
	(0.0211)	(0.0035)		(0.5600)	(0.5630)
$\omega_{1\ell}$	0.6666	0.1979	$\overline{s}_\ell$	5.3183	3.0550
	(0.0413)	(0.0364)		(0.5611)	(0.5672)
$\omega_{2\ell}$	0.0666	0.8163	m_ℓ	0.9707	0.3287
	(0.0124)	(0.0202)		(0.0108)	(0.0355)
$\omega_{3\ell}$	0.2958	0.6054			
	(0.0235)	(0.0279)			

	Occupation (k)				
Parameter	1	2	3	Parameter	
λ_k^0	0.0103	0.0277	0.0358	α	0.8126
	(0.0016)	(0.0043)	(0.0066)		(0.0201)
λ_k^1	0.0034	0.0137	0.0294	b	5.9594
	(0.0007)	(0.0026)	(0.0055)		(0.9832)
δ_k	0.0207	0.0246	0.0449	σ_ε	0.3027
	(0.0017)	(0.0013)	(0.0022)		(0.0015)
β_{k1}	0.7200	0.5889	1.0000†	π	0.5793
	(0.0963)	(0.0403)			(0.0265)
β_{k2}	1.5349	1.3512	1.0000†		
	(0.1449)	(0.0952)			
μ_k	0.1736	1.3873	0.0000†		
	(0.2590)	(0.3932)			
σ_k	0.2681	0.2380	0.2684		
	(0.0241)	(0.0113)	(0.0130)		

Loglikelihood value: -10794.60
Asymptotic standard errors in parentheses
† normalized values

0.50 for skills 1 and 2, respectively.[23] The distribution of the skill 1 acquisition probability is highly skewed to the right, with a median of 0.13. In contrast, the distribution of the skill 2 acquisition probability is symmetric and concentrated around the mean. The median of the distribution coincides with the mean, and the 5th and 95th percentiles of the distribution are 0.31 and 0.68. In the model, workers form their prior beliefs from the population distributions of the skill acquisition probabilities and their initial skill levels. Thus, these estimates indicate that there is a very low level of uncertainty for the skill 2 acquisition probability.

The estimates of $\underline{s}_1$ and $\overline{s}_1$ indicate a large gap between the upper and lower support points for the initial skill 1 level. This gap has substantial implications for wages. For example, a worker's first wage will be lower by 1.31 log point in occupation 3 if they are endowed with the lower initial skill 1 level than with the higher level.[24] The gap between the two support points for the initial skill 2 level distribution is narrower than for skill 1, but its implications for wages are still substantial. A worker's first wage will be lower by 0.35 log points in occupation 3 if they are endowed with the lower initial skill 2 level than otherwise. About one third of the sample population is initially endowed with the higher skill 2 level.

The first two rows of the lower panel of the table show differences in search frictions among the three occupations. Transitions to occupation 1 are hampered by the low values of λ_1^0 and λ_1^1. Occupation 3 presents the lowest degree of search frictions among the three occupations. For example, if unemployed workers exert the same job search effort level to all the three occupations, the likelihoods that a job offer comes from occupations 1, 2, and 3 are 0.140, 0.375, and 0.485, respectively.[25] The corresponding probabilities while

[23]Recall that the mean of a Beta distribution with parameters p_ℓ and q_ℓ is $p_\ell/(p_\ell + q_\ell)$.

[24]However, those endowed with the lower initial skill 1 level account for a small fraction of the sample population.

[25]The probability equals $\lambda_k^0(e)/(\lambda_1^0(e) + \lambda_2^0(e) + \lambda_3^0(e)) = \lambda_k^0/(\lambda_1^0 + \lambda_2^0 + \lambda_3^0)$ for each occupation k.

employed are 0.073, 0.295, and 0.632 for occupations 1, 2, and 3, respectively. However, the probability of exogenous job destruction is estimated to be the highest for occupation 3, far exceeding those for the other two occupations.

Estimated to be around 0.81, α indicates a high degree of diminishing returns to search efforts. Therefore, job search efforts are predicted to be spread out over different occupations rather than narrowly directed at a particular occupation. For example, about two-thirds of workers enter the labour market with initial skill levels $(s_1, s_2) = (5.31, 2.70)$, and are the largest group in the distribution of initial skill levels. The proportions of job search effort allocated to occupations 1, 2, and 3 by this group of workers are 0.19, 0.48, and 0.33, respectively.

The estimates of $\beta_{k\ell}$ show that occupation 1 provides higher returns to both skills 1 and 2 than does occupation 2. Curiously, occupation 3 provides the highest return to skill 1. As for the distributions of the employer-specific wage components, occupation 2 has the highest mean, and the large difference between μ_2 and μ_3 offsets occupation 3's advantage in return to skill 1. If evaluated at the mean values of the offer distributions, the initial wage is expected to be the lowest in occupation 3 for any possible pairs of the initial skill levels.

Overall, these estimates suggest that most workers value occupation 3 the least of the three occupations. Workers are expected to be paid less working in occupation 3 than in the other two occupations. Occupation 2 provides greater learning opportunities for both skills than occupation 3. Furthermore, workers face higher risk of exogenous job separation in occupation 3. While all of these differences give workers' less incentive to search for jobs in occupation 3, they are counterbalanced by the higher efficiency of search effort for occupation 3. The estimation results therefore suggest that jobs in occupation 3 serve as temporary stepping stones. The higher search efficiency for occupation 3 encourages transitions to this occupation from unemployment, but once workers are employed in occupation 3, they generally prefer to move to different occupations.

Table 2.7: Predicted Probability of Transitions from Jobs per Quarter

	1-10Q	11-20Q	21-40Q	41-60Q
Probability of making a	0.042	0.033	0.027	0.024
job-to-job transition	(0.063)	(0.044)	(0.043)	(0.031)
Probability of making an	0.023	0.017	0.014	0.012
occupation change	(0.026)	(0.021)	(0.015)	(0.011)
Probability of making a	0.031	0.029	0.028	0.026
job-to-unemployment transition	(0.053)	(0.042)	(0.022)	(0.017)

2.5.2 Model Fit

To examine the model's fit to the data, I generate a large sample of artificial work histories of workers using the estimated parameters, and compare the simulated data to the observed mobility patterns and wage profiles. For each simulated observation, I first sample skill acquisition probabilities and initial skill levels from their distributions among the worker population given by the parameter estimates. Then I record the predicted transition outcomes.

Table 2.7 presents the simulated counterpart of the transition probabilities in Table 2.4. The model matches the observed probabilities of job and occupational changes reasonably well. The model is able to generate the declining age profiles of these probabilities, though it underpredicts the job-to-job transition probabilities, and leaves a large portion of the declining profile of job-to-unemployment transitions unexplained.

Table 2.8 presents the predicted quarterly occupation transition matrices, together with the observed values given in parentheses. The predicted transition rates are close to the observed values in magnitude in most of the cells. A few mismatches are noted. First, the model overpredicts quarterly transitions from occupation 1 to unemployment because the estimate of δ_1 generates higher quarterly job separation rates than those observed. Second, while the diagonal elements for occupations 2 and 3 increase over time in the data, the

Table 2.8: Predicted Quarterly Occupational Transition Matrices

Occupation Origin	Period	UE	Occ 1	Occ 2	Occ 3
	1-10Q	0.711	0.035	0.132	0.122
		(0.721)	(0.023)	(0.107)	(0.145)
Unemployed	11-20Q	0.715	0.045	0.125	0.115
		(0.685)	(0.040)	(0.124)	(0.144)
(UE)	21-40Q	0.723	0.036	0.124	0.118
		(0.689)	(0.043)	(0.110)	(0.146)
	41-60Q	0.699	0.043	0.124	0.134
		(0.716)	(0.040)	(0.138)	(0.098)
	1-10Q	0.026	0.955	0.014	0.005
		(0.015)	(0.969)	(0.008)	(0.008)
Occupation 1	11-20Q	0.020	0.966	0.013	0.001
		(0.019)	(0.967)	(0.012)	(0.006)
(Occ 1)	21-40Q	0.021	0.967	0.010	0.002
		(0.011)	(0.973)	(0.010)	(0.004)
	41-60Q	0.020	0.970	0.009	0.001
		(0.014)	(0.976)	(0.005)	(0.006)
	1-10Q	0.023	0.003	0.971	0.002
		(0.051)	(0.006)	(0.925)	(0.017)
Occupation 2	11-20Q	0.026	0.004	0.969	0.001
		(0.045)	(0.007)	(0.938)	(0.008)
(Occ 2)	21-40Q	0.025	0.004	0.970	0.001
		(0.021)	(0.005)	(0.969)	(0.006)
	41-60Q	0.024	0.003	0.972	0.000
		(0.014)	(0.003)	(0.977)	(0.005)
	1-10Q	0.050	0.008	0.048	0.893
		(0.068)	(0.007)	(0.023)	(0.901)
Occupation 3	11-20Q	0.046	0.009	0.044	0.900
		(0.050)	(0.006)	(0.020)	(0.922)
(Occ 3)	21-40Q	0.046	0.008	0.042	0.903
		(0.032)	(0.005)	(0.015)	(0.950)
	41-60Q	0.045	0.011	0.047	0.898
		(0.022)	(0.006)	(0.009)	(0.965)

predicted values stay around 0.97 and 0.90, respectively. In addition, the model does not replicate the declines with age in transition rates to unemployment from occupations 2 and 3, respectively.

Figure 2.3 compares the observed and predicted fractions of those unemployed and working in the three occupations. The estimated model produces a good fit to the fractions of the unemployed and those working in occupation 1. If any, the predicted unemployment rate drops too fast initially, and stabilizes slightly above the observed level.

In contrast, the model has difficulty matching the shares of occupations 2 and 3 in the workers' employment, overpredicting occupation 2's share and underpredicting occupation 3's share. These discrepancies can be explained by the mismatches in the quarterly transition rates shown in Table 2.8. On the one hand, the inflow to occupation 2 from occupation 3 is overpredicted for the whole sample period, and this excess is not cancelled out by the predicted outflow from occupation 2. On the other hand, the predicted outflow from occupation 3 does not decline with age, resulting in the divergence between the predictions and data. Furthermore, the predicted transition to occupation 3 from the two other occupations is almost negligible. The excess outflow from occupation 3 to occupation 2 and the lack of inflow to occupation 3 from the other occupations reflect the point made earlier that occupation 3 acts as a transitory state in workers' movements from unemployment to the other occupations.

Finally, Figure 2.4 compares the observed and predicted log quarterly earnings profiles. The predicted earnings profile follows the observed counterpart closely, generating the wage growth overall, though it does not capture the observed concavity.

Overall, the model is able to match the data in several dimensions. In particular, it replicates the life-cycle patterns of job-to-job transition rates, the probability of changing occupations, and life-cycle wage growth. At the same time, however, the estimates have a tendency to miss the observed age effects in the occupational transition profiles.

Figure 2.3: Observed and Predicted Occupation Affiliations

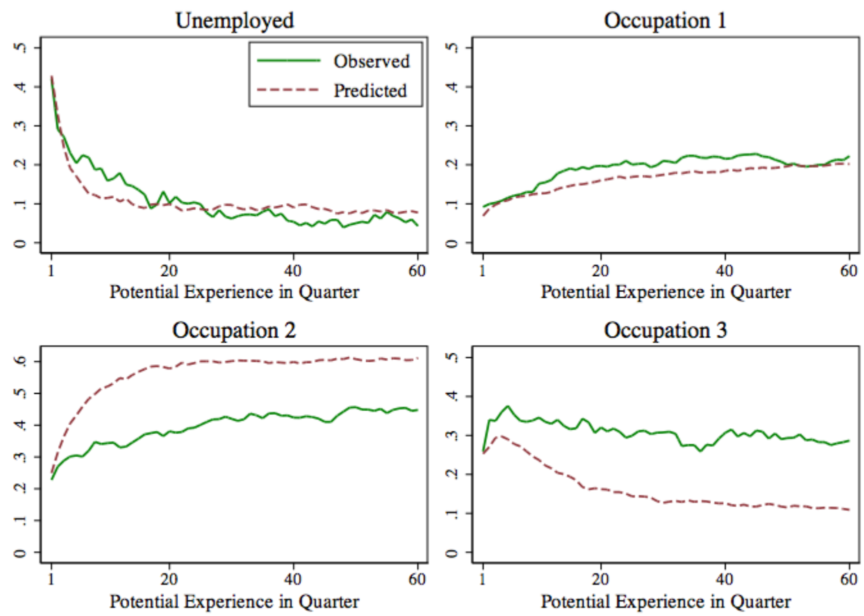
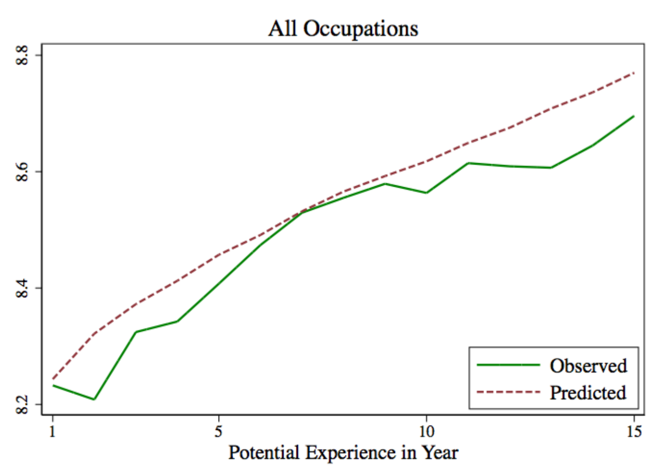


Figure 2.4: Observed and Predicted Log Wages



2.5.3 Augmented Model

The original model, called the baseline model from now on, substantially underpredicted occupation 3's employment share. Occupation 3 offers the least favourable opportunities in terms of both skill acquisition and earnings compared to the other occupations, so the transition rate out of occupation 3 is overpredicted. This discrepancy suggests that the baseline model is missing factors besides labour earnings and skill acquisitions that influence occupational choice. One way to ameliorate this is to model heterogeneity in workers' tastes for working in different occupations by allowing non-pecuniary utility to enter their labour market decision problem. Thus, I assume that a certain fraction of workers derive non-pecuniary utility from working in occupation 3. This is assumed to be constant among the group and over time, and denoted by z. This value is added to the pecuniary utility flow each period the workers work in occupation 3. The proportion of this group in the population is given by $\chi \in (0, 1)$. To facilitate discussions that follow, I call workers in this group type B workers, and the rest are called type A workers.

Moreover, the baseline model failed to capture declining transition probabilities from occupations 2 and 3 to unemployment with respect to age. Mirroring this problem, the probabilities of staying in occupations 2 and 3 did not rise. To address this problem, job destructions rates are allowed to depend on the potential experience of workers. Specifically, I assume that δ_k gives the exogenous job destruction probability in the first 20 quarters after a worker's labour market entry, δ'_k for the next 20 quarters, and δ''_k for the rest of his time in the labour market.[26]

Table 2.9 presents parameter estimates for the augmented model. Reflecting the observed transition patterns to unemployment, the estimated job destruction rates from occupations 2 and 3 exhibit decreasing trends with age. The decline is especially steep for occupation 2, with δ''_2 nearly a quarter of δ_2. The estimate of δ_1 is more in line with the observed quarterly transition rates from occupation 1 to unemployment in Table 2.5 than the value estimated in the baseline model.

[26] As shown in Table 2.5, there does not appear to be a trend in job separations from occupation 1 in terms of workers' ageing process. Therefore I restrict $\delta_1 = \delta'_1 = \delta''_1$.

With χ and z estimated to be 0.2620 and 1.6463, respectively, about one quarter of the population is estimated to be Type B workers. The non-pecuniary utility value they derive from working in occupation 3 is considerable; receiving the non-pecuniary utility is equivalent to valuing their labour earnings from occupation 3 five times higher.[27]

The augmented model provides the same qualitative implications as the baseline model in many respects. The distribution of the skill 1 acquisition probability is right-skewed with mean 0.18 and median 0.13. The distribution of the skill 2 acquisition probability is symmetric around mean 0.50, and tight around the mean with the 5th and 95th percentiles at 0.31 and 0.68, respectively. Occupation 1 provides the best skill acquisition and the fastest learning opportunities for skill 1, while occupation 2 is occupation 1's counterpart for skill 2. Occupation 3 presents the least search frictions. If unemployed workers exert the same level of job search effort to all three occupations, the likelihood that a job offer comes from occupation 3 is nearly 50% as opposed to 10% and 40% chances that a job offer is for employment in occupation 1 and occupation 2, respectively. With α estimated to be 0.8190, the estimated job offer arrival probabilities as a function of search efforts exhibit a high degree of diminishing returns.

Table 2.10 presents predicted transition probabilities between jobs, between occupations, and from job to unemployment for the augmented model. Compared with the baseline model's predictions, the augmented model fits the data better, especially for the quarterly job-to-unemployment transition probabilities. Table 2.11 presents the predicted quarterly occupation transition matrices from the augmented model. With the age-varying job destruction rates, predicted transitions to unemployment from occupations 2 and 3 decline with age. This also allows the model to predict that the likelihood of staying in occupations 2 or 3 rises with age. By allowing for workers with different tastes, the model is able

[27]Recall that to type B workers, the per-period utility from working in occupation 3 is the sum of the logarithm of labour earnings and the non-pecuniary utility term.

Table 2.9: Parameter Estimates – Augmented Model

	Skill (ℓ)			Skill (ℓ)	
Parameter	1	2	Parameter	1	2
$\gamma_{1\ell} \times 100$	9.4069	0.0000	p_ℓ	0.7425	9.3819
	(1.6735)	(1.2217)		(0.1880)	(3.2944)
$\gamma_{2\ell} \times 100$	0.3295	1.2726	q_ℓ	3.4546	9.4270
	(10.3300)	(0.1474)		(1.2076)	(3.4183)
$\gamma_{3\ell} \times 100$	0.0408	0.7682	$\underline{s}_\ell$	4.0183	2.6936
	(0.0147)	(0.0037)		(0.3265)	(0.3146)
$\omega_{1\ell}$	0.6665	0.1611	$\overline{s}_\ell$	5.3359	3.0259
	(0.0423)	(0.0335)		(0.3146)	(0.3161)
$\omega_{2\ell}$	0.0459	0.8480	m_ℓ	0.9736	0.3613
	(0.0107)	(0.0189)		(0.0109)	(0.0371)
$\omega_{3\ell}$	0.3947	0.4890			
	(0.0256)	(0.0285)			

	Occupation (k)				
Parameter	1	2	3	Parameter	
λ_k^0	0.0097	0.0286	0.0364	α	0.8190
	(0.0008)	(0.0018)	(0.0025)		(0.0059)
λ_k^1	0.0036	0.0201	0.0215	b	6.4576
	(0.0003)	(0.0016)	(0.0001)		(0.4996)
δ_k	0.0110	0.0474	0.0547	σ_ε	0.3047
	(0.0015)	(0.0036)	(0.0040)		(0.0015)
$\delta_k^{'}$	0.0110	0.0173	0.0392	π	0.5751
	(0.0015)	(0.0024)	(0.0047)		(0.0269)
$\delta_k^{''}$	0.0110	0.0130	0.0369	z	1.6463
	(0.0015)	(0.0021)	(0.0041)		(0.1661)
β_{k1}	0.7203	0.5898	1.0000†	χ	0.2605
	(0.0932)	(0.0507)			(0.0443)
β_{k2}	1.5213	1.3469	1.0000†		
	(0.1461)	(0.0870)			
μ_k	0.2333	1.3568	0.0000†		
	(0.2922)	(0.1964)			
σ_k	0.2827	0.2141	0.2612		
	(0.0225)	(0.0112)	(0.0114)		

Loglikelihood value: -10694.68

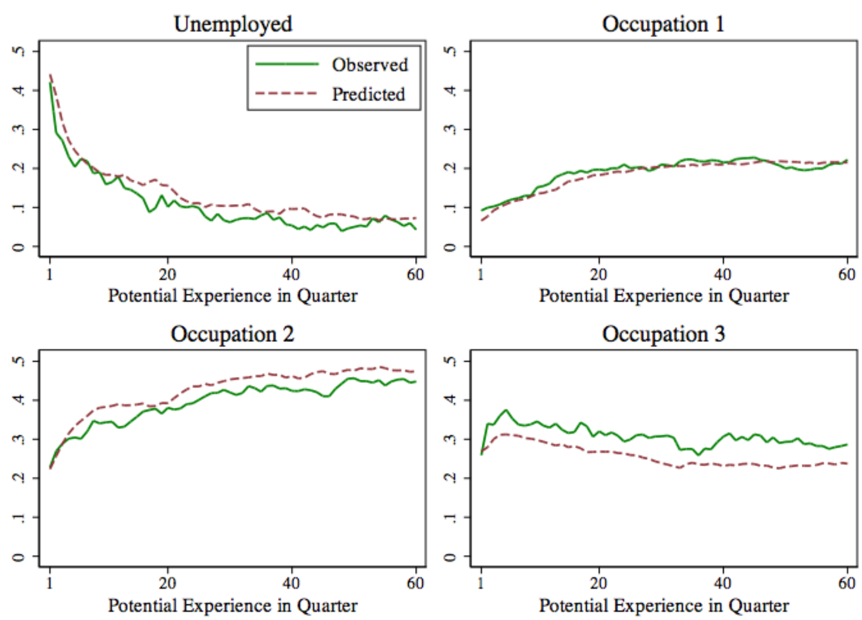
Asymptotic standard errors in parentheses

† normalized values

Table 2.10: Predicted Probability of Transitions from Jobs per Quarter

	1-10Q	11-20Q	21-40Q	41-60Q
Probability of making a	0.042	0.036	0.029	0.023
job-to-job transition	(0.063)	(0.044)	(0.043)	(0.031)
Probability of making an	0.019	0.016	0.012	0.007
occupation change	(0.026)	(0.021)	(0.015)	(0.011)
Probability of making a	0.046	0.041	0.024	0.019
job-to-unemployment transition	(0.053)	(0.042)	(0.022)	(0.017)

Figure 2.5: Observed and Predicted Occupational Choice – Augmented Model



to better match the outflow from occupation 3 to occupation 2 as well as the inflows to occupation 3 from the other occupations.

These improvements also produce a better fit to the proportions of workers working in the three occupations, as shown in Figure 2.5. The baseline model's large mismatches between the observed and predicted proportions of workers in occupations 2 and 3 are narrowed. The predicted fraction of workers unemployed follows the observed values closely.

Table 2.11: Predicted Quarterly Occupational Transition Matrices

Occupation Origin	Period	UE	Occ 1	Occ 2	Occ 3
		Destination			
	1-10Q	0.757	0.029	0.095	0.118
		(0.721)	(0.023)	(0.107)	(0.145)
Unemployed	11-20Q	0.770	0.032	0.092	0.105
		(0.685)	(0.040)	(0.124)	(0.144)
(UE)	21-40Q	0.773	0.029	0.087	0.112
		(0.689)	(0.043)	(0.110)	(0.146)
	41-60Q	0.764	0.023	0.086	0.127
		(0.716)	(0.040)	(0.138)	(0.098)
	1-10Q	0.023	0.961	0.008	0.007
		(0.015)	(0.969)	(0.008)	(0.008)
Occupation 1	11-20Q	0.012	0.973	0.012	0.003
		(0.019)	(0.967)	(0.012)	(0.006)
(Occ 1)	21-40Q	0.010	0.975	0.012	0.002
		(0.011)	(0.973)	(0.010)	(0.004)
	41-60Q	0.011	0.981	0.007	0.002
		(0.014)	(0.976)	(0.005)	(0.006)
	1-10Q	0.055	0.007	0.932	0.006
		(0.051)	(0.006)	(0.925)	(0.017)
Occupation 2	11-20Q	0.047	0.006	0.945	0.003
		(0.045)	(0.007)	(0.938)	(0.008)
(Occ 2)	21-40Q	0.019	0.006	0.973	0.002
		(0.021)	(0.005)	(0.969)	(0.006)
	41-60Q	0.013	0.004	0.982	0.001
		(0.014)	(0.003)	(0.977)	(0.005)
	1-10Q	0.063	0.004	0.040	0.893
		(0.068)	(0.007)	(0.023)	(0.901)
Occupation 3	11-20Q	0.051	0.006	0.022	0.922
		(0.050)	(0.006)	(0.020)	(0.922)
(Occ 3)	21-40Q	0.044	0.004	0.013	0.939
		(0.032)	(0.005)	(0.015)	(0.950)
	41-60Q	0.037	0.002	0.007	0.954
		(0.022)	(0.006)	(0.009)	(0.965)

Figure 2.6: Predicted Occupational Choice by Diiferent Utility Type

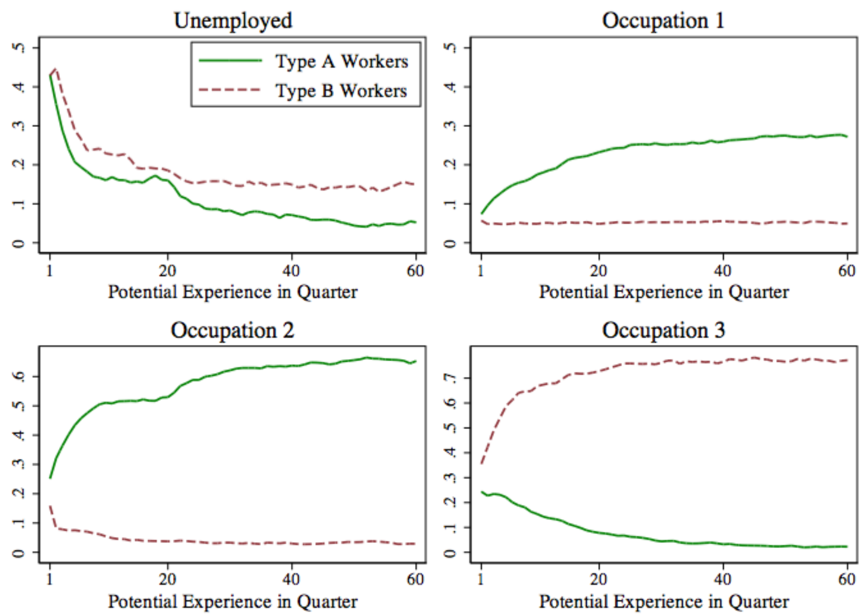
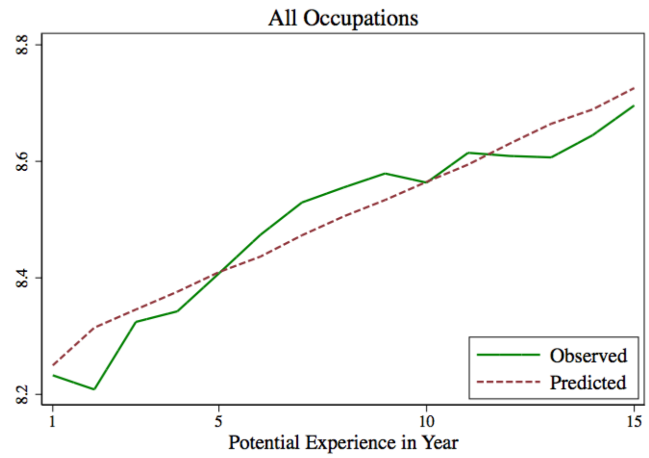


Figure 2.6 presents the distributions of occupational affiliation of type A and type B workers at given potential experience levels. Naturally, occupation 3 is a dominant choice among type B workers because it gives them the large non-pecuniary utility. More than 70% of this group are employed in occupation 3 within 20 quarters of their labour market entry. Because occupation 3 presents the highest unemployment risk, this group has a higher unemployment rate than the other group.

Figure 2.7 shows predicted earnings profiles from the augmented model. The fit appears good overall, but still is not able to produce a concave profile.

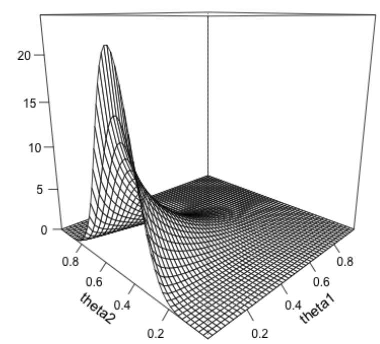
Figure 2.7: Predicted Log Wage



2.6 The Effect of Learning on Job Search Effort Decisions and Occupational Affiliation Outcomes

With the estimated augmented model fitting the data reasonably well, I utilize these estimates and perform several simulation exercises to examine how learning influences workers' job search decisions and occupational transition outcomes. I start with analyzing the effect of learning on workers' job search effort decisions, followed by an analysis of the effect of learning on their occupational transitions. With job search frictions playing a role in workers' transitions, changes in job search effort decisions induced by learning do not necessarily translate into comparable changes in occupational transition outcomes. Thus, how the estimated job search frictions affect learning and its impact on transition outcomes are also examined.

Figure 2.8: Joint Probability Density of θ_1 and θ_2



2.6.1 The Effect of Learning on Job Search Effort Decisions

Before describing how learning affects workers, I first present the initial state of workers' beliefs implied by the parameter estimates. The model envisions an environment where workers, initially uncertain about their skill acquisition probabilities, form their beliefs from the population distributions of these probabilities. Figure 2.8 plots the joint probability density of θ_1 and θ_2 among workers. Since the average value of θ_1 among workers is around 0.18, workers initially believe that the likelihood that their skill 1 levels will increase is not particularly high. In truth, the high positive skew of the distribution with respect to θ_1 means that a majority of workers' skill 1 acquisition probabilities are even lower than the mean of their prior beliefs. Moreover, since the distribution of skill acquisition probabilities is more spread out over θ_1 than θ_2, the initial beliefs provides more uncertainty for θ_1 than θ_2.

This initial belief distribution forms a basis for workers' job search decisions at labour market entry, and workers decide how much effort to exert in searching in different occupations. Table 2.12 presents the proportions of search effort exerted to the three occupations

Table 2.12: Predicted Job Search Effort Allocations and Distributions of Job Offered in the 2nd Quarter

	Search effort allocation		
Current Occupation	Occ 1	Occ 2	Occ3
Unemployed	0.280	0.417	0.303
Occupation 1	1.000	0.000	0.000
Occupation 2	0.755	0.240	0.005
Occupation 3	0.349	0.563	0.088

For workers with $(s_1^0, s_2^0) = (5.34, 2.69)$.

by workers a quarter after their labour market entry.[28] Search effort is spread out among the three occupations while workers are unemployed, with approximately two fifths of their effort allocated to occupation 2, and the rest almost evenly divided for occupations 1 and 3. Only a small fraction of job search effort is allocated to occupation 3 once the workers leave unemployment.[29] In occupations 1 and 2, effort is primarily devoted to occupation 1 and hardly to occupation 3. While working in occupation 3, approximately 60 percent of job search effort is devoted to find employment in occupation 2, 30 percent to occupation 1, and 10 percent to occupation 3.

Next, I examine how new information alters workers' job search effort allocations. Table 2.13 shows how the fraction of job search effort allocated to each occupation changes after one quarter of employment in occupation 3.[30] For example, the row labeled "No signal" refers to the case where workers encounter no skill acquisition opportunity, and the next row, labeled "Signal for skill 1, success," corresponds to the case where workers en-

[28] Throughout this section, I focus on a group of type A workers with initial skill levels of $(s_1, s_2) = (5.34, 2.69)$, which accounts for the largest fraction of workers in terms of initial conditions. The results for the other types of workers are given in Appendix C. General conclusions presented in this section hold for the other initial skill level configurations.

[29] The predicted values for employed workers assume that their employer-specific wage components are at the 25th percentile of the estimated distribution.

[30] Predicted changes in job search effort in occupation 2 are in general similar to those predicted for occupation 3. In contrast, effort allocations exerted in occupation 1 remain highly concentrated on occupation 1 after any possible learning outcomes in quarter 2.

Table 2.13: Predicted Changes in Job Search Effort Allocations and Distribution of Offered Jobs in the 3rd Quarter

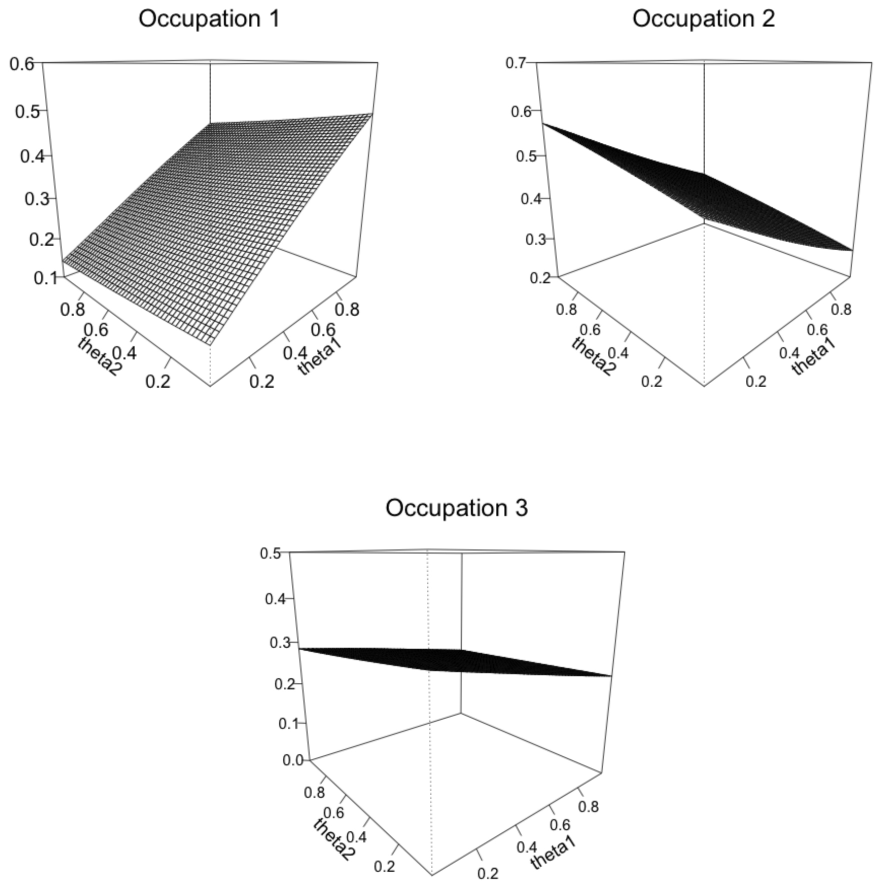
	Change in Search effort allocation					
	Occupation 3			Unemployed		
Skill dynamics scenario in $t = 2$	Occ 1	Occ 2	Occ 3	Occ 1	Occ 2	Occ 3
No signal	−0.004	0.005	0.000	−0.002	0.002	0.000
Signal for skill 1, success	0.150	−0.128	−0.021	0.129	−0.064	−0.064
Signal for skill 1, failure	−0.047	0.042	0.005	−0.032	0.017	0.015
Signal for skill 2, success	−0.015	0.017	−0.002	−0.005	0.007	−0.002
Signal for skill 2, failure	0.006	−0.008	0.002	0.002	−0.003	0.001

counter a skill 1 acquisition opportunity and their skill 1 levels have increased. The three columns under the label "Occupation 3" present the search effort allocations exerted by workers if they stay employed in occupation 3, while the three rightmost columns report predicted search effort allocations searching for a job while unemployed.

While no new information resulting in negligible changes in search effort allocations from the previous period is expected, new signals about skill 2 acquisition probabilities also produce only small changes. In contrast, a large change in search effort allocation is predicted in response to an increase in the skill 1 level, with workers reallocating their effort toward occupation 1, as well as a moderate change in job search effort away from occupation 1 if a skill acquisition opportunity for skill 1 fails to produce an increase. These heterogeneous responses stem from the shapes of the initial belief distributions. The workers' initial beliefs indicate that the odds of skill 1 acquisition is low. Therefore, an increase in the skill 1 level acts as a signal dramatically updating their beliefs. The initial belief distribution for θ_2 is symmetric, with the density highly concentrated around the mean. New information about θ_2 does not substantially alter their beliefs, and therefore generates only a small response.

Having examined the effect of a single signal, I now turn to how workers change their job search behaviour as learning progresses over time. To this end, I simulate a 10 year spell of labour market transitions since labour market entry for a large number of workers,

Figure 2.9: Predicted Search Effort Allocations in the 41st Quarter



and predict how the progress of learning during this period alters their job search effort allocations. Figure 2.9 presents the predicted fraction of job search effort that workers with given skill acquisition probabilities allocate to each of the three occupations in the 41st quarter if they are unemployed.

This simulation exercise shows that 10 years of labour market experience can change their job search behaviour sizeably. Workers with high skill 1 acquisition probabilities are predicted to allocate nearly half of their job search effort toward occupation 1 as opposed to the nearly one third they allocated at their labour market entry. Workers with low skill

1 acquisition probabilities change their effort allocation in the opposite direction, putting more than half of their effort to search for a job in occupation 2.

Figure 2.9 also shows that there is a weaker pattern between job search effort allocations and θ_2, such that skill 2 acquisition probabilities affect their job search effort decisions less than θ_1. This also reflects the earlier result that a signal about θ_2 has a much smaller effect on job search effort allocations than a signal about θ_1. In sum, the analysis indicates that learning about θ_1 affects workers far more than learning about θ_2, and learning largely affects the allocation of job search effort between occupations 1 and 2, with the fraction of effort allocated to occupation 3 largely unaffected.

2.6.2 The Effect of Learning on Transition Outcomes

Due to search frictions, job search effort is only a part of the equation determining occupational transition outcomes of workers. How learning affects the latter therefore needs to be examined separately. To this end, I first simulate labour market transitions for a large number of workers whose true acquisition probabilities are drawn from the population distributions of θ_1 and θ_2, and predict their occupational affiliation distribution after 10 years of (potential) experience for this simulated sample. Then, I compare the result with a counterfactual occupational affiliation distribution for the same period simulated under a scenario where learning is shut down. Specifically, I consider a scenario where workers do not update their initial belief distributions of θ_1 and θ_2 despite receiving new information, thus holding on to their initial beliefs.

Table 2.14 presents the simulated distributions of employed workers in the three occupations under the two scenarios with and without learning, respectively. Comparing these two distributions shows that learning increases the likelihood of working in occupation 2, with this difference offset by a comparable decrease in the proportion of employed workers

Table 2.14: Predited Occupational Choice Distributions in the 41st Period

	Occupation 1	Occupation 2	Occupation 3
With learning	0.277	0.624	0.099
Without learning	0.306	0.598	0.096
Difference	−0.029	0.026	0.003

in occupation 1. However, the differences between the two distributions are small even if a 10 year period of labour market experience is considered.

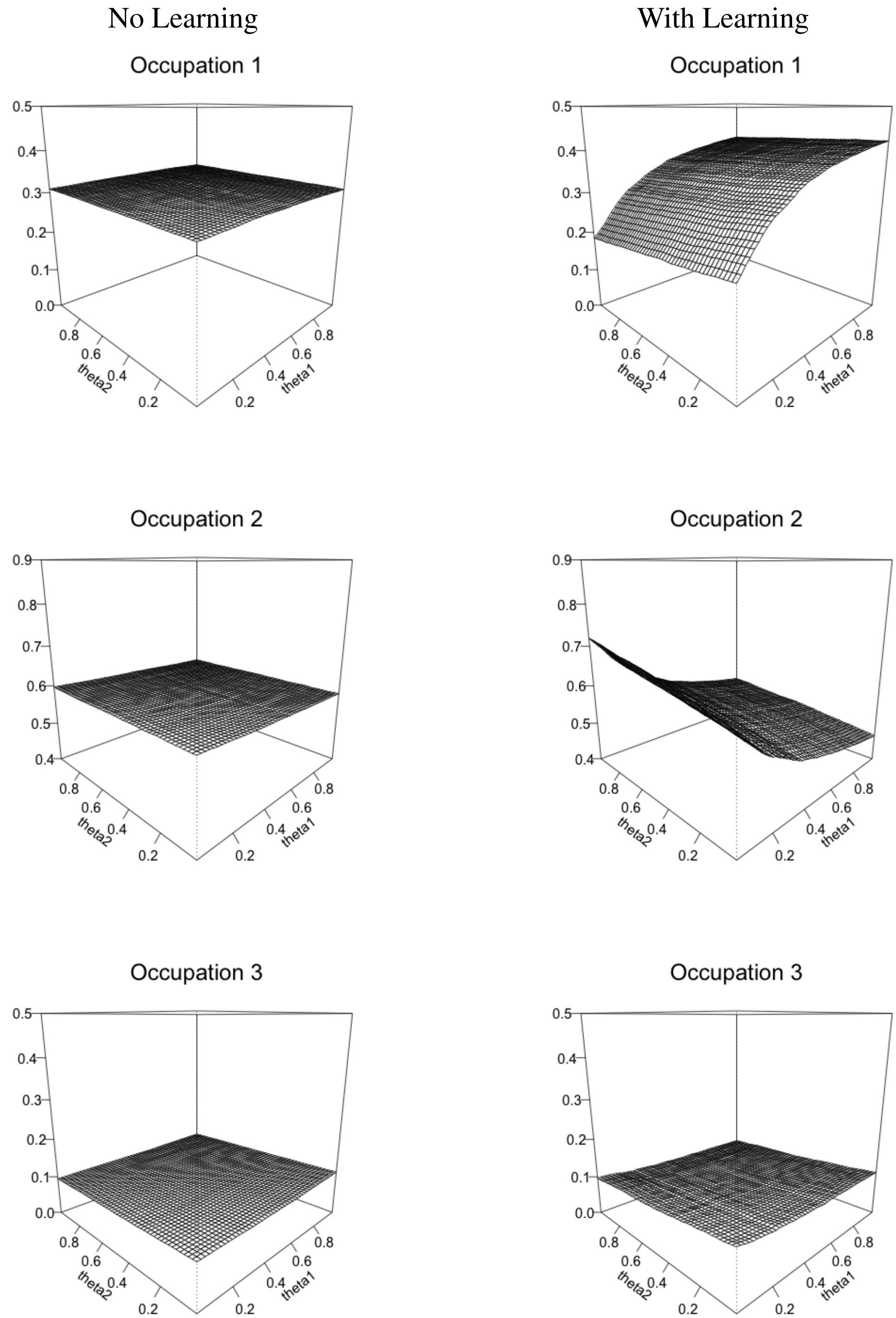
This small difference is partly due to the composition of workers with different skill acquisition probabilities in the simulated sample. With workers of different skill acquisition probabilities affected by the learning process differently, their changes can offset each other to some degree. Indeed, it is more illumining to analyze the effect at a more disaggregated level.

Therefore, I simulate occupational affiliation distributions for samples of workers with the same skill acquisition probabilities, and present the results in Figure 2.10. The three surface graphs on the left-hand side of the figure show that the fractions of workers working in given occupations in the 41st quarter since their labour market entry under the counter-factual no-learning scenario. All the graphs are flat, indicating that without learning, the occupational affiliation distribution vary little with skill acquisition probabilities.

The three graphs on the right-hand side of Figure 2.10 present the simulated occupational affiliation distribution in the 41st quarter with learning. Following the way learning influences workers' job search effort decisions presented in Figure 2.9, the proportion working in occupation 1 increases with θ_1, and this change is largely offset by a fall in the proportion working in occupation 2. There is a weaker pattern between occupational choice and θ_2.

Figure 2.9 showed that workers with high θ_1 learn to focus their job search effort on

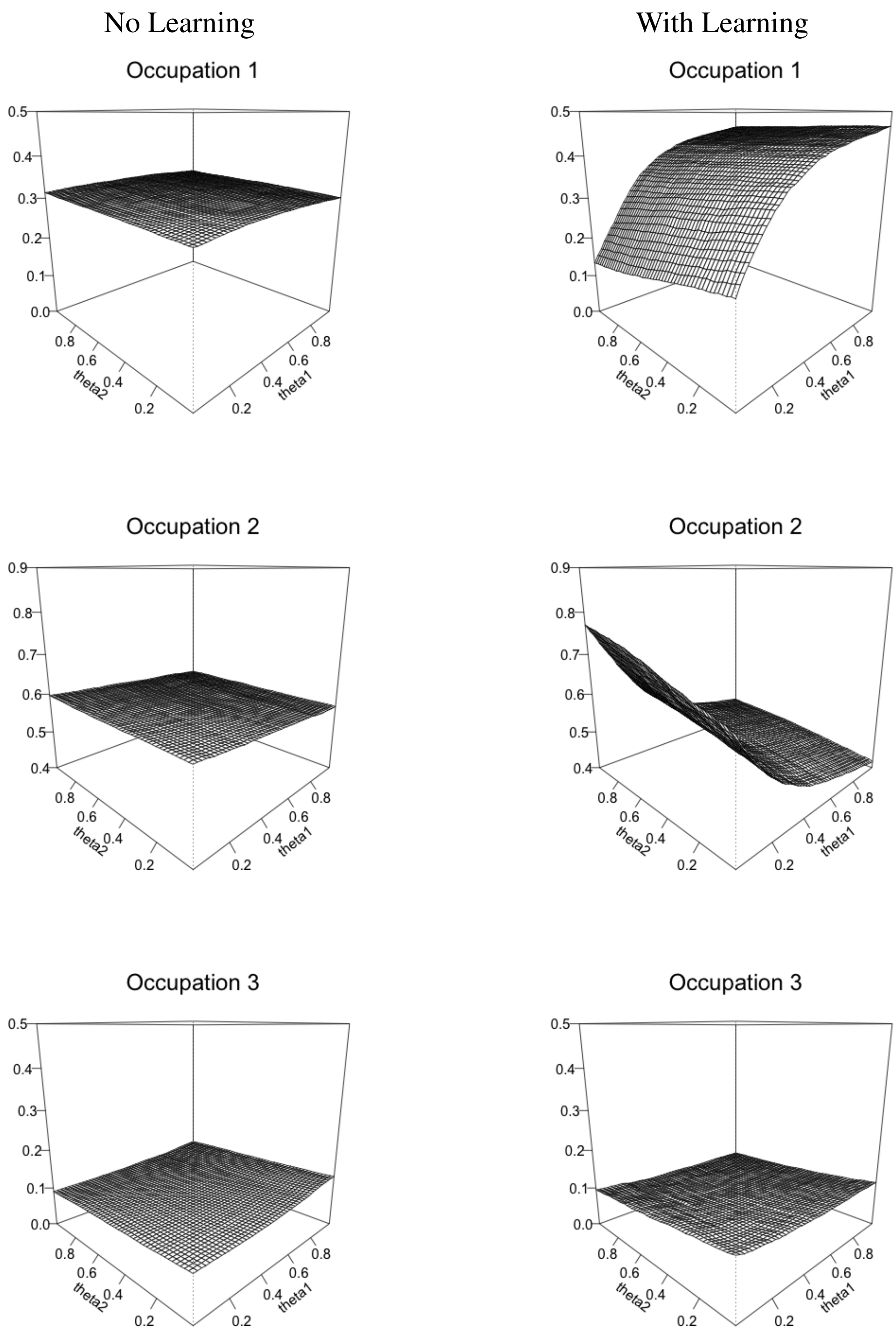
Figure 2.10: Predicted Occupational Affiliations in the 41st Quarter



occupation 1. However, such effort faces high search frictions associated with this occupation, and as a result, their occupational affiliation outcomes do not match their job search effort. Around two fifth of them are employed in occupation 1, with a majority of them are in occupation 2 instead. Because of search frictions, their occupational affiliation outcomes do not reflect the extent of the changes in job search effort allocations. In contrast, the search frictions are less of a concern for workers with low θ_1, because they direct their job search effort away from occupation 1 toward occupation 2 as learning progresses. They indeed attain occupational affiliation where more than three fifths of them are employed in occupation 2. The search frictions associated with occupation 2 are substantially lower than those associated with occupation 1, and job search effort to find work in occupation 2 is more likely to be successful in generating transitions to this occupation.

Not only does the low likelihood of finding a job in occupation 1 substantially affect occupational transitions of workers with high skill 1 acquisition probabilities, it also influences learning in a different way. Specifically, it slows down workers' learning about θ_1 because the likelihood of receiving a signal about θ_1 is the highest in occupation 1. This effect can be seen by simulating labour market transitions where the arrival of signal is made more frequent. For example, I consider an environment where signals for both θ_1 and θ_2 arrive every period when workers are employed in any occupation so that the difficulty to find work in occupation 1 does not hinder the learning process. Under this faster learning process, I simulate the occupational affiliation distributions in the 41st quarter, and compare them with the counterfactual distribution without learning. These distributions are presented in Figure 2.11. Comparing the distributions presented in Figure 2.11 with those in Figure 2.10 shows that the faster learning process makes larger differences in the workers' labour market outcomes. With the faster learning process, workers whose θ_1 are at either end of the skill 1 acquisition probability distribution attain more transitions to the occupations targeted by their job search effort. With more signals, the learning process

Figure 2.11: Predicted Occupational Affiliations in the 41st Quarter with the Faster Learning Process



accelerates changes in workers' job search effort decisions, which in turn generates larger differences in occupational transitions outcomes.

2.7 Concluding Remarks

This chapter develops and estimates a dynamic model in which skill acquisition, learning, and job search frictions influence workers' mobility decisions, and uses the estimates to measure the extent to which learning behaviour affects the workers' job search behaviour and transitions. The model is estimated using a sample of white male high school graduates in the U.S. The estimated parameters show large differences among occupations in search frictions, and skill acquisition and learning opportunities. Occupation 3, which mainly consists of low skill jobs and accounts for the largest share of the workers' employment immediately after their labour market entry, is associated with the least search frictions of the three occupations, but provides workers with the least favourable opportunities for human capital growth. The other two occupations, representing skilled white-collar and blue-collar occupations, provide favourable skill acquisition and learning environment. However, transition to the white-collar occupation is infrequent because of high search frictions.

Simulation exercises show that while workers initially spread out their job search effort across occupations, learning about the white-collar related skill acquisition can have a sizeable effect on their job search effort over time. A signal regarding the blue-collar skill, however, has a very small effect on workers' job search effort choice.

Because of job search frictions, changes in job search effort due to learning do not result in a comparable effect in occupational transition outcomes. Moreover, the high search frictions associated with the white-collar occupation slow down learning by limiting opportunities to acquire signals about the white-collar skill acquisition. This diminishes the effect of learning on workers' early labour market outcomes.

Bibliography

Antonovics, Kate, and Limor Golan. 2012. Experimentation and job choice. *Journal of Labor Economics* 30(2):333–66.

Bowlus, Audra J., and Huju Liu. 2011. The contribution of search and human capital in lifetime earnings growth. Workin paper.

Burdett, Kenneth, Carlos Carrillo-Tudela, and Melvyn G. Coles. 2011. Human capital accumulation and labor market equilibrium. *International Economic Review* 52(3):657–77.

Christensen, Bent Jesper, Rasmus Lentz, Dale T. Mortensen, George R. Neumann, and Axel Werwatz. 2005. On-the-job search and the wage distribution. *Journal of Labor Economics* 23(1):31–57.

DeGroot, Morris H. 1970. *Optimal Statistical Decisions*. McGraw-Hill.

Fraley, Chris, and Adrian E. Raftery. 2002. Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association* 97(458):611–31.

Gathmann, Christina, and Schönberg. 2010. How general is human capital? a task-based approach. *Journal of Labor Economics* 28(1):1–49.

Heckman, James J., and Guilherme Sedlacek. 1985. Heterogeneity, aggregation, and market wage functions: an empirical model of self-selection in the labor market. *Journal of Political Economy* 93(6):1077–1125.

Keane, Michael P, and Kenneth I. Wolpin. 1994. The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: monte carlo evidence. *The Review of Economics and Statistics* 76(4):648–72.

———. 1997. The career decisions of young men. *Journal of Political Economy* 105(3):473–522.

———. 2001. The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review* 42(4):1051–1103.

Light, Audrey, and Kathleen McGarry. 1998. Job change patterns and the wages of young men. *Review of Ecconomics and Statistics* 80(2):276–86.

McCall, Brian P. 1990. Occupational matching: a test of sorts. *Journal of Political Economy* 98(1):45–69.

Miller, Robert A. 1984. Job matching and occupational choice. *Journal of Political Economy* 92(6):1086–1120.

Neal, Derek. 1999. The complexity of job mobility among young men. *Journal of Labor Economics* 17(2):237–261.

Papageorgiou, Theodore. 2014. Learning your comparative advantages. *Review of Economic Studies* Forthcoming.

Pavan, Ronni. 2010. The role of career choice in understanding job mobility. *LABOUR* 24(2):107–27.

———. 2011. Career choice and wage growth. *Journal of Labor Economics* 29(3):549–87.

Poletaev, Maxim, and Chris Robinson. 2008. Human capital specificity: evidence from the dictionary of occupational titles and displaced worker surveys. *Journal of Labor Economics* 26(3):387–420.

Robinson, Chris. 2011. Occupational mobility, occupation distance, and specific human capital. Working paper.

Roy, A. D. 1951. Some thoughts on the distributions of earnings. *Oxford Economic Papers* 3(2):135–146.

Rubinstein, Yona, and Yoram Weiss. 2006. Post schooling wage growth: investment, search and learning. In *Handbook of the Economics of Education*, eds. Erik Hanushek and F. Welch, vol. 1, chap. 1, 1–67. Elsevier.

Sanders, Carl. 2012. Skill uncertainty, skill accumulation, and occupational choice. Working paper.

Sullivan, Paul. 2010. A dynamic analysis of educational attainment, occupational choices, and job search. *International Economic Review* 51(1):289–317.

Topel, Robert H., and Michael P. Ward. 1992. Job mobility and the careers of young men. *Quarterly Journal of Economics* 107(2):439–479.

Wolpin, Kenneth I. 1992. The determinant of black-white differences in early employment careers: search, layoffs, quits, and endogenous wage growth. *Journal of Political Economy* 100(3):535–560.

Yamaguchi, Shintaro. 2012. Tasks and heterogeneous human capital. *Journal of Labor Economics* 30(1):1–53.

Chapter 3

Motherhood, Part-time Work, and Skill Dynamics of British Women

3.1 Introduction

Generally, women take primary responsibility for childcare. Among different alternatives used to meet this responsibility, part-time employment is frequently chosen by mothers in many countries, which some argue enables them to balance their work and family lives. The popularity of part-time employment among mothers is widely documented in Britain. While most women in Britain start their career with full-time work, many of them make transitions to nonemployment or part-time employment upon the arrival of a child.[1] These transitions are, however, unlikely to be reversed later in their working lives with part-time work becoming prevalent after the birth of their children. Such moves are common even with the large wage differentials between full-time and part-time work widely reported in

[1]Percentage of women with no children who work full-time was 84.9 in study using BHPS by Paull (2008). Using the British Household Panel Survey (BHPS) between 1991 and 2004, Paull (2008) found that 43% of women who worked full-time before the birth of their first child switched to part-time jobs, while about 30% of them continued to work full-time.

the literature.[2]

With fewer hours worked and a lower wage rate, a substantial earnings loss is likely to follow when individuals transition from full-time to part-time employment. Thus, there appears to be a potent disincentive against part-time work, especially for mothers who need to arrange paid childcare. Therefore, the part-time wage penalty and prevalence of part-time work are difficult observations to explain simultaneously in standard labour supply models based on leisure-income trade-off. One view consistent with these two observations is to assume that motherhood induces a permanent preference shock changing workers' income-leisure tradeoff in favour of leisure. While plausible, this view has the limitation that the change should favour non-participation over part-time work.

An alternative explanation for the pervasive part-time employment among mothers is skill dynamics during career interruptions. First, the reason to remain attached to the labour market even when their family concerns are important may reflect their desire not to let their human capital atrophy and to maintain their earning capacities. Second, though it is customary to associate a spell of nonemployment with human capital depreciation, such risk may also be present even while workers are employed. Research has provided evidence that the pay gap between full-time and part-time work is attributable to the occupational segregation between them.[3] As a result, when workers switch to part-time work, they may end up accepting positions that do not require the skills used in their previous employment.[4] Mismatch of worker's skill and chosen occupation might lead to skill depreciation, causing workers difficulties in returning to the previous occupation later. When limited to job opportunities in less skilled occupations and therefore lower earnings potential, they may find it optimal to stay in part-time employment given their leisure-income tradeoff.

[2]Manning and Petrongolo (2008), for example, reports that the average hourly earnings of women working part-time was 26% below women working full-time in 2001.

[3]See, for example, Connolly and Gregory (2008); Manning and Petrongolo (2008); Hirsh (2005).

[4]Connolly and Gregory (2008) found that 40% of the women who changed employers to reduce their work hours experienced the occupational downgrading.

The extent to which skill depreciation may account for the persistent part-time employment among mothers hinges on whether they voluntarily make such moves in the first place despite the adverse effect on their lifetime earnings potentials. Naturally, the answer is likely to depend on the benefits and costs of such moves over their lifetimes, which are not only determined by the process of skill dynamics, but also by a variety of factors such as the wage structure, workers' income-leisure tradeoff and expectations of the lengths of childcare responsibilities and their labour market opportunities afterward. This paper develops a dynamic model of employment transitions between full-time, part-time and non-participation incorporating these aspects in order to capture the potential role of skill depreciation in British mothers' employment pattern.

The model is based on the search-matching model of the labour market in Albrecht and Vroman (2002). Their model incorporates differences in skill levels among workers and skill requirements among jobs. Due to these differences, the workers can only be matched with jobs for which they are qualified. Specifically, in the model, high-skilled and low-skilled workers, and employers looking to fill high-skill or low-skill job vacancies search for the other party to form a productive match. The high skilled workers can meet the skill requirements of both jobs, while the low-skilled workers can only satisfy the low-skill jobs' skill requirement. The current paper extends this modelling framework by adding hours of work as another job characteristic, and lets workers search for jobs in terms of hours of work as well as skill requirement. Search frictions are used in this paper to account for the scarcity of part-time jobs in the high-skill occupation. Because of the frictions, part-time job offers are expected to be predominantly low-skill jobs.

In addition, I add to the worker's problem two more dynamic aspects. The first comes from changes in family circumstances due to the arrival of a child, which changes the pecuniary returns from work in the labour market and potentially changes workers' preferred hours of work. The second is a skill depreciation process, which changes workers' ability to meet the skill requirements of different jobs. While family concerns are assumed to change exogenously, skill dynamics depend on workers' employment decisions, giving them a certain control over their skill levels.

After formulating the workers' problem and describing their transition behaviour, a series of numerical examples of the model are given to present model predictions under different parameter configurations. The goal of this exercise is twofold. First, using numerical examples, I illustrate the general properties of the workers' problem. Second, I examine whether the model can capture the stylized facts of the observed life-cycle employment transition patterns of British mothers. For this, the model needs to produce a prediction that workers are willing to downgrade their occupations in order to switch from full-time work to part-time work when childcare needs arise, and then prefer to stay employed part time should skill depreciation prevent them from reversing their occupational changes. The exercise finds a set of reasonable parameter values that produces this result.

The rest of the paper is organized as follows. In the next section, I present the model. In Section 3.3, I examine the problem of individual workers. In Section 3.4, I present the numerical illustration of the model. Concluding remarks follow.

3.2 Model

In the model, time is discrete, and all exogenous events occur at constant arrival probabilities. There are two occupations in the economy: a high-skill occupation and a low-skill occupation. Jobs are available for both occupations on a full-time and a part-time basis. Thus, four types of jobs can be offered to workers: high-skill full-time (HF), high-skill part-time (HP), low-skill full-time (LF), and low-skill part-time job (LP). Each period workers are employed, each job type, respectively, pays them the following amount: y_{HF}, y_{HP}, y_{LF}, and y_{LP}. I assume $y_{HF} > y_{HP}$, $y_{LF} > y_{LP}$, and $y_{HF} > y_{LF}$. The amount of hours required by full-time and part-time jobs are denoted by h_{FT} and h_{PT}, where $h_{FT} > h_{PT}$, respectively.

Workers are heterogeneous with respect to their valuation of leisure. I treat a worker's valuation of leisure as her type. The type is represented by a real number θ, and the valuation of leisure relative to income is assumed to be increasing in θ. Each worker has a time endowment of H in each period, and divides it between market work and leisure. Given the types of jobs that can be offered to workers, choices of time allocation are limited to the following: full-time work (h_{FT}), part-time work (h_{PT}), and nonemployment ($h_{NE} \equiv 0$).

Workers also differ in skill endowments. They have either low-skill (s_L) or high-skill (s_H). While high-skilled workers are able to work in both high-skill and low-skill occupations, low-skilled workers cannot perform high-skill jobs, and thus can only be employed in the low-skill occupation. However, skills are not permanent characteristics. High-skilled workers may become low-skilled workers if they are not working in the high-skill occupation. This occurs to workers with probability γ_{NE} if they are not employed, or with probability γ_{LS} if they are working in the low-skill occupation. It is a plausible assumption that the market-relevant skills are more easily maintained if workers are working in the labour market. To incorporate this idea, I assume that $\gamma_{NE} \geq \gamma_{LS}$.

Workers can experience three different life stages. In the first life stage, all workers enter the labour market and start searching for jobs as nonemployed workers. This life stage continues until the arrival of a child, which occurs as an exogenous random event with probability η_1, and workers enter the next life stage, life stage 2. In life stage 2, workers need to arrange for childcare. Two options are available to meet this need. First, they can look after their child by themselves by choosing to be nonemployed. Alternatively, they can procure child care service. The flow cost of the service is τ if she works full time, or $\psi\tau$ if she works part time, where $\psi \in (0, 1)$.

In any period in life stage 2, they transition to life stage 3 with probability η_2. Life stage 3 represents time when the child has grown and childcare is no longer required. The environment surrounding workers in life stage 3 is similar to that in life stage 1 except that workers do not expect any more life stage changes. In any given time, workers may exit from the labour market permanently with probability ξ regardless of life stage.

The objective of workers is to maximize their discounted lifetime utility. Assume that workers have a time-discount rate of r, and their flow utility from job type $e \in \{HF, HP, LF, LP\}$ in a period in life stage t is given by period-by-period utility function $u_e(t,\theta)$, which is defined by

$$u_e(t,\theta) = \begin{cases} y_e^{\varepsilon_y} + \mathbb{I}_{FT}(e)[\theta(H - h_{FT})^{\varepsilon_\ell}] + \mathbb{I}_{PT}(e)[\theta(H - h_{PT})^{\varepsilon_\ell}] & \text{if } t = 1, 3 \\ \mathbb{I}_{FT}(e)[(y_e - \tau)^{\varepsilon_y} + \theta(H - h_{FT})^{\varepsilon_\ell}] & \\ \quad + \mathbb{I}_{PT}(e)[(y_e - \psi\tau)^{\varepsilon_y} + \theta(H - h_{PT})^{\varepsilon_\ell}] & \text{if } t = 2 \end{cases}$$

where $\mathbb{I}_{FT}(e)$, and $\mathbb{I}_{PT}(e)$ are the indicator functions defined by the following:

$$\mathbb{I}_{FT}(e) = \begin{cases} 1 & \text{if} \quad e \in \{HF, LF\} \\ 0 & \text{otherwise} \end{cases}, \quad \mathbb{I}_{PT}(e) = \begin{cases} 1 & \text{if} \quad e \in \{HP, LP\} \\ 0 & \text{otherwise} \end{cases}.$$

Assume that search frictions exist in the labour market, and workers meet with potential employers through a random search technology. Assume that job search is a costless activity and a worker searches for a job every period regardless of her employment status. Job offer arrival probabilities are given by λ_0 while the workers are not employed, and λ_1 while they are employed. The fraction of job offers accounted for by job type e is given by ϕ_e for $e \in \{HF, HP, LF, LP\}$.

Match separations between workers and jobs occur for both exogenous and endogenous reasons. On the one hand, exogenous job separation occurs at the rate δ. On the other hand, a worker can voluntarily quit her current job at any time without cost.

Assume that no employer agrees to change hours of work. Because of this, if a worker wishes to change her hours of work, she has to search for a job that meets her preferred hours.

3.3 Individual Worker's Problem

As is standard in labour search models, an individual worker's problem can be written using a stochastic dynamic programming framework. Workers carry three state variables: e (employment state), s (skill level), and t (life stage). Partitioning the state space according to the values of s and t yields the following six sets:

$$S(s_H, 1) \equiv \{(e, s_H, 1)|\ e \in \{NE, HF, HP, LF, LP\}\},$$

$$S(s_H, 2) \equiv \{(e, s_H, 2)|\ e \in \{NE, HF, HP, LF, LP\}\},$$

$$S(s_H, 3) \equiv \{(e, s_H, 3)|\ e \in \{NE, HF, HP, LF, LP\}\},$$

$$S(s_L, 1) \equiv \{(e, s_L, 1)|\ e \in \{NE, LF, LP\}\},$$

$$S(s_L, 2) \equiv \{(e, s_L, 2)|\ e \in \{NE, LF, LP\}\}, \quad \text{and}$$

$$S(s_L, 3) \equiv \{(e, s_L, 3)|\ e \in \{NE, LF, LP\}\}.$$

$S(s, t)$ contains all states which workers with skill level s can occupy in life stage t. The last three sets reflect the assumption that low-skilled workers cannot be employed in the high-skill occupation.

The state vector can move either within a set, or between sets. Moves within the set reflect only a change of employment state, and are made either by endogenous employment transitions or exogenous job separations. Transitions across the sets are driven either by life stage changes, or by skill depreciation. These transitions may also lead workers to choose to move to the nonemployment state.

First, consider the low-skilled workers' problem. Let $V_e(s,t;\theta)$ denote the value of state (e,s,t) to type θ workers. The value of nonemployment in life stage 3 to the low-skilled workers, $V_{NE}(s_L,3;\theta)$, is given by

$$\begin{aligned}V_{NE}(s_L,3;\theta) =& \frac{1}{1+r}\Big[u_{NE}(3,\theta) + \sum_{j\in\{LF,LP\}} \lambda_0\phi_j \max\left[V_j(s_L,3;\theta), V_{NE}(s_L,3;\theta)\right] \\ &+ \xi\bar{V} + (1-\lambda_0\phi_{LF} - \lambda_0\phi_{LP} - \xi)V_{NE}(s_L,3;\theta)\Big],\end{aligned} \tag{3.1}$$

Equation (3.1) shows that the value of the state comprises the period-by-period utility and option value associated with the state. Every period workers are nonemployed, they gain the utility flow corresponding to nonemployment. They meet with job vacancies with probability λ_0, are offered a low-skill full-time job with probability ϕ_{LF} or a low-skill part-time job with probability ϕ_{LP}, and decide whether to accept the offer. With probability ξ, they exit from the labour market permanently. If no event occurs, workers keep their current positions. With loss of generality, assume $\bar{V} = 0$. Then equation (3.1) can be rearranged to

$$(r+\xi)V_{NE}(s_L,3;\theta) = u_{NE}(3;\theta) + \sum_{j\in\{LF,LP\}} \lambda_0\phi_j \max\left[V_j(s_L,3;\theta) - V_{NE}(s_L,3;\theta), 0\right]. \tag{3.2}$$

Similarly, the value of being employed either full-time or part-time for low-skilled workers in life stage 3 is given by

$$\begin{aligned}(r+\xi)V_e(s_L,3;\theta) =& u_e(3;\theta) + \sum_{j\in\{LF,LP\}} \lambda_1\phi_j \max\left[V_j(s_L,3;\theta) - V_e(s_L,3;\theta), 0\right] \\ &+ \delta[V_{NE}(s_L,3;\theta) - V_e(s_L,3;\theta)]\end{aligned} \tag{3.3}$$

for $e \in \{LF, LP\}$. Equation (3.3) has the same interpretation as equation (3.2) except that the job offer arrival probability is given by λ_1, and the possibility of an exogenous job separation with probability δ is included.

For the earlier life stages, i.e., $t = 1, 2$, the value to low-skilled workers is given by

$$\begin{aligned}(r+\xi)V_{NE}(s_L,t;\theta) =& u_{NE}(t;\theta) + \sum_{j\in\{LF,LP\}} \lambda_0\phi_j \max\left[V_j(s_L,t;\theta) - V_{NE}(s_L,t;\theta), 0\right] \\ &+ \eta_t\left[V_{NE}(s_L,t+1;\theta) - V_{NE}(s_L,t;\theta)\right]\end{aligned} \tag{3.4}$$

or

$$\begin{aligned}(r+\xi)V_e(s_L,t;\theta) =& u_e(t;\theta) + \sum_{j\in\{LF,LP\}} \lambda_1\phi_j \max\left[V_j(s_L,t;\theta) - V_e(s_L,t;\theta), 0\right] \\ &+ \delta\left[V_{NE}(s_L,t;\theta) - V_e(s_L,t;\theta)\right] \\ &+ \eta_t\left\{\max\left[V_{NE}(s_L,t+1;\theta), V_e(s_L,t+1;\theta)\right] - V_e(s_L,t;\theta)\right\}\end{aligned} \tag{3.5}$$

for $e \in \{LF, LP\}$. As in equations (3.2) and (3.3), equations (3.4) and (3.5) show that the value for the workers consists of the current utility flow and the option value of the employment state. However, in these life stages, a possible transition to the next life stage also contributes to the option value. This is reflected in the last terms of both equations (3.4) and (3.5). In life stage t, workers transition to life stage $t+1$ with probability η_t. If they are employed when a transition to the next life stage occurs, they decide whether to quit the current job to be nonemployed.

Next, the high-skilled worker's problem is considered. Their problem differs from the low-skilled workers' problem in two ways. First, the high-skilled workers can be matched with high-skill jobs as well as low-skill jobs. Second, they face a risk of skill depreciation when they are not employed in the high-skill occupation.

The value of nonemployment to the high-skilled workers in life stage 3 is given by

$$\begin{aligned}(r+\xi)V_{NE}(s_H,3;\theta) =& u_{NE}(3;\theta) + \sum_{\substack{j\in\{HF,HP,\\LF,LP\}}} \lambda_0\phi_j \max\left[V_j(s_H,3;\theta) - V_{NE}(s_H,3;\theta), 0\right] \\ &+ \gamma_{NE}\left[V_{NE}(s_L,3;\theta) - V_{NE}(s_H,3;\theta)\right].\end{aligned} \tag{3.6}$$

The second term of the equation's right-hand side reflects the workers' job search. With probability λ_0, they receive a job offer, and with probability ϕ_j, they are given a chance to work in employment state j. The last term on the equation's right-hand side reflects the contribution of skill destruction to the value of the nonemployment state. With probability γ_{NE}, a worker's skill depreciates and the skill level changes to s_L.

Similarly, the value of working in the low-skill occupations in life stage 3 is given by

$$\begin{aligned}(r+\xi)V_e(s_H,3;\theta) =& u_e(3;\theta) + \sum_{\substack{j\in\{HF,HP,\\LF,LP\}}} \lambda_1\phi_j \max\left[V_j(s_H,3;\theta) - V_e(s_H,3;\theta), 0\right] \\ &+ \delta\left[V_{NE}(s_L,3;\theta) - V_e(s_H,3;\theta)\right] \\ &+ \gamma_{LS}\left\{\max\left[V_{NE}(s_L,3;\theta), V_e(s_L,3;\theta)\right] - V_e(s_H,3;\theta)\right\}\end{aligned} \tag{3.7}$$

for $e \in \{LF, LP\}$. On the right-hand side of equation (3.7), the second term accounts for job search, the third term accounts for exogenous job separation, and the last term accounts for skill destruction. Skill destruction may induce workers to leave their current jobs if the value associated with the current job becomes lower than the value of nonemployment.

The workers do not face a risk of skill destruction while working in the high-skill occupation. Therefore, the option value associated with the high-skill jobs in life stage 3 consists of just a possible job-to-job transition and an exogenous job separation. For $e \in \{HF, HP\}$,

$V_e(s_H, 3; \theta)$ is given by

$$(r+\xi)V_e(s_H,3;\theta) = u_e(3,\theta) + \sum_{\substack{j\in\{HF,HP,\\LF,LP\}}} \lambda_1\phi_j \max\left[V_j(s_H,3;\theta) - V_e(s_H,3;\theta), 0\right] + \delta[V_{NE}(s_H,3;\theta) - V_e(s_H,3;\theta)]. \tag{3.8}$$

For life stages $t = 1, 2$, the high-skilled workers take into account a transition to the next life stage, which occurs with probability η_t. The value functions also account for skill destruction that workers may experience when they are not working in the high-skill occupation. The value of nonemployment to the high-skilled workers in these life stages, $V_{NE}(s_H, t; \theta)$, is given by

$$\begin{aligned}(r+\xi)V_{NE}(s_H,t;\theta) = u_{NE}(t;\theta) &+ \sum_{\substack{j\in\{HF,HP,\\LF,LP\}}} \lambda_0\phi_j \max\left[V_j(s_H,t;\theta) - V_{NE}(s_H,t;\theta), 0\right] \\ &+ \eta_t[V_{NE}(s_H,t+1;\theta) - V_{NE}(s_H,t;\theta)] \\ &+ \gamma_{NE}[V_{NE}(s_L,t;\theta) - V_{NE}(s_H,t;\theta)].\end{aligned} \tag{3.9}$$

For each job type $e \in \{LF, LP, HF, HP\}$, $V_e(s_H, t; \theta)$ is given by

$$\begin{aligned}(r+\xi)V_e(s_H,t;\theta) = u_e(t;\theta) &+ \sum_{\substack{j\in\{HF,HP,\\LF,LP\}}} \lambda_1\phi_j \max\left[V_j(s_H,t;\theta) - V_e(s_H,t;\theta), 0\right] \\ &+ \delta[V_{NE}(s_H,t;\theta) - V_e(s_H,t;\theta)] \\ &+ \eta_t\{\max\left[V_{NE}(s_H,t+1;\theta), V_e(s_H,t+1;\theta)\right] - V_e(s_H,t;\theta)\} \\ &+ \mathbb{I}_{LS}(e)\gamma_{LS}\{\max\left[V_{NE}(s_L,t;\theta), V_e(s_L,t;\theta)\right] - V_e(s_H,t;\theta)\}\end{aligned} \tag{3.10}$$

where $\mathbb{I}(e)$ is an indicator function taking 1 if $e \in \{LF, LS\}$ and 0 otherwise. The last term on the right-hand side of this equation captures the effect of skill destruction that may occur to high-skilled workers working in the low-skill occupation, and this term is irrelevant when they are working in the high-skill occupation.

In the remainder of this section, the workers' decision rules are discussed. Since the transitions of s and t are exogenous to workers, the workers' decision rules govern transition between the employment states given their skill levels and life stages. Given skill level s and life stage t, their decision rules are determined by how they rank the values of $V_e(s, t; \theta)$ over $S(s, t)$. Within $S(s, t)$, workers always transition to states that rank higher than their current state when they encounter such an opportunity, or remain in the present state otherwise.

Generally, low-skill employment is never preferred to high-skill employment with the same hours of work, as low-skill employment yields a lower earnings flow than high-skill employment and the likelihood of the arrival of an exogenous event is independent of hours worked. However, it is possible that low-skill full-time work is preferred to high-skill part-time work, depending on a worker's valuation of leisure relative to income. Similarly, low-skill part-time work may be preferred to high-skill full-time work by some workers.

There exists θ such that the values of two employment states with different hours of work coincide with each other given a skill level and a life stage. Given $s \in \{s_L, s_H\}$ and $t \in \{1, 2, 3\}$, these cutoff values, denoted by $\theta_{je}(s, t)$, are implicitly given by

$$V_j(s, t; \theta_{je}(s, t)) = V_e(s, t; \theta_{je}(s, t))$$

where j and e are a pair of employment states to be compared.[5] These values indicate the valuation of leisure of workers who are indifferent between the two relevant employment states.

For both high-skilled and low-skilled workers, the cutoff values differ across the three life stages for two reasons. First, the cutoff values relevant to life stage 2 differ from those

[5]For low-skilled workers, all the possible paris for (j, e) are (LF, LP), (LF, NE), and (LP, NE). In addition to them, there are five more pairs for high-skilled workers to consider: (LF, HP), (HF, NE), (HF, LP), (HF, HP) and (HP, NE).

pertaining to life stages 1 and 3 because the childcare costs reduce the pecuniary returns to employment in life stage 2. Second, while the flow utility from work is the same in both life stages 1 and 3, the cutoff values differ between these two life stages because of workers' forward-looking decision-making behaviour. In life stage 1, workers expect a transition to life stage 2, where they face different economic conditions. Anticipating this, they do not make decisions solely based on the current environment, but also on that of the next life stage. In contrast, such concern is absent in life stage 3 since there is no further life stage changes.

These cutoff values divide the type space into several intervals such that within each of the intervals, workers have a common ranking over the employment states. By the assumption that the value of leisure relative to income is increasing in θ, full-time employment ranks higher in an interval containing low values of θ than in those intervals containing high values of θ. Once θ is sufficiently high, workers effectively stay out of labour force by not accepting any job offer.

Interestingly, with a difference in job offer arrival probability between nonemployment and employment, deriving no value from leisure does not necessarily mean that nonemployment is ranked the lowest of all the possible employment states. If λ_0 is larger than λ_1, there may be two distinct cutoff values equating nonemployment with a particular job type. Then, workers with θ below the lower cutoff value and those with θ above the higher cutoff value prefer nonemployment to this particular job, while the rest values the former below the latter. This occurs because workers with sufficiently low valuation of leisure may reject a job offer with a low pay, preferring to stay nonemployed since the likelihood of receiving a job offer with a higher pay is higher while nonemployed.

Beyond these results, a complete analytical characterization of the workers' problem is difficult given the large number of parameters in the model, making it intractable to describe their transition behaviour under all parameter configurations. Instead, I present a

Table 3.1: Baseline Parameter Values

Hours	h_{FT}	1.00	Life cycle	η_1	0.0391
	h_{PT}	0.55		η_2	0.0083
	H	1.25		ξ	0.0028
Pay	y_{LF}	1.00	Skill	γ_{NE}	0.0333
	y_{HF}	1.20	depreciation	γ_{LS}	0.0167
	y_{LP}	0.55	Job search	λ_0	0.20
	y_{HP}	0.66		λ_1	0.04
Childcare	τ	0.25		δ	0.006
cost	ψ	0.55	Offer	ϕ_{LF}	0.25
Preference	ε_y	0.7	distribution	ϕ_{HF}	0.25
	ε_ℓ	0.7		ϕ_{LP}	0.45
				ϕ_{HP}	0.05

few numerical examples of the model in the next section.

3.4 Numerical Illustrations

This section presents numerical examples of the model. The goal of this exercise is twofold. First, using numerical examples, I illustrate the general properties of the workers' problem described in the previous section. Second, I examine whether the model can capture the stylized facts of mothers' life-cycle employment transition patterns under reasonable parameter values.

Table 3.1 presents the baseline parameter values. The hours of full-time work and part-time work are set at $h_{FT} = 1$ and $h_{PT} = 0.55$, respectively, while the total time endowment is set at $H = 1.25$. The baseline parameter values envision a 20% wage gap between the high-skill and low-skill occupations, and the difference in earnings between full-time and part-time jobs within the same occupation purely reflects the difference in hours worked. Therefore, with the normalization of $y_{LF} = 1$, I set $y_{HF} = 1.2$, $y_{HP} = 0.66$, and $y_{LP} = 0.55$. The childcare costs incurred to work full time are set at a quarter of the earnings of low-skill full-time work. Thus $\tau = 0.25$. The parameter governing part-time child care cost, ψ, is set at 0.55.

The parameters governing changes in life stage, skill destruction, and employment transitions are set as follows. The life-cycle parameters, η_1, η_2 and ξ, are set at 0.0139, 0.0083, and 0.0028, respectively.[6] The skill destruction probabilities, γ_{NE} and γ_{LS}, are set to 0.0333 and 0.0167, respectively.[7] The job offer arrival probabilities are set at $\lambda_0 = 0.2$ while nonemployed, and at $\lambda_1 = 0.04$ while employed.[8] The job destruction probability is set at $\delta = 0.006$.

The job offer parameters are set at values that reflect the scarcity of high-skill part-time work in a simple way. Specifically, the fraction of high-skill part-time job offers among all offers, ϕ_{HP}, is set at 0.05, while ϕ_{LP} is set at 0.45, and the rest, ϕ_{HF} and ϕ_{LF}, are both set at 0.25.

Tables 3.2 and 3.3 present the cutoff values $\theta_{je}(s,t)$ defined in the previous section. In each table, the second and third columns indicate the two employment states to be compared, and the remaining columns list the cutoff values for given life stages under particular parameter configurations.

The fourth column in Tables 3.2 and 3.3 presents the results under the baseline parameter values. There are two distinct cutoff values equating low-skill part-time work and nonemployment for high-skilled workers in each life stage. The interval bounded by these two values contains valuations of leisure that rank low-skill part-time work higher than nonemployment. For any other value of θ, nonemployment is ranked higher than lot-skill part-time work.

[6] These parameter values imply that the first and second life stages are on average 72 and 120 periods long, respectively, and the average time spent in the labour market is 360 periods. Using months as a time unit, these numbers yield 6, 10 and 30 years.

[7] With these parameters, the expected length of time before a high-skilled worker experience skill destruction is 2.5 years while nonemployed , and is 5 years while employed in the low-skill occupation.

[8] These probabilities imply that on average, it takes 5 months to receive a job offer while non employed, and 25 month while employed.

The rest of the cutoff values indicate that θ below them ranks employment state in the second column higher than the one in the third column, as the employment state listed in the former is associated with less hours of leisure, and the valuation of leisure is increasing in θ. For example, $\theta_{LFHP}(s_H, 1) = 0.334$ means that high-skilled workers with $\theta < 0.334$ prefer the low-skill full-time work to high-skill part-time work in life stage 1, and those with $\theta > 0.334$ prefer the latter to the former.

Because of the differences in the cutoff values across different life stages, a transition to the next life stage induces some workers to change their transition behaviour. For example, low-skilled workers with $\theta = 0.7$ prefer full-time work to part-time work in life stage 1, but the opposite is true in life stage 2. Thus upon moving to life stage 2, low-skilled workers with $\theta = 0.7$ working full time wish to move to part-time work. The ranking between low-skill full-time and part-time work is reversed again for these workers in life stage 3, inducing them to move back to full-time work.

These cutoff values divide workers into several groups having a common ranking over the different employment states. Table 3.4 shows how the ranking of the employment states changes according to θ under the baseline parameter values. The table consists of six parts, each of which shows how different levels of valuation of leisure are divided into several intervals for a particular pair of skill level and life stage. For instance, the upper left part of the table shows how high-skilled workers rank the employment states in life stage 1.[9]

The extent to which the model captures mothers' occupational downgrading and moves to persistent part-time employment in response to the arrival of a child hinges on the predicted employment transition patterns of high-skilled workers. Indeed, the middle left panel of Table 3.4 shows that some high-skilled workers find it preferable to work in low-skill part-time jobs over both nonemployment and full-time work in life stage 2 ($\theta \in$

[9] Values considered for θ range from 0 to 1.917. High-skilled workers with θ greater than 1.917 prefer nonemployment to any employment type in any life stage, effectively choosing to be labour market non-participants.

Table 3.2: Cutoff Values - High-Skilled Workers

Life Stage	Employment States		Parameter Configuration			
			baseline	γ_{LS} = 0.0139	ψ = 0.6	η_2 = 0.0093
1	LP	NE	0.715	0.653	0.722	0.717
	LP	NE	1.630	1.636	1.632	1.629
	LF	NE	0.913	0.937	0.911	0.912
	HP	NE	1.917	1.917	1.917	1.917
	HF	NE	1.327	1.326	1.331	1.326
	LF	LP	0.841	0.839	0.844	0.842
	HF	LP	1.218	1.214	1.221	1.217
	LF	HP	0.334	0.360	0.342	0.336
	HF	HP	0.922	0.922	0.931	0.924
2	LP	NE	0.856	0.767	0.927	0.860
	LP	NE	1.311	1.319	1.282	1.314
	LF	NE	0.697	0.727	0.703	0.700
	HP	NE	1.746	1.746	1.730	1.754
	HF	NE	1.177	1.177	1.187	1.180
	LF	LP	0.734	0.735	0.757	0.737
	HF	LP	1.150	1.147	1.167	1.153
	LF	HP	0.234	0.261	0.257	0.238
	HF	HP	0.864	0.864	0.884	0.867
3	LP	NE	0.781	0.723	0.781	0.781
	LP	NE	1.601	1.609	1.601	1.601
	LF	NE	0.892	0.916	0.892	0.892
	HP	NE	1.917	1.917	1.917	1.917
	HF	NE	1.309	1.309	1.309	1.309
	LF	LP	0.855	0.855	0.855	0.855
	HF	LP	1.229	1.227	1.229	1.229
	LF	HP	0.390	0.415	0.390	0.390
	HF	HP	0.971	0.971	0.971	0.971

$(1.150, 1.311)$). Since these workers prefer high-skill full-time work to low-skill part-time work in life stage 1, their predicted transitions between these two employment states are reversed upon the arrival of a child. Thus the model predicts voluntary transitions from high-skill full-time work to low-skill part-time work. Furthermore, these moves may have a permanent effect on their transition behaviour because the low-skill employment carries

Table 3.3: Cutoff Values - Low-Skilled Workers

Life Stage	Employment States		Parameter Configuration			
			baseline	γ_{LS} = 0.0139	ψ = 0.6	η_2 = 0.0093
1	LF	NE	1.060	1.060	1.061	1.060
	LF	LP	0.817	0.817	0.824	0.818
	LP	NE	1.687	1.687	1.687	1.687
2	LF	NE	0.879	0.879	0.890	0.881
	LF	LP	0.741	0.741	0.762	0.744
	LP	NE	1.402	1.402	1.375	1.405
3	LF	NE	1.042	1.042	1.042	1.042
	LF	LP	0.855	0.855	0.855	0.855
	LP	NE	1.687	1.687	1.687	1.687

the risk of skill depreciation. When skill destruction hits these workers and makes them low-skilled, low-skill part-time work becomes their preferred employment option in the remainder of their labour market careers.

These workers have moderately high values of leisure, and are sensitive to economic incentives arising from both childcare needs and skill depreciation. A reduction in net earnings due to the childcare costs has a large negative impact on their incentives to work. Without a fear of skill depreciation, they would simply choose the nonemployment state in response to the arrival of a child. However, they still hope to work full time in the high-skill occupation when they no longer face childcare responsibilities, and nonemployment presents the largest risk of skill depreciation. It is this trade-off that makes part-time employment attractive to these workers. Part-time work has an advantage over full-time work in the provision of child care, and part-time work has an advantage in preserving skill over nonemployment.

These results are specific to the baseline parameters. Next, I examine how $\theta_{je}(s,t)$ changes in response to changes in some of the parameters controlling key features of the model: skill depreciation, childcare cost, and the length of childrearing period. First, I

Table 3.4: Ranking of Value Function in the Benchmark Parameter Configuration

	$s = s_H$						$s = s_L$			
	θ interval	Ranking					θ inverval	Ranking		
t		1	2	3	4	5		1	2	3
1	(0.000, 0.334)	HF	LF	HP	NE	LP	(0.000, 0.817)	LF	LP	NE
	(0.334, 0.715)	HF	HP	LF	NE	LP	(0.817, 1.060)	LP	LF	NE
	(0.715, 0.841)	HF	HP	LF	LP	NE	(1.060, 1.687)	LP	NE	LF
	(0.841, 0.913)	HF	HP	LP	LF	NE	(1.687, 1.917)	NE	LP	LF
	(0.913, 0.922)	HF	HP	LP	NE	LF				
	(0.922, 1.218)	HP	HF	LP	NE	LF				
	(1.218, 1.327)	HP	LP	HF	NE	LF				
	(1.327, 1.630)	HP	LP	NE	HF	LF				
	(1.630, 1.917)	HP	NE	LP	HF	LF				
2	(0.000, 0.234)	HF	LF	HP	NE	LP	(0.000, 0.741)	LF	LP	NE
	(0.234, 0.697)	HF	HP	LF	NE	LP	(0.741, 0.879)	LP	LF	NE
	(0.697, 0.734)	HF	HP	NE	LF	LP	(0.879, 1.402)	LP	NE	LF
	(0.734, 0.856)	HF	HP	NE	LP	LF	(1.402, 1.917)	NE	LP	LF
	(0.856, 0.864)	HF	HP	LP	NE	LF				
	(0.864, 1.150)	HP	HF	LP	NE	LF				
	(1.150, 1.177)	HP	LP	HF	NE	LF				
	(1.177, 1.311)	HP	LP	NE	HF	LF				
	(1.311, 1.746)	HP	NE	LP	HF	LF				
	(1.746, 1.917)	NE	HP	LP	HF	LF				
3	(0.000, 0.390)	HF	LF	HP	NE	LP	(0.000, 0.855)	LF	LP	NE
	(0.390, 0.781)	HF	HP	LF	NE	LP	(0.855, 1.042)	LP	LF	NE
	(0.781, 0.855)	HF	HP	LF	LP	NE	(1.042, 1.687)	LP	NE	LF
	(0.855, 0.892)	HF	HP	LP	LF	NE	(1.687, 1.917)	NE	LP	LF
	(0.892, 0.971)	HF	HP	LP	NE	LF				
	(0.971, 1.229)	HP	HF	LP	NE	LF				
	(1.229, 1.309)	HP	LP	HF	NE	LF				
	(1.309, 1.601)	HP	LP	NE	HF	LF				
	(1.601, 1.917)	HP	NE	LP	HF	LF				

consider a case where the low-skill occupation presents a lower risk of skill depreciation by decreasing γ_{LS} to 0.0139 from the baseline value of 0.0167. Even if workers choose to experience occupational downgrading, they are likely to look for occupations that require similar type of skills to perform similar lines of work. In this case, the risk of skill depreciation may be mitigated. The cutoff values for high-skilled workers under the new parameter value configuration are presented in the fifth column in Table 3.2.[10] The decrease in γ_{LS} leads the cutoff values equating the high-skill occupation and the low-skill occupation to decrease. It also expands the interval where low-skill part-time work is ranked higher than nonemployment at each life stage, since the distance between the two relevant cutoff values increases. Overall, the change leads more high-skill workers to accept occupational downgrading to switch part-time work in life stage 2 .

Second, the baseline parameter value of $\psi = 0.55$ implies that the childcare costs are proportional to hours of work. However, there is evidence that the hourly price of childcare increases at a decreasing rate in hours (Duncan et al., 2001). Thus, ψ is increased to 0.6 to consider potential nonlinearity of the childcare cost. This change decreases the pecuniary returns from part-time work in life stage 2, creating an effect of lowering the value of part-time employment relative to the other employment states. Comparing the cutoff values in the fourth and sixth columns in Tables 3.2 and 3.3 reveals that the increase in the part-time childcare costs narrows the range for θ where part-time employment ranks higher than full-time employment or nonemployment, implying fewer transitions to part-time work with more workers choosing to stay working full-time or leaving the labour market in life stage 2.

Lastly, η_2 is increased to 0.0093 to consider a case where the expected duration of the period requiring childcare is shortened, with the new parameter value implying that the

[10]This change does not affect low-skilled workers' problem since skill destruction does not affect their problem.

average length of life stage 2 is 108 periods instead of 120 periods. The seventh column in Tables 3.2 and 3.3 shows that the cutoff values equating full-time work to the other employment states have increased in life stage 2, indicating that the increase in η_2 raises the value of full-time work relative to other employment states in this life stage. This change stems from the workers' forward-looking behaviour. Due to search frictions, workers cannot change their employment states immediately when a shock changes their economic conditions, and if their current employment states differ from their optimal employment states after the shock, they have to forgo values that could have been attained but for the frictions. Therefore, there is value to occupying the employment state which will be optimal after a shock hits workers, and such value increases with the likelihood of the shock. In this model, workers tend to prefer to work full time in life stage 3 so that the higher likelihood of transitioning from life stage 2 to life stage 3 increases the value of full-time work in life stage 2.

Overall, the baseline parameter values reasonably reflect the labour market environment facing British women. First, using data from BHPS, Paull (2008) finds that the distribution of hours worked by female workers in Britain has two peaks, located around 16-20 and 36-40 hours per week, respectively. By treating them as typical hours of part-time and full-time work, respectively, the ratio between them yields a similar value implied by the baseline parameter values, i.e., $h_{pt}/h_{ft} = 0.55$. The 20% wage gap between the high-skill and low-skill occupations closely reflects the wage rate reduction associated with a common episode of occupational downgrading experienced by British women, as presented in Connolly and Gregory (2008). With τ set at 0.25, the childcare costs account for a quarter of earnings for low-skilled workers, and approximately two-fifths for high-skilled workers. These values are in line with the ratio between the average hourly payment of childcare and the average hourly wage reported in Viitanen (2005).[11]

[11] Using the British Family Resources Survey from 1997 and 2004, Viitanen (2005) calculated the average

Second, the chosen values for the skill destruction parameters provide quantitatively reasonable implications for the expected wage loss and the likelihood to return to the high-skill occupation for high-skilled workers who experience occupational downgrading or nonemployment. Consider a high-skilled worker who switches to low-skill part-time work when she enters life stage 2, but prefers to go back to high-skill full-time work once she leaves this life stage. The baseline parameter configuration implies that the probability that she maintains her skill level during this life stage is 0.34.[12] This value closely represents the probability of returning to her previous occupation. The value is in line with the findings in Dex and Bukodi (2012) that only one third of the occupational downgrading experienced by British women changing from full-time work to part-time work was reversed later in their careers.

The baseline value of γ_{NE} implies that the probability of skill depreciation resulting from one year spell of nonemployment is 0.2870.[13] With the 20% wage gap between the high-skill and low-skill occupations, this probability implies that the expected wage loss associated with one year of nonemployement for high-skilled workers is 5.9%, which is in line with the findings in Mincer and Ofek (1982).

Third, the values for the parameters governing employment transitions are also reasonable. The baseline value of $\lambda_0 = 0.2$ implies that the per-period probability of exiting from nonemployment is 0.2 for high-skilled workers accepting any type of job offers, and 0.15 for low-skilled workers accepting any eligible job offers, implying that the average dura-

hourly payment of childcare based on actual payments made to childcare workers or nurseries. Thus, the ratio between this value and the hourly wage adequately represents the size of childcare costs relative to earnings. Chevalier and Viitanen (2002) also calculated the average hourly payment to formal childcare in an analogous manner for the time period between 1994 and 1998 from the same survey, obtaining a comparable value. The figure was also comparable with the cost of a typical nursery placement reported in Daycare Trust (2001).

[12]Recall that in the model, the childrearing period, i.e., life stage 2, ends with probability η_2 in a given period, and skill depreciation occurs with probability γ_{LS} in any given period working in the low-skill occupation. These two processes are independent. Under these assumption that the probability that skill depreciation occurs before the end of life stage 2 is given by $\gamma_{LS}/(\gamma_{LS} + \eta_2)$.

[13]The value corresponds to $1 - (1 - \gamma_{NE})^{12}$.

tions spent out of work for these two groups of workers are 5 and 6.7 months, respectively. These values are close to the average unemployment duration of female workers based on data from BHPS in Böheim and Taylor (2000). λ_1 is set one fifth the job offer arrival probability while nonemployed. The same ratio between λ_1 and λ_0 is chosen by Mortensen (1994). Moreover, $\lambda_1/\lambda_0 = 0.2$ is in the range of the values estimated by Bowlus and Grogan (2009) using data from BHPS. The value of δ, the per-period job separation probability, matches the corresponding estimates in Bowlus and Grogan.

Some limitation exists in the model, however. Recall that ε_ℓ and ε_y govern the degrees of diminishing marginal utilities with respect to leisure and income, respectively. Not too surprisingly, the model's ability to produce transitions to part-time work involving occupational downgrading is somewhat sensitive to them. For example, with an increase in ε_ℓ from 0.7 to 0.8, which reduces the degree of diminishing marginal utility with respect to leisure, the model fails to find high-skilled workers preferring low-skill part-time work to both high-skill full-time work and nonemployment in life stage 2, ruling out voluntary occupational downgrading at the onset of life stage 2. This invalidates the model's applicability to the observed transition patterns of British mothers.[14]

Modifying the job search process may be an interesting way to make the model more robust. Increasing the efficiency of on-the-job search will increase the value of employment relative to the value of nonemployment, switching some high-skills workers' ranking in favour of low-skill part-time work over nonemployment. Allowing endogenous search effort may achieve this without contradicting the documented differences in job offer arrival probabilities between nonemployment and employment. As shown in Christensen et al. (2005), incorporating endogenous job search effort can explain these differences even with a common search technology. Extending the current model in this dimension may be an

[14]In contrast, lowering ε_ℓ or ε_y induces more concavity in the utility function, and increases the value of part-time work because it is associated with intermediate outcomes both in hours and earnings.

interesting avenue for future work.

Lastly, the present model is admittedly stylized and abstracts from a number of important aspects. These include earnings growth on the job, restoration of skills, and joint decision-making within a household. Incorporating these aspects may provide the model with better predictive performance as well.

3.5 Concluding Remarks

It is common for British women to switch from full-time to part-time employment when they become mothers, even though such moves are often associated with occupational downgrading. Interestingly, not many reverse these changes later in their careers. In this chapter, I offer a potential explanation for these patterns which emphasizes the role of skill depreciation. In the model developed here, workers' endogenous choices over hours of work and occupations in different life stages can produce the patterns under reasonable parameter values.

The employment patterns studied in this chapter come from a time period when the UK government primarily viewed childcare as the responsibilities of parents (Randall, 1995). More recently, several policy changes have been enacted to improve the balance between work and family lives of workers, and the government is likely to continue playing a proactive role in childcare provision (Lewis, 2003; Lewis and Campbell, 2007).[15] Therefore, it is important for future work to incorporate the elements of those policies to account for their impact on mothers' preferences over various employment options and career choices beyond the childcare period.

In particular, in 2003, an employment law that grants parents with children under age

[15] The policies recently enacted to improve workers' work-life balance include free public part-time early childhood education for children aged between three and four, and the right to request flexible working patterns for parents with children aged under six.

six the right to request a flexible working arrangement from employers was introduced. While Grainger and Holt (2005) reported that the most popular request from women using this right was to switch to part-time work, the model suggests other potential impacts of this policy. For example, it may encourage mothers to return to full-time employment later in their working lives by allowing them to stay in the same occupation, and thus avoid occupational downgrading and the associated skill depreciation. As a result, part-time employment may become more of a temporary state. The policy may also encourage women to pursue career paths where the continuity of job experience is important for career progression. Thus, the model can provide a framework to evaluate family-friendly polices from the view point of life-cycle skill dynamics, hours and employment.

Bibliography

Albrecht, James, and Susan Vroman. 2002. A matching model with endogenous skill requirements. *International Economic Review* 43(1):283–305.

Böheim, René, and Mark P. Taylor. 2000. Unemployment duration and exit states in Britain. Working paper.

Bowlus, Audra J., and Louise Grogan. 2009. Gender wage differentials, job search, and part-time employment in the UK. *Oxford Economic Papers* 61(2):275–303.

Chevalier, Arnaud, and Tajra K. Viitanen. 2002. The supply of childcare in Britain: do mothers queue for childcare? Working paper.

Christensen, Bent Jesper, Rasmus Lentz, Dale T. Mortensen, George R. Neumann, and Axel Werwatz. 2005. On-the-job search and the wage distribution. *Journal of Labor Economics* 23(1):31–57.

Connolly, Sara, and Mary Gregory. 2008. Moving down: women's part-time work and occupational change in Britain 1991–2001. *The Economic Journal* 118:F52–76.

Daycare Trust. 2001. The price parents pay: sharing the cost of childcare. Publication No. 3.

Dex, Shirley, and Erzsébet Bukodi. 2012. The effects of part-time work on women's occupational mobility in Britain: evidence from the 1958 birth cohort study. *National Institue Economic Review* 222(1):R20–R37.

Duncan, Alan, Gillian Paull, and Jayne Taylor. 2001. Price and quality in the UK childcare market. Working Paper.

Grainger, Heidi, and Heather Holt. 2005. Results of the second flexible working employee survey. *Labour Market Trends* .

Hirsh, Barry T. 2005. Why part-time workers earn less? The role of worker and job skills. *Industrial and Labor Relations Review* 58:1–27.

Lewis, Jane. 2003. Developing early years childcare in England, 1997-2002: the choices for (working) mothers. *Scial Policy and Administration* 37(3):219–238.

Lewis, Jane, and Mary Campbell. 2007. Work/family balance policies in the UK since 1997: a new departure? *Journal of Social Policy* 36(3):365–381.

Manning, Allan, and Barbara Petrongolo. 2008. The part-time pay penalty for women in Britain. *The Economic Journal* 118:F28–51.

Mincer, Jacob, and Haim Ofek. 1982. Interrupted work careers: depreciation and restration of human capital. *The Journal of Human Resource* 17:3–24.

Mortensen, Dale T. 1994. The cyclical behavior of job and worker flows. *Journal of Economic Dynamics and Control* 18:1121–1142.

Paull, Gillian. 2008. Children and women's hours of work. *The Economic Journal* 118:F8–27.

Randall, Vicky. 1995. The irresponsible state? The politics of child care provision in Britain. *British Journal of Political Science* 25(3):327–348.

Viitanen, Tarja K. 2005. Cost of childcare and female employment in the UK. *Labour* 149–170.

Chapter 4

Immigrant Job Search Assimilation in Canada

4.1 Introduction

Immigration has always played a major role in Canada. It serves as an important source of labour for the country, and therefore successful integration of immigrants in the labour market has direct implications for the Canadian economy. There is a large literature documenting and analyzing Canadian immigration patterns and the relative success of Canadian immigrants by landing cohort and immigrant type in the labour market.[1] The vast majority of this research, and for that matter research on immigrant assimilation in all countries, has been based on the standard human capital model.[2]

The lower earnings of immigrants upon arrival is attributed to their lower levels of human capital specific to the host country, and the catch-up of their earnings is explained by their stronger incentives to invest in host-country specific human capital. Numerous papers

[1]See, for example, Abbott and Beach (2011), Green and Worswick (2010), Xue (2010), Baker and Benjamin (1994), and Sweetman and Warman (2013).

[2]Chiswick (1978); Borjas (1999)

have been written studying the relationship between immigrants' earnings and observable characteristics related to their human capital acquired in the source country as well as that acquired in the host country. The set of characteristics often includes information on educational attainment and labour market experience in each country as well as age and literacy level at immigration.[3]

A new approach to the immigrant-native earnings gap has recently emerged and is asking how much of the gap is due to differences in job search behavior.[4] There are a number of reasons to expect that new immigrants face a different search environment than natives. First, it is natural to think that newly arrived immigrants are simply not accustomed to local practices of job search. Second, the same factors accounting for the native-immigrant human capital gap may contribute to job search differentials between natives and immigrants. Examples include qualification recognition and language fluency. Possible difficulties of having foreign credentials recognized by employers may slow the application process, while language fluency likely plays an important role in the job search and interview processes.[5]

It is also important to consider the role of social networks in job search. Differences in networks formed and network usage between natives and immigrants are likely to be reflected in differences in job search outcomes between natives and immigrants. A number of papers have documented that workers utilize friends and relatives when searching for a job, and recent work indicates that social networks are especially important for newly-arrived immigrants. Using the Longitudinal Survey of Immigrants to Canada (LSIC), Goel and Lang (2012) find that having close friends or relatives in Canada significantly increases the likelihood that recently arrived immigrants will find employment within the first six

[3]See, for example, Skuterud and Su (2012).

[4]Examples include Chassamboulli and Palivos (2010), Ortega (2000) and Liu (2010), Lessem and Sanders (2012), and Gupta and Kromann (2013)

[5]Oreopoulos (2011) and Dechief and Oreopoulos (2012) provide evidence that employers discriminate against job applicants with foreign-sounding names in hiring, associating with them low local language skills.

months of their arrival.[6] Other examples include Munshi (2003) who documents that use of social networks is very common in acquiring employment for Mexican migrant workers in the United States, and Frijters et al. (2005) who document that immigrants in the U.K. tend to rely on their social networks to obtain a job more than natives.

Taking this idea of differences in job search behaviour one step further this paper uses search theory to examine the role of job search in immigrant assimilation. The idea being that not only can immigrants invest in human capital once they enter the host country to achieve earnings assimilation but that they can also learn more about the labour market and how to better search for a job. This idea is similar to that proposed by Daneshvary et al. (1992) who take the view that immigrants acquire more information about the host country's labour market the longer they have been there and this results in earnings that are much closer to their potential or maximum attainable earnings. Here we take a different approach and develop a Burdett-Mortensen style general equilibrium search model with two types of workers: immigrant workers and native workers. Importantly, new immigrants face more search frictions than natives but over time can assimilate such that they face the same search frictions as natives.

The Burdett-Mortensen equilibrium search model has been used to explain earnings differentials between many different groups, including the male-female earnings gap (Bowlus, 1997), the black-white earnings gap (Bowlus et al., 2001; Bowlus and Eckstein, 2002), and the family earnings gap (Zhang, 2012). This paper extends this tradition to the immigrant earnings gap. The Burdett-Mortensen search model can generate earnings differentials because how fast workers are able to generate job offers has direct implications for their earnings. More specifically, the theory predicts that the distribution of the earnings of a group with higher job offer arrival rates dominates that of a group with lower rates. If new

[6] Interestingly the structural estimates of their search model reveal that the presence of close ties in Canada increases the likelihood of receiving a job offer but does not help in finding better offers.

immigrants have search behaviour that is characterized by lower job offer arrival rates than those of native born workers, the model will predict an immigrant earnings gap. However, if there is an assimilation process such that immigrants can learn how to search more effectively over time and increase their arrival rates of job offers, then the model predicts that this assimilation of the search process will provide a mechanism for earnings convergence.

To conduct our analysis we build on the model developed by Zhang (2012) to study the difference in earnings between mothers and non-mothers. Zhang extended the Burdett-Mortensen equilibrium search model in two important ways. First, her model has two groups of workers with different job offer arrival rates conducting job search in a single labour market.[7] Second, her model allows one of the groups (in her case non-mothers) to transition to the other group (mothers). Firms then take these transitions into account when posting offers and equilibrium earnings differentials result. This model setting fits our purposes well as it is important for us to model both the job search differences between immigrants and natives, but also the possibility that immigrants can assimilate and become like natives in their job search behaviour. In addition, we also capture the general equilibrium effects of both the presence of the immigrants in the labour market and their assimilation behaviour on not only their earnings levels but also on those of natives.

We estimate the model using duration and earnings data from the Canadian Survey of Labour and Income Dynamics (SLID) to measure the difference in job offer arrival rates between new immigrants and natives, and estimate how long it takes immigrants to acquire the same job search parameters as natives. Then, we study the implications of this assimilation process for the earnings gap between immigrants and natives, and immigrants' earnings growth.[8]

[7]Prior to Zhang (2012) the standard method was to assume that the two groups were operating in separate markets. Because of the assimilation process this standard method is not suitable for our setting.

[8]Papers using SLID to study immigrant assimilation include Skuterud and Su (2012), Hum and Simpson (2004) and Hum and Simpson (2000).

Our estimation results indicate that there are substantial and significant differences in job offer arrival and job destruction rates between natives and newly arrived immigrants. Job offer arrival rates for immigrants are 36% lower while unemployed and 93% lower while employed. The latter figure has substantial consequences for the amount of search frictions faced by immigrants. Immigrants receive almost no job offers during an employment spell compared to natives who receive nearly 2 offers. Importantly, these differences are able to account for three quarters of the observed earnings differential between natives and immigrants.

Our results also indicate that it takes immigrants 13 years, on average, to assimilate and acquire the same search parameters as natives, with counterfactual exercises indicating that job search assimilation accounts for the vast majority of earnings growth immigrants experience after migration. Studies of earnings assimilation using the human capital framework have reported a wide range of estimates of the time it takes for assimilation. In a recent study on assimilation of Canadian immigrants Skuterud and Su (2012) find an initial gap of 0.29 log points is halved after 8 years and a slower narrowing subsequently. Thus our results on the search component are consistent in magnitude with the human capital based assimilation literature.

The remainder of the paper is organized as follows. The next section describes the model and its implications for earnings. Section 4.3 discusses the estimation strategy and data. The estimation results are given in Section 4.4 followed by concluding remarks in Section 4.5.

4.2 Equilibrium Search Model

4.2.1 Environment

As noted above our model is a version of the equilibrium search model in Zhang (2012). Here we describe the model noting the differences between immigrant and native workers. In the model, time is continuous and the economy is in a steady state. There are a large number of firms and workers in the labour market, with the population of the workers normalized to 1. All workers are either native or immigrant workers. The measure of immigrant workers is denoted by $\mu \in (0, 1)$.

There are two types of immigrant workers. A type 1 immigrant worker represents an individual new to the country and unfamiliar with its labour market. A type 2 immigrant worker represents an immigrant who has lived in the country for a sufficiently long period and as a result has acquired the same level of knowledge of the local labour market as native workers. A type 1 immigrant worker may become a type 2 immigrant worker over time. This event is modeled as a Poisson process with arrival rate η.

All workers are either unemployed or employed, and search for jobs both on and off the job. If workers are unemployed, they receive b, the flow value of non-market time while unemployed. The arrival of job offers is modeled as a Poisson process. For native-born workers and type 2 immigrant workers, the job offer arrival rate is λ_0 if they are unemployed and λ_1 if they are employed. Because of their lack of knowledge about the labour market, it takes type 1 immigrant workers longer to receive a job offer, on average. For type 1 immigrant workers, the job offer arrival rate is $\alpha_0\lambda_0$ while unemployed and $\alpha_1\lambda_1$ while employed, where $0 < \alpha_i < 1$ for $i = 0, 1$.

All firms have constant returns-to-scale production technologies with labour being their sole input. We assume that both native and immigrant workers are equally productive within a firm. However, the productivity of a worker differs across firms. Specifically, we assume that there are Q types of firms that differ in their marginal productivity of labour. For a type j firm, p_j denotes the per-worker output of the firm, where we assume $p_j < p_{j+1}$ for $j = 1, 2, \ldots, Q-1$.

Each firm posts a wage offer w to attract workers. We assume that it must post the same offer to both native and immigrant workers.[9] The distribution of offers in equilibrium is given by $F(w)$ with support $[\underline{w}, \overline{w}]$, and workers and firms meet each other through random search. This implies that a worker draws a wage offer from $F(w)$ when receiving a job offer. A worker-firm match can be terminated exogenously at rate δ_1, forcing the worker into unemployment.

To ensure that both types of immigrants are present in the steady-state, we assume that workers exit from the labour market permanently according to a Poisson process with arrival rate δ_2. Exiting native workers are replaced by unemployed native workers, while all exiting immigrant workers are replaced by unemployed type 1 immigrant workers.

4.2.2 Workers' Problem

After receiving an offer, the worker decides whether to accept or reject it. Let V_n^u denote the value of unemployment to native workers, and let $V_n^e(w)$ denote the value to native workers of being employed at wage w. Then V_n^u is characterized by the following Bellman equation:

$$rV_n^u = b + \lambda_0 \int_{\underline{w}}^{\overline{w}} \Big[\max\left(V_n^e(x), V_n^u\right) - V_n^u\Big] dF(x) - \delta_2 V_n^u. \tag{4.1}$$

The left-hand side of equation (4.1) represents the payoff to unemployment, while the right-hand side shows how the payoff is derived. The first term on the right-hand side is the

[9]That is immigrants and natives are competing in the same labour market. If firms could post separate offers, then effectively the two markets would be separate. This assumption does tie our hands somewhat in the degree to which the model can generate an immigrant earnings differential. However, we prefer this specification of direct competition both in terms of the assimilation process and in terms of understanding the impact the lower search frictions of the immigrants has on the native workers.

monetary value of non-market time. The remainder is the option value of unemployment due to possible transitions out of unemployment. There are two possible transitions from unemployment. First, at arrival rate λ_0, a worker receives a job offer, and accepts it if it yields a higher value than unemployment. Second, at arrival rate δ_2, a worker permanently leaves the labour market and receives the value of zero thereafter.

Similarly, $V_n^e(w)$ is characterized by the following Bellman equation:

$$rV_n^e(w) = w + \lambda_1 \int_{\underline{w}}^{\overline{w}} \Big[\max\left(V_n^e(x), V_n^e(w)\right) - V_n^e(w) \Big] dF(x) \\ + \delta_1\big(V_n^u - V_n^e(w)\big) - \delta_2 V_n^e(w). \quad (4.2)$$

The left-hand side of equation (4.2) represents the payoff to a native working at wage w. The first term on the right-hand side is the wage flow. The remainder indicates the option value accruing from possible transitions to different states. There are three possible transitions from the current job. First, at arrival rate λ_1, a worker receives a new job offer, and decides whether to accept the offer. Second, at arrival rate δ_1, a worker is separated from the current job. Third, a worker permanently leaves the labour market at rate δ_2.

The optimal search strategy of workers has a reservation wage property. While employed, workers accept any job offer that specifies a higher wage than the current job. While unemployed, workers optimally set a reservation wage, and accept any job that offers a wage above the reservation wage, and reject offers otherwise. Let R_n denote the optimal reservation wage for unemployed native workers. R_n satisfies condition $V_n^e(R_n) = V_n^u$. Together with equations (4.1) and (4.2), the condition yields

$$R_n = b + \int_{R_n}^{\overline{w}} \frac{(\lambda_0 - \lambda_1)\overline{F}(x)}{\rho + \lambda_1 \overline{F}(x)}\, dx, \quad (4.3)$$

where $\overline{F}(w) \equiv 1 - F(w)$ and $\rho \equiv r + \delta_1 + \delta_2$. Notice that type 2 immigrant workers face the same problem as native workers. Thus this group of immigrants acts according to the same reservation wage strategy as native workers.

Type 1 immigrant workers face a slightly different process. Let V_m^u denote the value of being unemployed and $V_m^e(w)$ denote the value of being employed at wage w to type 1 immigrant workers. Then V_m^u is determined by the following equation:

$$rV_m^u = b + \alpha_0\lambda_0 \int_{\underline{w}}^{\overline{w}} \Big[\max(V_m^e(x), V_m^u) - V_m^u \Big] dF(x) + \eta(V_n^u - V_m^u) - \delta_2 V_m^u. \tag{4.4}$$

When compared with equation (4.1), equation (4.4) shows that the option value of unemployment for type 1 immigrants differs from that of natives and type 2 immigrants in two ways. First, as shown in the second term on the right-hand side, the arrival rate of job offers is scaled by α_0. Second, as the third term on the right-hand side shows, a possible change in type also contributes to the option value of unemployment. Similar points are made for $V_m^e(w)$, which satisfies the following Bellman equation:

$$\begin{aligned} rV_m^e(w) = w + \alpha_1\lambda_1 \int_{\underline{w}}^{\overline{w}} &\Big[\max\left(V_m^e(x), V_m^e(w)\right) - V_m^e(w) \Big] dF(x) \\ &+ \eta\Big[\max\left(V_n^e(w), V_n^u\right) - V_m^e(w) \Big] + \delta_1(V_m^u - V_m^e(w)) - \delta_2 V_m^e(w). \end{aligned} \tag{4.5}$$

Again the differences from equation (4.2) are that the job offer arrival rate is scaled by α_1, and the optimal response to a change in type is accounted for in the option value.

When a type 1 immigrant working at wage w becomes a type 2 immigrant worker, it is optimal to quit the current job if $V_n^e(w) < V_n^u$. Since $V_n^u = V_n^e(R_n)$, and $V_n^e(w)$ is increasing in w, such behavior is optimal only if $w < R_n$.

It is straightforward to show that $V_m^e(w)$ is increasing in w, implying that the optimal search strategy of type 1 immigrant workers is a reservation wage strategy as well. Employed type 1 immigrant workers accept any job offer that pays a higher wage than the current job, and unemployed type 1 immigrant workers optimally set a cut-off wage level, denoted by R_m, and accept any job offering a wage that exceeds R_m and reject otherwise.

The following result shows the equation characterizing R_m:

Proposition 4.2.1 *R_m solves the following equation:*

$$R_m = b + \int_{R_m}^{\overline{w}} \frac{(\alpha_0\lambda_0 - \alpha_1\lambda_1)\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)} dx + \int_{R_n}^{\overline{w}} \frac{(\alpha_0\lambda_0 - \alpha_1\lambda_1)\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)} \frac{\eta}{\rho + \lambda_1\overline{F}(x)} dx$$
$$- I(R_m \geq R_n) \int_{R_n}^{R_m} \frac{\rho + \eta + \alpha_0\lambda_0\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)} \frac{\eta}{\rho + \lambda_1\overline{F}(x)} dx,$$

where R_n is given by equation (4.3).

Proof. See Appendix D.

When $\eta = 0$, type 1 immigrant workers never assimilate by changing their types. In this case, their problem is similar to that of native workers except for the differences in job offer arrival rates. That is, R_m satisfies essentially the same nonlinear equation as R_n, since the second and third terms on the right-hand side of equations (4.4) and (4.5) vanish. When $\eta > 0$, in contrast, type 1 immigrant workers need to take into account events that may occur after their type changes in order to set their reservation wage strategies.

The condition determining the ranking between R_m and R_n is given in Lemma 1.

Lemma 4.2.2 *$R_m \leq R_n$ if and only if*

$$\int_{R_n}^{\overline{w}} \left(\frac{\rho + \eta + \lambda_0\overline{F}(x)}{\rho + \lambda_1\overline{F}(x)} - \frac{\rho + \eta + \alpha_0\lambda_0\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)} \frac{\rho + \eta + \lambda_1\overline{F}(x)}{\rho + \lambda_1\overline{F}(x)} \right) dx \geq 0,$$

Proof. See Appendix D.

The above condition is difficult to verify analytically except in a small number of cases. However, the following two examples show that the ranking between R_m and R_n is in general ambiguous. First suppose $\alpha_1 = 1$, then the left-hand side becomes

$$\int_{R_n}^{\overline{w}} \frac{\rho + \eta + (1-\alpha_0)\lambda_0\overline{F}(x)}{\rho + \lambda_1\overline{F}(x)} dx \geq 0,$$

establishing $R_m \leq R_n$. In contrast, if $\alpha_0 = 1$, then the left-hand side of the equation becomes

$$\int_{R_n}^{\overline{w}} \frac{(\alpha_1 - 1)(\rho + \lambda_1\overline{F}(w))\lambda_1\overline{F}(w)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(w)} < 0,$$

concluding $R_n < R_m$.

4.2.3 Firms' Problem and Equilibrium Wage Offer Distribution

The equilibrium wage offer distribution is derived from firms' optimal behaviour. Given the technology, a firm's profit is the product of the per-worker profit margin and the stock of workers in the firm. Letting $l(w)$ denote the steady-state measure of workers available to a firm offering wage w, the steady state profit flow of a type-j firm is written as

$$\pi_j(w) = (p_j - w)l(w).$$

Each firm posts a wage to maximize the steady-state profit flow.

Lemma 1 indicates that in general R_m is not equal to R_n. As pointed out in Mortensen (1990), the equilibrium wage offer distribution may have gaps in the support when workers with different reservation wages operate in a single labour market. To avoid complications resulting from this feature of the wage offer distribution, we follow Zhang (2012) and assume that there is an exogenously set minimum wage in the economy, denoted by w_{min} which is set higher than both R_m and R_n. Given this simplifying assumption, all wage offers will be accepted by all unemployed workers.

The labour stock at a firm paying w can be divided into three components based on the characteristics of workers:

$$l(w) = l_n(w) + l_{m1}(w) + l_{m2}(w)$$

where $l_n(w)$ denotes the steady-state mass of natives working for a firm offering wage w, and for $y \in \{1, 2\}$, $l_{my}(w)$ denotes the steady-state mass of type y immigrants working for a firm offering wage w. These three objects are derived by balancing the steady-state flows generated by job-to-job, unemployment-to-job and job-to-unemployment transitions made by workers. In particular, native workers make these transitions in the same way as modeled in Burdett and Mortensen (1998) with the exception that permanent exit from the labour market is a possible transition. The per-firm native worker stock is therefore given by

$$l_n(w) = \frac{\lambda_0 \delta(\lambda_1 + \delta)}{(\lambda_0 + \delta)(\delta + \lambda_1 \overline{F}(w))^2}(1 - \mu)$$

where $\delta \equiv \delta_1 + \delta_2$. For immigrant workers, the steady-state flow analysis is analogous to that of Zhang (2012) such that the per-firm stocks of type 1 and type 2 immigrant workers are respectively given by[10]

$$l_{m1}(w) = \frac{\alpha_0 \lambda_0 \delta_2 (\delta + \eta)(\delta + \eta + \alpha_1 \lambda_1)}{(\alpha_0 \lambda_0 + \delta + \eta)(\delta_2 + \eta)\left(\delta + \eta + \alpha_1 \lambda_1 \overline{F}(w)\right)^2}\mu,$$

and

$$\begin{aligned} l_{m2}(w) = & \frac{\lambda_0 \eta(\delta + \lambda_1)(\delta(\delta + \eta) + \delta_1 \alpha_0 \lambda_0)}{(\eta + \delta_2)(\delta + \lambda_0)(\eta + \delta + \alpha_0 \lambda_0)\left(\delta + \lambda_1 \overline{F}(w)\right)^2}\mu \\ & + \frac{\lambda_0 \eta \delta_2 \alpha_0 (\delta + \eta)((\delta + \lambda_1)(\delta + \eta + \alpha_1 \lambda_1) - \alpha_1 \lambda_1^2 F(w)^2)}{(\eta + \delta_2)(\eta + \delta + \alpha_0 \lambda_0)\left(\delta + \lambda_1 \overline{F}(w)\right)^2\left(\delta + \eta + \alpha_1 \lambda_1 \overline{F}(w)\right)^2}\mu. \end{aligned}$$

[10] Unlike Zhang (2012), the present paper does not model transitions in and out of nonparticipation of workers. Therefore, the results in this paper are obtained by setting the relevant transition parameters to zero in her model.

Note that $l_n(w)$, $l_{m1}(w)$ and $l_{m2}(w)$ all depend on w only through $F(w)$, the quantile of w in the offer distribution. Because of this property, it is convenient to express the per-firm labour force in terms of the wage quantile when characterizing the equilibrium offer distribution. To this end, define $l^*(y)$ with domain $[0, 1]$ by $l^*(y) = l(F^{-1}(y))$.

Following Mortensen (1990), several properties of the equilibrium hold. First, all firms with the same productivity level earn the same steady-state profit flow in equilibrium. Second, the wage offer from a higher productivity firm should be at least as high as the one from a lower productivity counterpart. Third, the highest wage offer made by a type j firm corresponds to the lowest wage offer made by a type $j + 1$ firm. Fourth, $\underline{w} = w_{min}$. Finally, $F(w)$ is implicitly defined by the following equal profit conditions:

$$(p_j - w_{Lj})l^*(F(w_{Lj})) = (p_j - w)l^*(F(w)) \text{ for } w \in [w_{Lj}, w_{Hj}], \tag{4.6}$$

with $w_{L1} = \underline{w}$ and $w_{Hj} = w_{Lj+1}$ for $j \in \{1, 2, ..., Q - 1\}$.

Let $\gamma_0 = 0$, and for $j = 1, 2, \ldots, Q$, let γ_j denote the fraction of firms whose labour productivity is p_j or less. Given the equilibrium condition, firms with labour productivity equal to or less than p_j offer wages equal to or less than w_{Hj}, and firms with the higher productivity offer wages above w_{Hj}. Therefore $F(w_{Hj}) = \gamma_j$ for $j \in \{1, ..., Q\}$, and equation (4.6) yields

$$(p_j - w_{Hj-1})l^*(\gamma_{j-1}) = (p_j - w_{Hj})l^*(\gamma_j) \text{ for } j \in \{1, \ldots, Q\}. \tag{4.7}$$

For each firm type j, equation (4.7) characterizes the upper- and lower-bounds of possible wage offers made by firms of the given type. This equation can be used to recover p_j from observed wage data. This property of the equilibrium is exploited in estimation.

4.2.4 Implications for Native-Immigrant Wage Gap

Given the offer distribution of wages $F(w)$, the earnings distribution of workers is characterized by the flows generated by the steady-state job-to-job, unemployment-to-job, job-to-unemployment transitions. Let $G_n(w)$ represent the steady-state earnings distribution among native workers. This takes the standard form:

$$G_n(w) = \frac{\delta F(w)}{\delta + \lambda_1 \overline{F}(w)}. \tag{4.8}$$

Analogously, let $G_{m1}(w)$ and $G_{m2}(w)$ denote the steady-state earnings distributions for type 1 and type 2 immigrants, respectively. From Zhang (2012), we have that these distributions are, respectively, given by

$$G_{m1}(w) = \frac{(\delta + \eta) F(w)}{\delta + \eta + \alpha_1 \lambda_1 \overline{F}(w)}. \tag{4.9}$$

and

$$G_{m2}(w) = \frac{(\delta(\eta + \delta) + \alpha_0 \lambda_0 \delta_1)(\delta + \eta + \alpha_1 \lambda_1 \overline{F}(w)) + (\delta + \eta)(\delta + \lambda_0)\delta_2 \alpha_0}{(\delta + \lambda_1 \overline{F}(w)(\alpha_0 \delta_2 + \alpha_0 \lambda_0 + \delta + \eta)(\delta + \eta + \alpha_1 \lambda_1 \overline{F}(w))} F(w). \tag{4.10}$$

Equations (4.8), (4.9) and (4.10) establish that these three distributions can be ranked unambiguously in terms of the first-order stochastic dominance. The following proposition states this result.

Proposition 4.2.3 *$G_{m2}(w)$ first-order stochastically dominates $G_n(w)$. $G_n(w)$ first-order stochastically dominates $G_{m1}(w)$.*

Proof. See Appendix D.

The first-order stochastic dominance of $G_{m2}(w)$ over $G_n(w)$ may not be intuitive since both of the relevant groups share the same search parameters. The reason can be thought of as an age-effect. Since an immigrant worker starts as a type 1 immigrant and later become a type 2 immigrant, type 2 immigrants have, on average, more labour market experience and, therefore, more time to improve their earnings through job-to-job transitions. In contrast, the native worker group includes workers who are new to the labour market and have not had a sufficient time to move up the earnings distribution. Therefore the difference in the average length of time spent in the labour market between these two groups generates the earnings gap between them. Analogously, the native group has on average has longer labour market experience than the type 1 immigrant group. This difference contributes to the ranking between the earnings distributions of these groups. Moreover, the higher job offer arrival rate on the job for the native group also widens the earnings gap between natives and type 1 immigrants.

Even though the native earnings distribution lies between those of type 1 and type 2 immigrant workers, it first-order stochastically dominates the earnings distribution for the whole immigrant population once we account for the steady-state composition of the immigrant population. Proposition 3 establishes this result.

Proposition 4.2.4 *Let $G_m(w)$ denote the earnings distribution for all immigrant workers. Then $G_n(w)$ first-order stochastically dominates $G_m(w)$.*

Proof. See Appendix D.

Finally, we examine the implications of the model for the earnings dynamics for both natives and immigrants by deriving the age profiles of their earnings distributions, and show that the model can generate earnings convergence between native and immigrant workers.

To this end, let $G_n(w; a)$ be the earnings distribution for native workers with labour market experience given by a. Similarly let $G_m(w; a)$ be the earnings distribution for immigrant workers with host country labour market experience given by a. The following proposition establishes that $G_n(w; a)$ and $G_m(w; a)$ have the same limit.

Proposition 4.2.5 *For given* $w \in [\underline{w}, \overline{w}]$,

$$\lim_{a\to\infty} G_n(w;a) = \lim_{a\to\infty} G_m(w;a) = \frac{\delta_1 F(w)}{\delta_1 + \lambda_1 \overline{F}(w)}.$$

Proof. See Appendix D.

The result in Proposition 4 is driven by the gradually increasing share of type 2 immigrants among all immigrants with the same level of (potential) experience. The model predicts that the share of type 1 immigrants among all immigrants with a given labour market experience declines as the time spent in the host country's labour market increases. Therefore, a group of immigrants with sufficiently long labour market experience mostly consists of type 2 immigrants, who behave like native workers. As a result, the earnings distributions converge.

4.3 Estimation

4.3.1 Estimation Procedure

The parameters to be estimated are those governing event arrivals (λ_0, λ_1, δ_1, α_0, α_1 and η), wage cuts ($\underline{w}$, $\{w_{Hj}\}_{j=1}^{Q}$), and those representing the productivity heterogeneity of firms ($\{(p_j, \gamma_j)\}_{j=1}^{Q}$). Estimation is performed by the maximum likelihood procedure developed by Bowlus et al. (1995, 2001) and Bowlus (1998) and then modified by Zhang (2012).[11]

First, we use the lowest and highest wages observed in data for the estimates of $\underline{w}$ and w_{HQ}. Then the following two-stage optimization routine is repeated until the log-likelihood value converges. In the first stage, while fixing the event arrival parameters, we maximize

[11] There are two other parameters in the model: δ_2 and μ. These are calibrated before performing the estimation procedure described in this section. Specifically, μ, the proportion of immigrants among the worker population, is set at 0.212 based on the comparable number for 2001. The arrival rate of permanent exit from the labour market is set at 0.0024, which implies that workers on average spend 35 years in the labour market.

the log-likelihood function by sampling values from the wages earned by native workers and use them for the estimates of $w_{H1}, ..., w_{HQ-1}$. In the second stage, while fixing the wage-cut levels, the log-likelihood function is maximized over the event arrival parameters with a standard iterative optimization routine. Note that every time the objective function is evaluated at a new guess, γ_j and p_j are calculated from the other parameters. Specifically, given equation (4.8), γ_j is calculated by

$$\gamma_j = \frac{(\delta + \lambda_1)\hat{G}_n(\hat{w}_{Hj})}{\delta + \lambda_1 \hat{G}_n(\hat{w}_{Hj})},$$

where $\hat{G}_n(w)$ is the empirical cumulative distribution function of the wages earned by native workers, and p_j is calculated by

$$p_j = \frac{w_{Hj} l^*(\gamma_j) - w_{Hj-1} l^*(\gamma_{j-1})}{l^*(\gamma_j) - l^*(\gamma_{j-1})}.$$

from equation (4.7).

Let θ denote the set of parameters that we aim to estimate, and x_i be a list of variables of individual i. The log-likelihood function is written as

$$\ell(\theta) = \sum_{i=1}^{N} \left[(1 - \chi_i) \ln L_n(\theta; x_i) + \chi_i \ln L_m(\theta; \tau_i, x_i)\right]$$

where $L_n(\theta; x_i)$ and $L_m(\theta; \tau_i, x_i)$, respectively, denote the likelihood contributions for native workers and immigrant workers; χ_i is an indicator variable taking 1 if individual i is an immigrant and 0 otherwise; and τ_i denotes individual i's time since migration. We write $L_m(\theta; \tau_i, x_i)$ as the mixture of likelihood contributions for immigrants of different types. Let $L_{m1}(\theta; x_i)$ and $L_{m2}(\theta; x_i)$ denote the likelihood contribution of type 1 and type 2 immigrants, respectively, and let π_i denote the probability that immigrant i is type 1 at the start of the survey. Then the likelihood contribution for an immigrant worker is given by $L_m(\theta; \tau_i, x_i) =$

$\pi_i L_{m1}(\theta; x_i) + (1 - \pi_i)L_{m2}(\theta; x_i)$, and the log-likelihood function can be rewritten as

$$\ell(\theta) = \sum_{i=1}^{N} \left[(1 - \chi_i)\ln(L_n(\theta; x_i)) + \chi_i \ln\left(\pi_i L_{m1}(\theta; x_i) + (1 - \pi_i)L_{m2}(\theta; x_i)\right)\right]. \tag{4.11}$$

The structural model dictates that the length of time in which an immigrant remains as a type 1 since migration is an exponential random variable with parameter η. Thus $\pi_i = e^{-\eta\tau_i}$. The expressions for $L_n(\theta; x_i)$, $L_{m1}(\theta; x_i)$ and $L_{m2}(\theta; x_i)$ are given in Appendix E.

4.3.2 Identification

Identification of the structural parameters other than α_0, α_1 and η follows Bowlus et al. (1995) with the parameters governing firms' productivity heterogeneity identified from the observed earnings distribution and the native search parameters identified from the relevant duration and transition data. Specifically, λ_0 is identified from unemployment durations of the natives in the data. Job durations and transitions at the end of job spells help to identify λ_1 and δ_1.

The immigrant job search parameters, α_0 and α_1 are identified from differences in unemployment durations and job durations between natives and recently immigrated individuals as well as differences in unemployment rates. η is identified from variation in spell durations of immigrants with respect to years since migration as well as changes in the earnings distribution with respect to years since migration.

4.3.3 Data

To estimate our model we need panel data on both immigrant and native populations. We, therefore, make use of the Canadian SLID, which is a household longitudinal survey containing a wide range of information on the labour market experiences, educational activities and attainment, and demographic characteristics of individuals residing in the country. The survey has several waves, each of which follows respondents for 6 years. The first wave started in 1994, and a new wave was introduced every three years such that two contiguous waves overlap for 3 years. Every January, the survey asks respondents about their labour market activities and/or schooling in the previous year, enabling the construction of their employment histories.

For this paper, the third and fourth waves of the survey are used to construct the estimation sample in order to ensure that it contains a sufficient number of immigrant observations. The third wave covers the period from 1999-2004 and the fourth wave from 2002-2007. Instead of pooling the 9 years of data we construct employment histories only from 2002 on for both waves. This results in a shorter panel for the third wave, but aids in maintaining the stationarity assumption of the model by not introducing large business cycle effects between the late 1990s and early 2000s.[12] We use cross-section sample weights from 2002 to address issues of attrition.

We restrict the estimation sample to male individuals aged between 20 and 55 in the beginning of 2002, and exclude respondents who were institutionalized for more than 6 months or who died during the survey period. Respondents are also excluded from the sample if information on their educational attainment or key demographic characteristics, such as the country of birth or years since migration, are missing. In addition, to keep the sample population as homogeneous as possible, we impose a restriction on the educational attainment of individuals. Specifically, the estimation sample contains only respondents who had some post-secondary schooling or a post-secondary diploma excluding master's degree or above. Finally, our model concerns individuals who are active labour force participants. Thus, we attempt to include only those who are either working or searching for

[12]In practice, however, survey non-response or missing information did not allow us to follow every respondent for the intended period. Rather than excluding these respondents from the estimation sample, if we encounter a problem, we censor the job history at that point.

jobs at any given time. To this end, we exclude individuals who were mainly in school, were in retirement, or were disabled or had a long-term illness.

The above model does not consider schooling decisions and, therefore, only captures immigrant assimilation through post-schooling labour market experience. Thus, it is only relevant for immigrants who completed their schooling before moving to Canada. However, not every immigrant respondent in SLID meets this modeling assumption. In fact, there are ample cases in which we suspect that individuals moved to Canada and then went through schooling activities.[13] Including those individuals in the estimation may, therefore, be problematic as various studies argue that there are differences in labour market outcomes between immigrants who were educated in Canada and those who were not.[14] Although SLID contains information on schooling activities during the survey, it provides less information on schooling undertaken before the survey.[15] To address this complication, we exclude immigrants who migrated to Canada before age 20 as a rough approximation for the desired sampling restriction.

We construct individual labour market histories by first identifying all the jobs held during the survey period. We define a job by an employment relationship with a particular employer, and to be counted as a job spell in our data set an employment relationship needs to last for more than 30 days and have 30 or more usual weekly hours of work.

While the model in this paper does not consider multiple job holding, it is not uncommon to observe individuals who worked for more than one employer simultaneously. To reconcile the difference, when observing an instance of multiple job holding, we assume that the spell that started later did not begin until the one that started earlier ended, and

[13] For a number of immigrant respondents, the reported age at immigration is lower than the typical age of the reported completed schooling level. For example, immigrants reporting a college degree who migrated at age 18 likely obtained their degrees in Canada.

[14] See, for example, Skuterud and Su (2012) and Ferrer and Riddell (2008).

[15] There is a survey question asking respondents where they did most of their elementary and high-school education. However, no such information is available for post-secondary schooling.

adjust the starting date of the latter job accordingly. This treatment of multiple job-holding is common in the literature.[16]

When a job spell is completed, the type of transition made at the end of the spell is based on how long it takes the worker to start a new job. We determine that an individual makes a job-to-job transition if the gap between two jobs is less than 14 days. Otherwise, the gap is treated as an unemployment spell and the transition is recorded as a job-to-unemployment transition.

Wage information is converted into monthly terms based on the reported unit of pay, and converted into real terms with year 2002 as the base year. In order to exclude extreme observations, we trimmed the top 3% and bottom 2% of the earnings distribution.[17]

The above steps yield the final estimation sample of size 3877, with 228 immigrant observations. Of the constructed labour market histories during the survey, the following set of information is used to calculate the likelihood contributions. The first pieces of information that enter into the likelihood are the employment status and residual duration of the first spell. If it is a job spell, the wage earned on the job and, if applicable, the type of transition made at the end of the spell are also included in the likelihood. If the initial spell is a complete unemployment spell, the characteristics of the following job spell enter into the likelihood as well. If the initial spell is a complete job spell, the duration of the next spell is also included in the likelihood but only if it is an unemployment spell.

Table 4.1 shows the sample statistics from the estimation sample.[18] Row 2 shows that the fractions of individuals who were initially unemployed were 0.053 for natives and 0.134 for immigrants, yielding a gap of 0.081. In addition to a higher unemployment rate, immigrants also have a longer mean unemployment duration than natives. The mean job

[16] See, for example, Bowlus et al. (2001).

[17] Trimming is standard in this literature to avoid estimates of reservation wages that are too small and productivity estimates that are too large. See Bowlus et al. (2001).

[18] As noted above the reported statistics are weighted by the 2002 cross-section sample weights.

durations are about 36 months and 30 months for natives and immigrants, respectively, and job spells exhibit high censoring rates for both groups.[19] While two fifths of the completed job spells ended with transitions to a new job for natives, the corresponding number for immigrants was less than a quarter.

The descriptive statistics suggest that immigrants are facing different search frictions with longer unemployment durations and lower job-to-job transition rates. Immigrants also have much higher unemployment rates and shorter job durations. Differences in job offer arrival rates alone may readily account for higher unemployment rates, but may have more difficulty in simultaneously explaining the lower job-to-job transition rates. Immigrants may also be facing higher job destruction rates. Therefore, in what follows, we also estimate a model specification that allows for different job destruction rates for the two groups.[20] This improves the model fit substantially.

There is a sizable gap in monthly earnings between the two groups. As shown in row 8, the mean monthly earnings for immigrant was \$3510.50 as opposed to \$4021.24 for natives, yielding roughly a \$500 earnings gap. Consistent with the job search model with on-the-job search, the mean monthly earnings out of unemployment is lower than the mean monthly earnings for both groups. There is a \$443.73 earnings gap in the mean accepted earnings level between natives and immigrants. This gap is quite large and is incompatible with the modeling assumptions that natives and immigrants face the same offer distribution and accept all offers. However, the standard errors (134.14 and 259.32 for natives and immigrants, respectively) are rather large because of the small numbers of observations.

[19] When reading the values reported in rows 3 to 7, it is important to keep in mind that spells may be censored at different dates for two reasons. First, the estimation sample is from unbalanced panel data. Second, the second spell inevitably has a shorter sample window than the first spell.

[20] The job destruction rate is not allowed to change when immigrants assimilate in the model. That is, the job offer arrival rates change but immigrants continue to face a higher job destruction rate. This is partly done to match the observed data, but also for simplicity in solving the equilibrium wage offer distribution. In addition, it is not clear that changes in the job destruction process should be part of assimilation due to learning about how to search more effectively.

Table 4.1: Summary Statistics from the Estimation Sample

		Natives	Immigrants
1	Number of observations	3649	228
2	Fraction of individuals initially unemployed	0.053	0.134
3	Mean unemployment duration (in month)	5.06	6.84
4	Fraction of censored spells among unemployment spells	0.15	†
5	Mean job duration (in month)	35.59	30.21
6	Fraction of censored spells among job spells	0.75	0.70
7	Fraction of completed job spells ending with a job-to-job transition	0.40	0.23
8	Mean monthly earnings	4021.24	3510.50
9	Mean monthly wage accepted out of unemployment	2907.32	2463.59

† This statistics was not released due to Statistics Canada's disclosure rules regarding confidential survey data.

4.4 Estimation Results

4.4.1 Parameter Estimates

We follow Bowlus et al. (2001) in determining the number of firm types Q. Their method yields $Q = 7$ for our estimation sample. Levels beyond seven yielded no further improvements in the likelihood and productivity parameter estimates that were substantially higher at the top. In addition, the estimated search parameters were stable once the number of firm type was increased to this level.

The estimation results are presented in Tables 4.2, 4.3 and 4.4. Table 4.2 presents the estimated values for the search parameters. The first column shows the estimation results with both natives and immigrants facing a common job destruction rate, and the second column shows the estimation results with natives and immigrants facing job destruction rates of δ_1 and δ_1^m, respectively. The parameter estimates reveal a large difference between

Table 4.2: Parameter Estimates: Event Arrival Rates

Parameter	Common job destruction rate	Separate job destruction rates
$\delta_2{}^*$	0.0024	0.0024
μ^*	0.2120	0.2120
λ_0	0.1556	0.1546
	(0.0094)	(0.0094)
λ_1	0.0089	0.0086
	(0.0010)	(0.0010)
δ_1	0.0047	0.0044
	(0.0003)	(0.0003)
δ_1^m		0.0075
		(0.0017)
α_0	0.6288	0.6471
	(0.1927)	(0.2100)
α_1	0.0735	0.0702
	(0.1146)	(0.1453)
η	0.0055	0.0064
	(0.0008)	(0.0010)
Log-likelihood	−37402.59	−37391.24

Note: Bootstrap standard errors are presented in parentheses.
* Values assigned outside estimation

δ_1 and δ_1^m, though δ_1^m is not precisely estimated. Allowing the job destruction rates to differ between the two groups results in an improvement in the log-likelihood value and fits the observed duration and transition data better. In the following discussion, therefore, we focus on the parameter estimates with separate job destruction rates.[21]

The estimate for λ_0 shows that receiving a job offer is a fairly frequent event for native workers during unemployment with an implied mean unemployment duration of 6.4

[21] Once we allow for different job destruction rates for the two groups, the propositions stated in Section 2 need to be modified because they rely on the assumption of common job destruction rate. However, if $\delta_1 < \delta_1^m$, the result presented in Proposition 3 remains intact and the earnings distribution of natives first-order stochastically dominates the one for immigrants. In contrast, in Proposition 4 $G_n(w,a)$ and $G_m(w,a)$ will have different limits with respect to a. As a result, earnings convergence will be reduced such that the limit of $G_n(w,a)$ will still first-order stochastically dominate the one for $G_m(w,a)$.

Table 4.3: Parameter Estimates: Wage Cuts

Parameter	Common job destruction rate	Separate job destruction rates
w_{min}	1315.80	1315.80
	(3.19)	(2.92)
w_{H1}	2580.00	2580.00
	(433.26)	(452.36)
w_{H2}	3440.00	3440.00
	(479.38)	(473.70)
w_{H3}	4000.00	4000.00
	(425.22)	(436.20)
w_{H4}	4733.44	4733.44
	(500.63)	(525.96)
w_{H5}	5000.00	5000.00
	(585.64)	(637.88)
w_{H6}	5825.00	5825.00
	(759.34)	(793.43)
w_{H7}	8520.79	8520.79
	(0.82)	(0.82)

Note: Bootstrap standard errors are presented in parentheses.

months. In contrast, the job offer arrival rate while employed and job destruction rate are estimated to be very low for natives. The ratio of λ_1 and δ_1 gives a measure of the expected number of job offers during an employment spell and is often used as a measure of search frictions. It is also a measure of how much earnings growth the model will generate as individuals move up the job ladder through on-the-job search. For the natives, this ratio is 1.98, which is higher than the value found for Canada in Bowlus (1998), but by international comparisons is relatively low.[22] Given the high censoring rate of job spells observed in data, these low values can be expected.[23]

[22]The fact that our ratio is higher than that in Bowlus (1998) is not surprising given our estimation sample contains more educated Canadians. However, in both cases the ratio for Canada is low compared to other countries. For example, the value estimated for U.S. males with educational attainment comparable to those in our analysis sample ranges from 1.75 to 4.62 (Bowlus and Seitz (2000) and Flinn (2002)).

[23]In order to examine the effect of the high job censoring rate on the parameter estimates, we estimated

Table 4.4: Parameter Estimates: Firm Productivity Levels and Distribution

Parameter	Common job destruction rate	Separate job destruction rates	Parameter	Common job destruction rate	Separate job destruction rates
p_1	5100.91	5113.64	γ_1	0.3598	0.3627
	(468.10)	(496.40)		(0.1141)	(0.1198)
p_2	5517.98	5533.66	γ_2	0.6026	0.6057
	(545.92)	(594.65)		(0.1146)	(0.1121)
p_3	6384.91	6440.80	γ_3	0.7281	0.7306
	(1031.13)	(1161.29)		(0.0768)	(0.0762)
p_4	7645.26	7728.36	γ_4	0.8462	0.8478
	(1703.17)	(1928.94)		(0.0572)	(0.0573)
p_5	9070.35	9188.67	γ_5	0.8769	0.8783
	(2716.80)	(3292.69)		(0.0410)	(0.0436)
p_6	11908.38	12100.86	γ_6	0.9354	0.9362
	(5897.89)	(6736.32)		(0.0260)	(0.0267)
p_7	24989.68	25553.25	γ_7	1	1
	(8838.72)	(9672.67)			

Note: Bootstrap standard errors are presented in parentheses.

Although not precisely estimated, the point estimates for α_0 and α_1 reveal differences in the job search process between natives and newly-arrived immigrants. The estimates for α_0 and λ_0 together imply that the job offer arrival rate for immigrants is 0.1002 giving an unemployment duration of 10.0 months for type 1 immigrants. The estimate for α_1 implies that the job offer arrival rate for employed type 1 immigrants is less than one tenth the native job offer arrival rate on the job and one tenth their own job destruction rate. All of this suggests that type 1 immigrants face substantial search frictions while employed and are, therefore, much more likely to have their jobs end with transitions to unemployment than are natives. Once immigrants acquire the same search parameters as natives, their job offer arrival rate while employed slightly exceeds their job destruction rate giving a ratio of

the model restricting the native worker sample to age 20 to 35, which had a lower job censoring rate than the original sample. We also attempted a different estimation strategy, used in Bowlus and Seitz (2000), that omits job duration data and relies on the initial unemployment rates and the transition data at the end of job spells to identify λ_1 and δ_1. Although unreported here, in both cases, the estimates yielded higher values not only for λ_1 but also for δ_1 resulting in ratios and earnings growth predictions that were hardly altered.

1.17.

The estimate for η is 0.0064, which implies that it takes newly arrived immigrants 13 years to acquire the native search parameters. Interpreted slightly differently, 47% of a cohort of immigrants who immigrated 10 years previously have acquired native search parameters.

The finding that search assimilation for immigrants takes, on average, more than a decade is in line with some of the previous search assimilation findings. For example, based on their estimation results, Daneshvary et al. (1992) argue that immigrants reach "information parity" with natives after about 12 years since migration in the United States.[24] In contrast, based on their duration analysis of unemployment, Frijters et al. (2005) extrapolate that it takes immigrants more than 40 years to attain the same hazard rate out of unemployment as their native peers in Britain.

Studies taking a human capital approach to assimilation also often find that it takes immigrants decades to catch up with natives, though a common feature is a much faster initial rate of catch up. Skuterud and Su (2012), for example, report results for a sample of recent Canadian immigrants (arrival cohort 1990-2002). These were compared with similarly aged native workers. In their preferred specification the initial wage gap of 0.29 log points was more than halved after 8 years, declined further to year 13, but then remained roughly constant at 0.09 log points thereafter. Skuterud and Su argue that this pattern of strong decreasing relative returns to host country experience reflects what might be expected from "language acquisition or acculturation processes" (p.1124). However, our results indicate that search assimilation may be responsible for much of this convergence and that more conventional human capital models may overestimate the role of human capital assimila-

[24] Their view is that the amount of information available to workers searching for jobs is related to the ratio between the actual and potential earnings where potential earnings means the upper support of the wage offer distribution. They found that this ratio was below natives' level for newly arrived immigrants, but it caught up to the same level after roughly 12 years since migration.

tion in not taking into account search assimilation.

Tables 4.3 and 4.4 show the estimates related to the firm productivity distribution and the resulting wage cuts. The productivity distribution is right skewed with the lowest two levels accounting for the majority of the productivity distribution. The implied average monthly productivity level is $7526.56. The large values for the highest productivity levels, needed to meet the equal-profit condition at the upper end of the wage distribution, are a common outcome of this model.

4.4.2 Model Fit

We examine how well the model fits the observed data by comparing the summary statistics reported in Table 4.1 with model predictions. When predicting the moments of the duration and transition data, it is important to account for the fact that our estimation sample is unbalanced panel data and spells can be censored at different dates. To control for this issue we simulate a sample of a large number of job histories matching the survey response outcomes in the estimation sample. One exception to this is that the predicted values for the earnings outcomes given in row 8 are obtained by numerically calculating the mean of the offer distribution rather than from the simulations. The results are presented in Table 4.5.

The model matches the duration and transition data of natives and immigrants well overall. It underpredicts the unemployment rates for both groups, and somewhat overpredicts the unemployment durations. Unfortunately the model cannot match both of these moments since to match the first the job offer arrival rate while unemployed needs to be lower and to fix the second it needs to be higher.

Not surprisingly, as shown in Figure 4.1, the predicted distribution of natives' earnings fits the observed distribution very closely. The model produces an earnings gap between

Table 4.5: Predicted Moments from the Estimation Result with Separate Job Destruction Rates for Natives and Immigrants

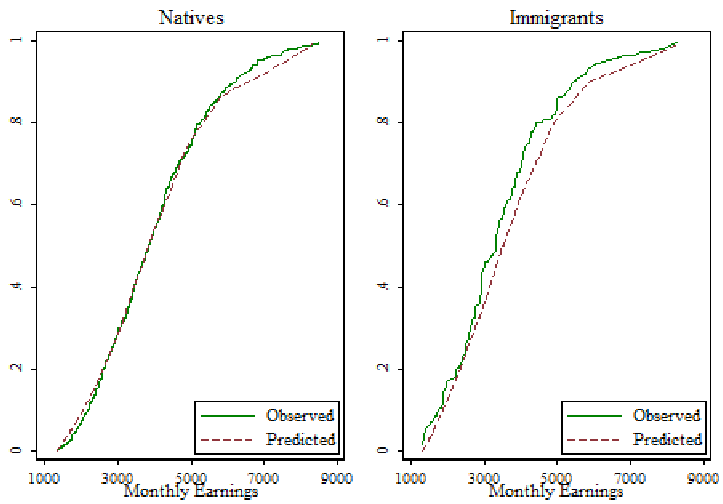
		Natives		Immigrants	
		Observed	Predicted	Observed	Predicted
1	Fraction of individuals initially unemployed	0.053	0.042	0.134	0.072
2	Mean unemployment duration (in month)	5.06	5.83	6.84	7.37
3	Fraction of censored spells among all unemployment spells	0.15	0.10	†	0.14
4	Mean job duration (in month)	35.59	35.87	30.21	32.33
5	Fraction of censored spells among all job spells	0.75	0.74	0.70	0.69
6	Fraction of completed job spells ending with a job-to-job transition	0.40	0.41	0.23	0.20
7	Mean monthly earnings	4021.24	4063.54	3510.50	3676.74
8	Mean monthly wage accepted out of unemployment	2907.73	3341.87	2463.59	3341.87

† This statistics was not released due to Statistics Canada's disclosure rules regarding confidential survey data.

natives and immigrants of approximately \$390, which captures three quarters of the observed gap. The same figure shows that the predicted earnings distribution for immigrants lies slightly to the right of the observed distribution reflecting the unexplained portion of the observed earnings gap.

Row 8 of Table 4.5 shows that the model is having a difficult time fitting the offer distribution, substantially overpredicting it. The offer distribution is identified from earnings observations accepted out of unemployment and the number of such observations is modest in the data. Therefore, fitting the data in this dimension does not seem to have influenced the estimation substantially. The difficulty in capturing enough difference between the earnings and offer distributions reflects the estimated low value of λ_1/δ_1, and also points to the model's problem in generating sufficient earnings growth. This suggests that the model is missing important earnings growth mechanisms. The most obvious factor

Figure 4.1: Observed and Predicted Earnings Distributions for Natives and Immigrants



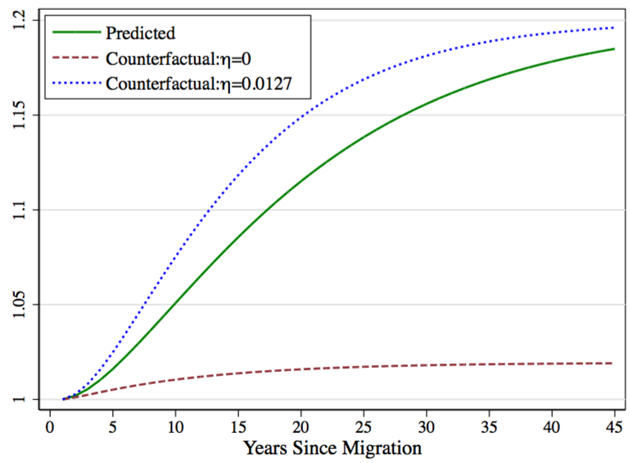
omitted by the model is human capital accumulation, and enriching the model in this direction is an interesting avenue for future work.[25] Finally, the observed gap in accepted wages out of unemployment between natives and immigrants may reflect productivity differences between the two groups. Therefore, it may be important to investigate whether there are separate labour markets for natives and immigrants.[26]

4.4.3 Implications for Earnings Assimilation

The estimation results imply that it takes newly arrived immigrants, on average, 13 years to acquire the native search parameters. We examine the implications of this estimated job search assimilation process for immigrants' life cycle earnings growth. In Figure 4.2, the solid-lined curve shows the predicted mean monthly earnings profile of immigrants since migration, relative to the earnings level in the first year since migration. On average, immigrants are predicted to experience about 5.1% earnings growth in the first 10 years

[25]The addition of human capital can substantially complicate the equilibrium solution to the model.
[26]This difference may also be aggravated by our assumption of a common minimum wage.

Figure 4.2: Predicted Immigrants' Earnings Growth



since migration, and about 17.8% earnings growth over 40 years since migration. The earnings profile is S-shaped, originally exhibiting slower earnings growth because of the lower rates of job offer arrival when they are new to the host country.

Indeed, the estimates reveal that newly arrived immigrants search under a very low job offer arrival rate while employed. This limits their chances of finding better paying jobs. The broken lined curve in Figure 4.2 shows a counterfactual scenario in which η is set to 0 and therefore immigrants do not experience job search assimilation. Under this scenario, the same 40 year period produces barely 1.9% earnings growth, showing that nearly 90% of the earnings growth for immigrants is attributable to job search assimilation.

In contrast, if immigrants could search as effectively as natives sooner than the estimates indicate, they would experience faster earnings growth. The dotted lined curve in the same figure shows a counterfactual scenario in which η is doubled to 0.0127, halving the average length of time needed to spend in the host country before acquiring the native job search efficiency. The counterfactual earnings profile shows faster and larger earnings growth, with 7.5% and 19.3% increases in 10 and 40 year periods since migration, respectively.

Figure 4.3: Predicted Immigrant Earnings Relative to Native Earnings

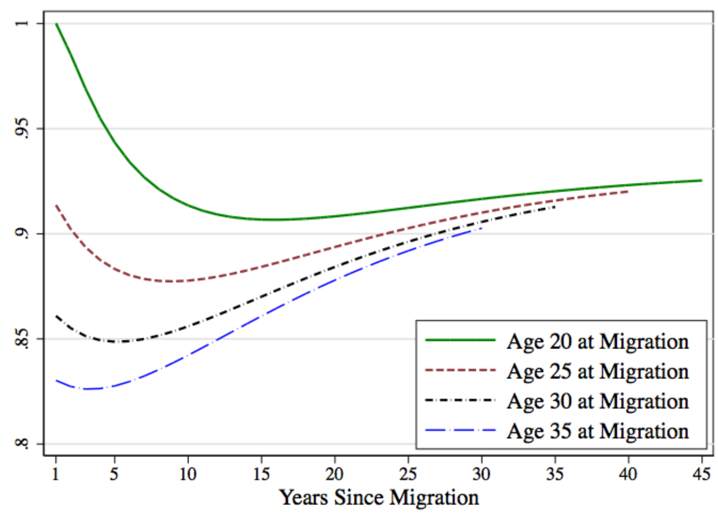
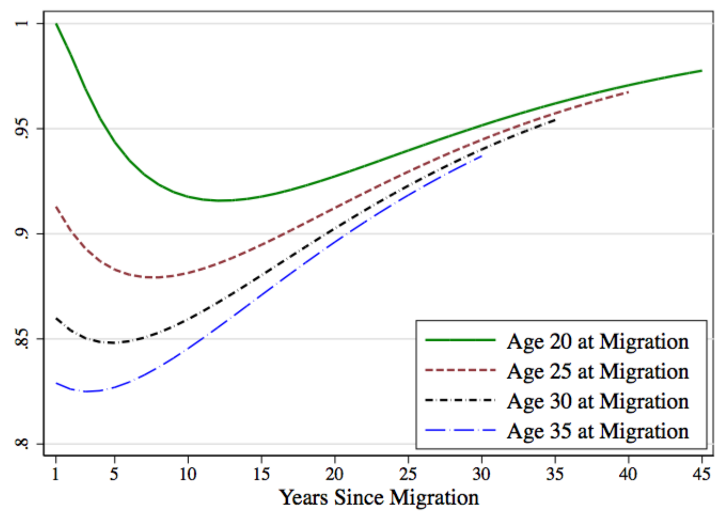


Figure 4.3 shows predicted earnings profiles of immigrants entering the host country at different ages. These profiles are measured relative to their native counterparts, with an assumption that native workers enter the labour market at age 20. The earnings are followed until immigrants reach age 65. For each age group, the initial earnings gap reflects the natives' advantage of having more time operating in the host country labour market. For any age group in Figure 4.3, the relative earnings decline initially for two reasons. First, newly arrived immigrants search with lower job search efficiency. Second, natives are experiencing robust earnings growth during their early years in the labour market. As immigrants age, their earnings level converges to a level about 8% below the native earnings, failing to achieve the earnings parity. This failure of earnings convergence is because of the immigrants' higher job destruction rate. The same exercise is conducted assuming that immigrants face the native job destruction rate, and this counterfactual experiment produces the relative earnings profile in Figure 4.4. In this counterfactual scenario, immigrants close the earnings gap more than in the previous case. However, even as they turn 65 they have yet to achieve earnings parity.

Figure 4.4: Counterfactual Immigrant Earnings Relative to Native Earnings

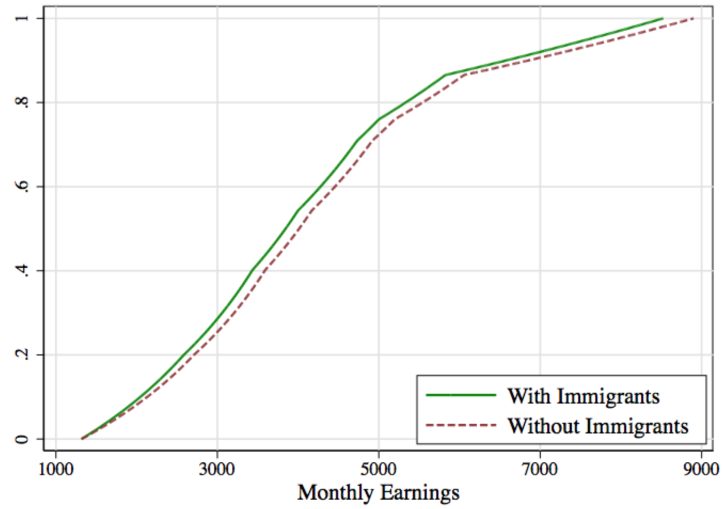


It is also interesting to ask how natives are affected by having to compete with immigrants in the same labour market. The presence of new immigrants produces an equilibrium effect, which reflects a change in firms' wage posting strategies. They search at lower job offer arrival rates, resulting in an increase in the fraction of workers with less propensity to make a job-to-job transition. This leads to an increase in firms' monopsony power in the model, allowing them to post lower wages and shift the offer distribution to the left in equilibrium.[27] Without immigrants, the firms' monopsony power would be reduced, and the equilibrium wage offer distribution would shift to the right.[28] To look at this effect on natives' earnings, we solve the equilibrium with $\mu = 0$, and compare the native earnings distribution under this equilibrium with the one under the estimated parameters. These distributions are presented in Figure 4.5. The mean difference in these two distributions are \$175.88, implying a 4% reduction in the mean earnings of natives due to the presence of

[27]See Burdett and Mortensen (1998) for a discussion on how workers' likelihood to make a job-to-job transition as opposed to a job-to-unemployment transition affects the monopsony power of firms.

[28]We do not account for the effect of the immigrant labour force on native job offer arrival rates and any resulting equilibrium effects.

Figure 4.5: Predicted Native Earnings Distributions



immigrants in the labour market.

4.5 Concluding Remarks

Immigrant assimilation is a major issue in many countries. There is a very large literature that studies assimilation primarily through a human capital framework. While a variety of these studies, going back to Chiswick (1978), refers to immigrants accumulating host country specific knowledge as well as skills following migration, the accumulation of knowledge of how the host country labour market works and how to search efficiently has received relatively little attention. In this paper, we use a search model to study assimilation via this potentially important host-country specific knowledge. Specifically, we present and estimate an equilibrium search model of immigrants operating in the same labour market as natives using Canadian panel data.

Assimilation via acquisition of knowledge of how the host country labour market works and how to search efficiently in it takes place in the model by having immigrants initially face a (potentially) lower arrival rate of job offers, and allowing them to acquire the same job offer arrival rate according to a stochastic process. The estimation results show substantial differences in job offer arrival rates between natives and newly arrived immigrants, as well as a difference in the job destruction rate between natives and immigrants. These differences are able to account for three quarters of the observed earnings differential between natives and immigrants.

The results also imply that it takes immigrants, on average, 13 years to acquire the same search parameters as natives. The parameter estimates reveal that newly arrived immigrants have a hard time generating earnings growth because of their very low job offer arrival rate while employed, with counterfactual exercises indicating that the vast majority of the earnings growth for immigrants is due to job search assimilation. Moreover, if the time needed to acquire the native job search process were halved, immigrants would experience a faster and larger earnings growth. This has important implications for policy initiatives to encourage immigrant assimilation.

Although the model is able to fit various dimensions of the observed data well, it is at odds with the observed data in some dimensions. Particularly, the model is not able to capture the difference in accepted wages out of unemployment between natives and immigrants and it has difficulty generating sufficient earnings growth. These two findings may point to productivity differences and the role of human capital accumulation, and enriching the model in this dimension is an interesting avenue for future research. In particular, given the large previous literature emphasizing the role of human capital accumulation in immigrant assimilation, it is important to understand the relative roles of human capital and search in this process.

Finally, given the modest number of immigrant observations in the estimation sample, the estimates pertaining to the immigrant job search parameters are not precisely estimated. In addition, it was not possible to allow for initial job offer arrival rates to depend on potentially relevant factors, such as the degree of similarity between the labour markets in the source and host countries. An important next step is to incorporate heterogeneity in initial Canadian Labour market knowledge through the use of alternative data sources such as the LSIC.

Bibliography

Abbott, Michael, and Charles Beach. 2011. Immigrant earnings differences across admission categories and landing cohorts in Canada. CLSRN Working paper No.81.

Baker, Michael, and Dwayne Benjamin. 1994. The performance of immigrants in the Canadian labor market. *Journal of Labor Economics* 12(3):369–405.

Borjas, Geroge J. 1999. The economic analysis of immigration. In *Handbook of Labor Economics*, eds. Orley C. Ashenfelter and David Card, vol. 3A, 1697 – 1760. Elsevier.

Bowlus, Audra J. 1997. A search interpretation of male-female wage differentials. *Journal of Labor Economics* 15(4):625–57.

———. 1998. U.S.-Canadian unemployment rate and wage differences among young, low-skilled males in 1980s. *Canadian Journal of Economics* 31(2):437–464.

Bowlus, Audra J., and Zvi Eckstein. 2002. Discrimination and skill differences in an equilibrium search model. *International Economic Review* 43(4):1309–45.

Bowlus, Audra J., Nicholas M. Kiefer, and George R. Neumann. 1995. Estimation of equilibrium wage distributions with heterogeneity. *Journal of Applied Econometrics* 10(S1):S119–S131.

———. 2001. Equilibirum search models and the transition from school to work. *International Economic Review* 42(2):317–344.

Bowlus, Audra J., and Shannon Seitz. 2000. Search friction in the U.S. labor market: equilibrium estimates from the PSID. In *Panel Data and Structural Labour Market Models*, eds. H Bunzel, B. J. Christensen, N. M. Keifer, and D. T. Mortensen. North Holland.

Burdett, Kenneth, and Dale T. Mortensen. 1998. Wage differentials, employer size, and unemployment. *International Economic Review* 39(2):257 – 73.

Chassamboulli, Andri, and Theodore Palivos. 2010. Give me your tired, your poor, so I can prosper: immigration in search equilibrium. Working paper.

Chiswick, Barry R. 1978. The effect of Americanization on the earnings of foreign born men. *Journal of Political Economy* 86(5):897–921.

Daneshvary, Nasser, Henry W. Herzog, Richard A. Hofler, and Alan M. Schlottmann. 1992. Job search and immigrant assimilation: an earnings frontier approach. *The Review of Economics and Statistics* 74(3):482–92.

Dechief, Diane, and Philip Oreopoulos. 2012. Why do some employers prefer to interview Matthew but not Samir? New evidence from Toronto, Montreal and Vancouver. Working paper.

Ferrer, Ana, and W. Craig Riddell. 2008. Education, credentials, and immigrant earnings. *The Canadian Journal of Economics* 41(1):186–216.

Flinn, Christopher J. 2002. Labour market strcture and inequality: A comparision of Italy and the U.S. *Review of Economic Studies* 69:611–645.

Frijters, Paul, Michael A. Shields, and Stephen Wheatley Price. 2005. Job search methods and their success: a comparison of immigrants and natives in the U.K. *The Economic Journal* 115:F359–76.

Goel, Deepti, and Kevin Lang. 2012. Social ties and the job search of recent immigrants. Working paper.

Green, David, and Christopher Worswick. 2010. Entry earnings of immigrant men in Canada: The role of labour market entry effects and returns to foreign experience. In *Canadian Immigration: Economic Evidence for a Dynamic Policy Environment*, eds. T. McDonald, E. Ruddick, A. Sweetman, and C. Worswick, 77–110. McGill-Queen's University Press.

Gupta, Nabanita Datta, and Lene Kromann. 2013. An equilibrium search model of the labor market entry of second-generation immigrants and ethnic danes. Working paper.

Hum, Derek, and Wayne Simpson. 2000. Closing the wage gap: Economic assimilation of Canadian immigrants reconsidered. *Journal of International Migrations and Integration* 1(4):427–441.

———. 2004. Reinterpreting the performance of immigrant wagess from panel data. *Empirical Economics* 29(1):129–147.

Lessem, Rebecca, and Carl Sanders. 2012. The immigrant-native wage gap in the United States. Tepper School of Business Paper 1451.

Liu, Xiangbo. 2010. On the macroeconomic and welfare effects of illegal immigration. *Journal of Economic Dynamics and Control* 34(12):2547–67.

Mortensen, Dale T. 1990. Equilibrium wage distribuions: a synthesis. In *Panel Data and*

Labor Market Studies, eds. Joop Hartog, Geert Fidder, and Jules Theeuwes, 279–96. Elsevier.

Munshi, Kaivan. 2003. Networks in the modern economy: Mexican migrants in the U.S. labor market. *Quarterly Journal of Economics* 118(2):549–599.

Oreopoulos, Philip. 2011. Why do skilled immigrants struggle in the labor market? A field experiment with thirteen thousand resumes. *American Economic Journal: Economic Policy* 3(4):148–71.

Ortega, Javier. 2000. Pareto-improving immigration in an economy with equilibrium unemployment. *The Economic Journal* 110:92–112.

Skuterud, Mikal, and Mingcui Su. 2012. The influence of measurement error and unobserved heterogeneity in estimating immigrant returns to foreign and host-country sources of human capital. *Empirical Economics* 43(3):1109–1141.

Sweetman, Arthur, and Casey Warman. 2013. Canada's immigration selection system and labour market outcomes. *Canadian Public Policy* 39:S141–S164.

Xue, Li. 2010. A comprehensive look at the employment experience of recent immigrants during their first four years in Canada. In *Canadian Immigration: Economic Evidence for a Dynamic Policy Environment*, eds. T. McDonald, E. Ruddick, A. Sweetman, and C. Worswick, 11–40. McGill-Queen's University Press.

Zhang, Yahong. 2012. A search interpretation of the family gap. *Labour Economics* 19(2):186–197.

Chapter 5

Conclusion

This thesis studies job and occupational transition patterns of different groups of workers. The second chapter examines the effect of learning on young workers' job search and transition patterns by developing and estimating a model unifying learning, skill accumulation and directed search processes. The estimation results using U.S. data show large differences in search frictions, returns to skills, skill acquisition rates, and learning opportunities across occupations. Simulation exercises show that while learning can have a sizeable effect on young workers' job search, changes in search effort allocation due to learning do not result in a comparable effect in occupational transition outcomes because of search frictions. In particular, high search frictions associated with the white-collar occupation limit opportunities to receive signals that have a large impact on workers' search decisions. As a result, they reduce the effect of learning in workers' early career occupational mobility.

The third chapter studies the role of skill depreciation in explaining the stylized life-cycle employment transition patterns of British women, which highlight pervasive part-time employment that follows occupational downgrading at the start of motherhood. A dynamic model of employment transitions between full-time work, part-time work and nonemployment is developed, and numerical exercises with the model find reasonable parameter values that can produce these patterns as a result of workers' choices over hours and occupations in the life cycle.

The fourth chapter develops and estimates an equilibrium search model of immigrants operating in the same labour market as natives, where newly arrived immigrants have lower job offer arrival rates than natives but can acquire the same arrival rates according to a stochastic process. Using Canadian panel data, substantial differences in job offer arrival and destruction rates between natives and immigrants are found that are able to account for three quarters of the observed earnings gap. The estimates imply that immigrants take, on average, 13 years to acquire the native search parameters. Counterfactual exercises show that the vast majority of earnings growth immigrants experience after migration is due to the job search assimilation process. This is due to the large difference in on-the-job job offer arrival rates between new immigrants and natives.

Appendix A

A Case Allowing for Correlation between the Skill Acquisition Probabilities and the Initial Skill Levels

In this appendix, I consider a case where a particular form of correlation between the skill acquisition probability and the initial skill level is allowed. While maintaining the assumption that the initial skill level takes on two possible values, $\overline{s}_\ell$ or $\underline{s}_\ell$, I now assume that the distribution of the initial skill level is given by the following conditional distribution:

$$Pr(s_0 = \overline{s}_\ell \mid \theta_\ell) = m_\ell + n_\ell \theta_\ell, \tag{A.1}$$

where $0 \leq m_\ell \leq 1$, $0 \leq n_\ell \leq 1$, and $0 \leq m_\ell + n_\ell \leq 1$. This specification allows the initial skill levels and skill acquisition probabilities to be positively correlated. Note that a case with $n_\ell = 0$ is equivalent to the original model.

For each skill $\ell = 1, 2$, two initial belief distributions emerge for the skill acquisition probabilities because there are two possible initial skill levels. Suppose that a worker's initial level is $\overline{s}_\ell$. By Bayes' rule, the posterior distribution of θ_ℓ conditional on $\overline{s}_\ell$ is given by

$$\Pr(\theta_\ell\,|s_0 = \overline{s}_\ell) = \frac{(m_\ell + n_\ell\theta_\ell)\theta_\ell^{p_\ell-1}(1-\theta_\ell)^{q_\ell-1}}{\int_0^1 (m_\ell + n_\ell\theta)\theta^{p_\ell-1}(1-\theta)^{q_\ell-1}d\theta}. \tag{A.2}$$

In contrast, if a worker's initial skill level is $\underline{s}_\ell$, then the initial belief distribution is given by

$$\Pr(\theta_\ell\,|s_0 = \underline{s}_\ell) = \frac{(1 - m_\ell - n_\ell\theta_\ell)\theta_\ell^{p_\ell-1}(1-\theta_\ell)^{q_\ell-1}}{\int_0^1 (1 - m_\ell - n_\ell\theta)\theta^{p_\ell-1}(1-\theta)^{q_\ell-1}d\theta}. \tag{A.3}$$

It can be shown that the distributions in equations (A.2) and (A.3) are mixtures of two Beta distributions.

Proposition A.0.1 *Define probability density functions $\phi(\theta, p, q, m, n)$ and $\psi(\theta, p, q, m, n)$, respectively, by*

$$\phi(\theta, p, q, m, n) \equiv \frac{m(p+q)f(\theta; p, q) + npf(\theta; p+1, q)}{m(p+q) + np}$$

and

$$\psi(\theta; p, q, m, n) \equiv \frac{(1-m-n)(p+q)f(\theta; p, q) + nqf(\theta; p, q+1)}{(1-m-n)(p+q) + nq},$$

where $f(\theta; p, q)$ is the density function of the Beta distribution with parameters p and q. Then the initial belief distributions shown in equations (A.2) and (A.3) are respectively rewritten as $Pr(\theta_\ell|s_0 = \overline{s}_\ell) = \phi(\theta_\ell; p_\ell, q_\ell, m_\ell, n_\ell)$ and $Pr(\theta_\ell|s_0 = \underline{s}_\ell) = \psi(\theta_\ell; p_\ell, q_\ell, m_\ell, n_\ell)$ for $\ell = 1, 2$.

Proof. Simple algebraic manipulations on equations (A.2) and (A.3) establish the claim. ∎

Both $\phi(\cdot)$ and $\psi(\cdot)$ are mixtures of two Beta distributions with mixing proportions given by functions of m_ℓ, n_ℓ, p_ℓ and q_ℓ. On the one hand, if $n_\ell = 0$, both of them are reduced to the Beta distribution with parameters p_ℓ and q_ℓ. In this case, the initial skill level does not provide any new information regarding θ_ℓ because they are uncorrelated. As a result, the initial belief distribution coincides with the population distribution of θ_ℓ. On the other hand, if $m_\ell = 0$ and $n_\ell = 1$, then the initial skill level has the same information content regarding θ_ℓ as a skill acquisition outcome. In this case, $\phi(\cdot)$ is the Beta distribution with parameters $p_\ell + 1$ and q_ℓ, while $\psi(\cdot)$ is the Beta distribution with parameters p_ℓ and $q_\ell + 1$.

It has been shown that mixtures of the Beta distributions are natural conjugates for the Bernoulli distribution.[1] This is also true in the present case, and the following result is presented to provide the Bayesian updating rules in this case.

Proposition A.0.2 *Consider a Bernoulli random variable with parameter θ and suppose the true value of θ is unknown. If the prior distribution of θ is given by $\phi(\theta; p_\ell, q_\ell, m_\ell, n_\ell)$, then the posterior distribution is given by $\phi(\theta; p_\ell + 1, q_\ell, m_\ell, n_\ell)$ in the event of success, and $\phi(\theta; p_\ell, q_\ell + 1, m_\ell, n_\ell)$ in the event of failure. Similarly, if the prior distribution is $\psi(\theta; p_\ell, q_\ell, m_\ell, n_\ell)$. Then the posterior distribution is given by $\psi(\theta; p_\ell + 1, q_\ell, m_\ell, n_\ell)$ in the event of success, and $\psi(\theta; p_\ell, q_\ell + 1, m_\ell, n_\ell)$ in the event of failure.*

Proof. Application of the Bayes' rule and straightforward algebraic steps prove the claim. ∎

The above statement means that the workers' learning process can be captured just by incrementing either the first or second parameter of $\phi(\cdot)$ or $\psi(\cdot)$, depending on the outcome of skill acquisition. In the event of success, the parameter p_ℓ is incremented by 1, while in the event of failure, q_ℓ is incremented by 1. Therefore the Bayesian updating rule in this model is as simple as in the standard Bernoulli learning model.

Using these two mixture distributions instead of the Beta distribution in the workers' dynamic problem only requires the change in equation (2.12). Specifically, the equation

[1] See, for example, Bernardo and Smith (1994).

will be replaced by

$$V_t^k(S,r) = r + y_k(s_1,s_2) + \frac{1}{1+\rho}\Big[(1-\omega_{k1}-\omega_{k2})\max[W_{t+1}^k(S,r), V_{t+1}^0(S)] \\ + \sum_{\ell=1}^{2}\omega_{k\ell}\sum_{a=0}^{1} p_\ell(a;s_\ell^0)\max[W_{t+1}^k(S_{k\ell}(a),r), V_{t+1}^0(S_{k\ell}(a))]\Big], \quad \text{(A.4)}$$

where $p_\ell(a; s_\ell^0)$ is given by

$$p_\ell(a;s_\ell^0) = \overline{\theta}_\ell(\eta_\ell, \upsilon_\ell, s_\ell^0)^a\,[1-\overline{\theta}_\ell(\eta_\ell,\upsilon_\ell,s_\ell^0)]^{1-a},$$

with $\overline{\theta}_\ell(\eta_\ell, \upsilon_\ell, s_\ell^0)$ denoting the mean of the belief distribution.[2] This value is given by

$$\overline{\theta}_\ell(\eta_\ell,\upsilon_\ell,s_\ell^0) = \eta_\ell + \frac{n_\ell\,\eta_\ell(1-\eta_\ell)\upsilon_\ell}{(m_\ell+n_\ell\eta_\ell)^{\mathbb{I}(s_\ell^0=\overline{s}_\ell)}(m_\ell+n_\ell\eta_\ell-1)^{\mathbb{I}(s_\ell^0=\underline{s}_\ell)}}.$$

Table A.1 presents the parameter estimates of the augmented model allowing the correlation. The parameters controlling the correlation between s_ℓ^0 and θ_ℓ for both $\ell = 1,2$ are estimated to be just around 0.02 and statistically insignificant at the 5% level. Furthermore, the likelihood ratio test does not reject the hypothesis of no correlation.

Bibliography

Bernardo, José M., and Adrian F.M. Smith. 1994. *Bayesian Theory*. John Wiley & Sons.

[2]Workers' initial skill levels (s_1^0, s_2^0) enter $p_\ell(a, s_\ell^0)$ because their initial skill levels influence their initial belief for (θ_1, θ_2).

Table A.1: Parameter Estimates - Skill Acquisition, Initial Skill Level

	Skill (ℓ)			Skill (ℓ)	
Parameter	1	2	Parameter	1	2
$\gamma_{1\ell} \times 100$	9.7683	0.0000	p_ℓ	1.4749	11.4096
	(1.3633)	(1.1996)		(0.3956)	(4.3484)
$\gamma_{2\ell} \times 100$	0.6423	1.1930	q_ℓ	6.5627	10.7274
	(15.0685)	(0.1358)		(1.9365)	(4.1178)
$\gamma_{3\ell} \times 100$	0.0323	0.8927	$\underline{s}_\ell$	4.0180	2.6952
	(0.0164)	(0.0040)		(0.6410)	(0.6221)
$\omega_{1\ell}$	0.6805	0.1520	$\overline{s}_\ell$	5.3316	3.0351
	(0.0452)	(0.0366)		(0.6238)	(0.6256)
$\omega_{2\ell}$	0.0254	0.9005	m_ℓ	0.9691	0.3481
	(0.0085)	(0.0170)		(0.0127)	(0.1093)
$\omega_{3\ell}$	0.3945	0.4902	n_ℓ	0.0238	0.0223
	(0.0257)	(0.0278)		(0.0135)	(0.1991)

	Occupation (k)				
Parameter	1	2	3	Parameter	
λ_k^0	0.0097	0.0289	0.0353	α	0.8183
	(0.0014)	(0.0038)	(0.0050)		(0.0171)
λ_k^1	0.0036	0.0209	0.0220	b	6.7831
	(0.0007)	(0.0035)	(0.0036)		(0.6139)
δ_k	0.0135	0.0458	0.0596	σ_u	0.3032
	(0.0000)	(0.0034)	(0.0041)		(0.0015)
$\delta_k^{'}$	0.0135	0.0138	0.0383	π	0.5894
	(0.0000)	(0.0025)	(0.0047)		(0.0264)
$\delta_k^{''}$	0.0135	0.0138	0.0383	ε	1.6260
	(0.0018)	(0.0022)	(0.0045)		(0.1667)
β_{k1}	0.7243	0.5881	1.0000†	χ	0.2586
	(0.0998)	(0.0480)			(0.0424)
β_{k2}	1.5315	1.3447	1.0000†		
	(0.1452)	(0.0807)			
μ_k	0.1477	1.3836	0.0000†		
	(0.2747)	(0.3444)			
σ_k	0.2665	0.2185	0.2562		
	(0.0232)	(0.0111)	(0.0126)		

Loglikelihood value: $-10.682.34$
Asymptotic standard errors in parentheses
† normalized value

Appendix B

Occupation Classification

Suppose that the population of occupations is comprised by $K > 0$ groups. The proportion of each group in all occupations is given by p_k for $k = 1, 2, \ldots, K$. Each occupation j has a vector of characteristics x_j. Assume that for occupations belonging to group k, their characteristics are drawn from distribution with density function $f(x_j; \Theta_k)$, where Θ_k is the parameters corresponding to group k. Under these assumptions, the likelihood of observing x_j is given by

$$\sum_{k=1}^{K} p_k f(x_j; \Theta_k).$$

The probability that occupation j belongs to group k is given by

$$\frac{p_k f(x_j; \Theta_k)}{\sum_{g=1}^{K} p_g f(x_j; \Theta_g)}. \tag{B.1}$$

Suppose that for $k = 1, \ldots, K$, p_k and Θ_k are estimated with a sample of occupations. Then the sample analog of equation (B.1) is given by

$$\frac{\hat{p}_k f(x_j; \hat{\Theta}_k)}{\sum_{g=1}^{K} \hat{p}_g f(x_j; \hat{\Theta}_g)}. \tag{B.2}$$

Table B.1: Parameter Estimates from Cluster Analysis

Cluster (k)	p_k		μ_k		σ_k
1	0.2007	1.2049	−0.3164	0.5609	0.2333
2	0.0303	1.8826	−0.0414	0.8952	0.0007
3	0.0904	0.7909	1.8348	0.2882	0.3516
4	0.0579	−0.8813	1.5840	0.8648	0.0724
5	0.1232	−0.1967	1.5079	−0.5164	0.1955
6	0.1204	−0.9033	0.6461	1.4244	0.4242
7	0.2602	−0.5398	−0.1368	−0.8937	0.4179
8	0.1168	−1.0776	−0.8484	0.3581	0.3929

Occupation j is assigned to group k that maximizes the value given in equation B.2.

I follow Banfield and Raftery (1993)'s framework, and assume that $f(\cdot;\Theta_k)$ is the multivariate normal distribution. The model is estimated with various specifications of the covariance matrix of the distribution, and with various values of K. The specification that attains the highest BIC level is chosen as the best fitting model. The actual estimation is performed with the mclust program for R.[1]

When grouping the 1970 census 3-digit occupation codes this way, their scores for the first three factors extracted by Robinson (2011)'s factor analysis are used as their observed characteristics.[2] The estimation results find that the best fitting model has $K = 8$ clusters, with each mixed distribution $f(\cdot;\Theta_k)$ specified as the multivariate normal distribution with mean vector μ_k and covariance matrix $\sigma_k I$, where I denotes the identity matrix of dimension three. Table B.1 presents the estimates for μ_k and σ_k for the best fitting model.

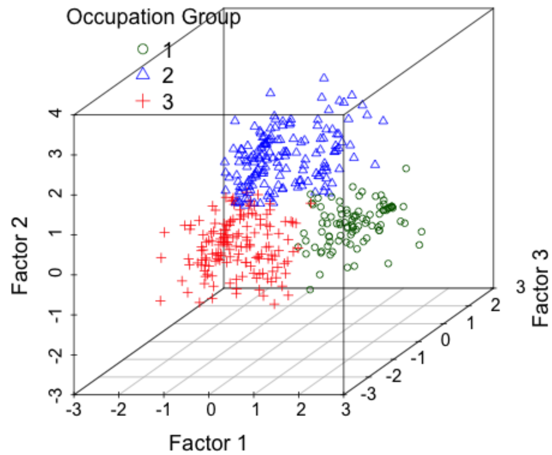
Unfortunately, it is highly difficult to estimate the worker's problem with eight occupation groups. Thus, these clusters are combined further into three clusters.[3] Specifically,

[1] See Fraley and Raftery (2003)

[2] While 4 factors are retained in his analysis, the last factor is omitted since adding of this factor to the present cluster analysis produced a result that was unintuitive than those obtained from the three factors. Leaving out this factor might be justified on the grounds that this factor seems to have marginally identified as a significant factor.

[3] One can estimate the multivariate normal model for a pre-specified number of clusters. The estimation result with three clusters provide similar occupation classification as the one used in this model. It appears

Figure B.1: Factor Scores of the 3-digit Occupations by Aggregated Occupation Group



clusters 1 and 2 are combined to form occupation group 1, clusters 3, 4, 5 and 6 into occupation group 2, clusters 7 and 8 into occupation group 3. Figure B.1 plots the factor scores of the 3-digit occupations by aggregated occupation group, and Table B.2 lists the 3-digit occupation codes in each aggregated occupation group.

Finally, to handle cases where reported 3-digit occupation codes for a particular job spell are classified into different aggregate occupation groups, the following three rules are applied in turn. First, if more than two-thirds of the reported 3-digit occupation codes are classified into a single occupation group, the corresponding job spell is assigned to this occupation group. Second, if reported 3-digit occupation codes are either classified into occupation 1 or occupation 3, but not to occupation 2, then the corresponding job spell is assigned to occupation 1. Analogously, if the reported 3-digit codes are either classified to occupation 2 or occupation 3, but not to occupation 1, then the corresponding job spell is classified to occupation 2. Third, if the reported 3-digit occupations are classified

that as the number of cluster is increased by one, a new cluster is formed by dividing an existing cluster into two.

into occupation 1 not less than occupation 2, then the corresponding job spell is assigned occupation 1. Otherwise, the job spell is assigned into occupation 2.

Bibliography

Banfield, Jeffrey D., and Adrian E. Raftery. 1993. Model-based gaussian and non-gaussian clustering. *Biometrics* 49:803–821.

Fraley, Chris, and Adrian E. Raftery. 2003. Enhanced software of model-based clustering, density estimation, and discriminant analysis: MCLUST. *Journal of Classification* 20:263–86.

Robinson, Chris. 2011. Occupational mobility, occupation distance, and specific human capital. Working paper.

Table B.2: List of the 1970 Census Occupation Codes in Each Aggregated Occupation Group

	Aggregated Occupation Group		
Occupation Category	1	2	3
Professional	1, 3-5, 13, 22 24, 30-32, 34-36 54-56, 74, 86-102 104-122, 25-133 135-144, 174 181, 184, 192-195	2, 6-12, 14-21 23, 25, 26, 33 42-53, 61-72 75-81, 83, 85 103, 123, 124, 134 145-155, 161-170 172, 175, 180, 182 183, 185-191	82, 84, 156, 171 173
Managers	201-212, 215-220 222-245	213, 221	
Sales workers	260, 261, 265 270-282, 284, 285		262, 264, 266, 283
Clerical workers	312, 313, 326, 363		301-310, 314-325 330-362, 364-396
Craftmen	441	403, 404, 410-440 444, 445, 452-495 502, 504-515 520-531, 534-540 543-571, 575	401, 402, 405, 442 443, 446, 450, 501 503, 516, 533, 542 572
Operatives		603, 614, 615, 635 640, 661, 680	601, 602, 604-612 620-634, 641-660 662-674, 681-695
Transport equipment operatives		701, 702, 710, 712 714, 715	704-706, 711, 713
Laborers		750, 761	740, 751-760 762-785
Farmers		801,802, 821	822, 823, 824
Service workers	940, 950, 952, 954 982	903, 935, 944, 945 961, 964, 965	901, 902, 910-934 941-943, 953, 960 962, 963, 980, 981 983, 984

Appendix C

Supplementary Results on Predicted Effects of Learning on Search Effort

Table C.1: Predicted Job Search Effort Allocations in the 2nd Quarter – Type A Workers

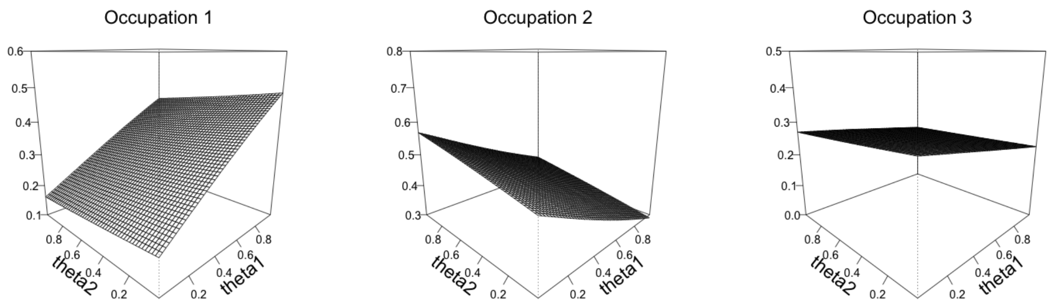
	Search effort allocation		
Current Occupation	Occ 1	Occ 2	Occ3
$(s^0_1, s^0_2) = (5.34, 3.03)$			
Unemployed	0.301	0.417	0.282
Occupation 1	1.000	0.000	0.000
Occupation 2	0.798	0.201	0.001
Occupation 3	0.358	0.575	0.067
$(s^0_1, s^0_2) = 4.02, 3.03)$			
Unemployed	0.347	0.563	0.090
Occupation 1	1.000	0.000	0.000
Occupation 2	0.734	0.266	0.000
Occupation 3	0.227	0.736	0.037
$(s^0_1, s^0_2) = (4.02, 2.69)$			
Unemployed	0.331	0.583	0.086
Occupation 1	1.000	0.000	0.000
Occupation 2	0.675	0.325	0.000
Occupation 3	0.207	0.748	0.044

Table C.2: Predicted Job Search Effort Allocations in the 2nd Quarter – Type B Workers

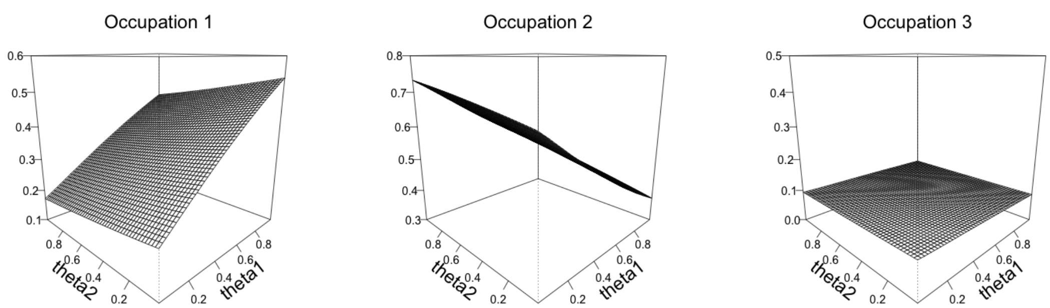
	Search effort allocation		
Current Occupation	Occ 1	Occ 2	Occ3
$(s_1^0, s_2^0) = (5.34, 3.03)$			
Unemployed	0.124	0.198	0.678
Occupation 1	0.137	0.000	0.867
Occupation 2	0.088	0.033	0.879
Occupation 3	0.211	0.000	0.789
$(s_1^0, s_2^0) = (5.34, 2.69)$			
Unemployed	0.094	0.171	0.735
Occupation 1	0.046	0.000	0.954
Occupation 2	0.052	0.030	0.918
Occupation 3	0.080	0.000	0.920
$(s_1^0, s_2^0) = (4.02, 3.03)$			
Unemployed	0.149	0253	0.598
Occupation 1	0.511	0.000	0.489
Occupation 2	0.156	0.077	0.767
Occupation 3	0.364	0.002	0.634
$(s_1^0, s_2^0) = (4.02, 2.69)$			
Unemployed	0.113	0.222	0.665
Occupation 1	0.124	0.003	0.873
Occupation 2	0.089	0.064	0.847
Occupation 3	0.208	0.000	0.792

Figure C.1: Predicted Search Effort Allocations in the 41st Quarter – Type A Workers

Workers with $(s_1^0, s_2^0) = (5.34, 3.03)$



Workers with $(s_1^0, s_2^0) = (4.02, 3.03)$



Workers with $(s_1^0, s_2^0) = (4.02, 2.69)$

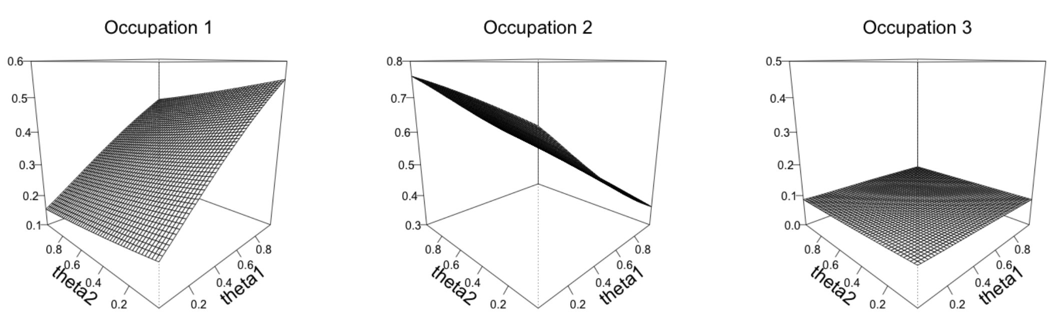
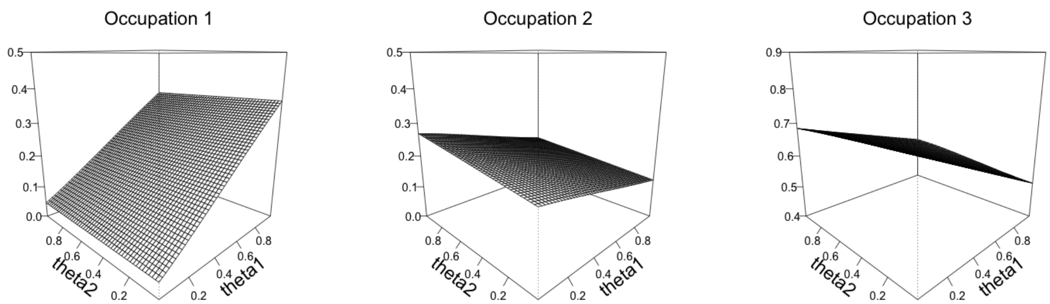


Figure C.2: Predicted Search Effort Allocations in the 41st Quarter – Type B Workers

Workers with $(s_1^0, s_2^0) = (5.34, 3.03)$



Workers with $(s_1^0, s_2^0) = (5.34, 2.69)$

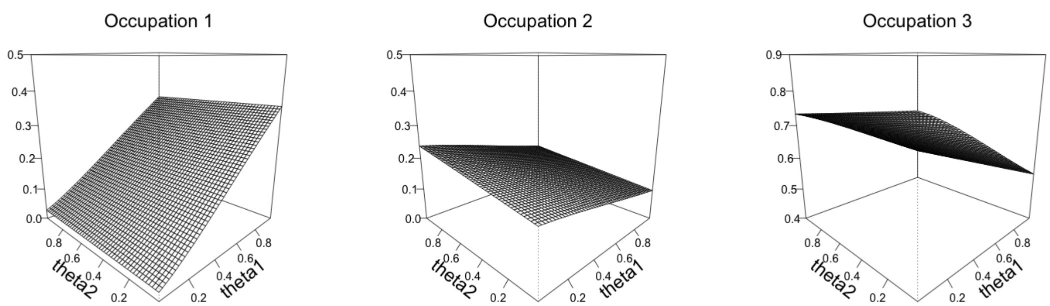
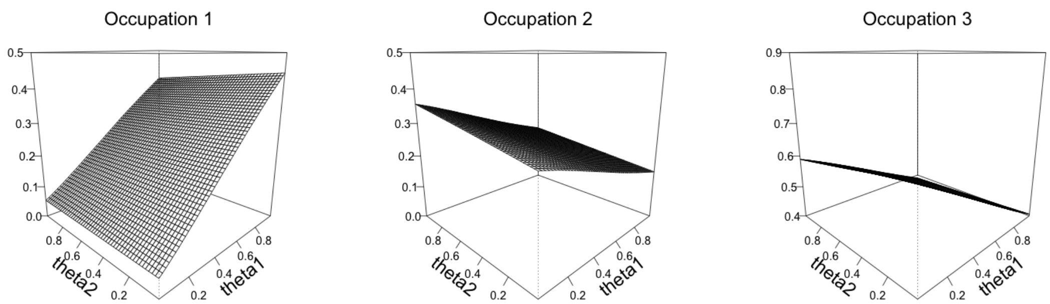
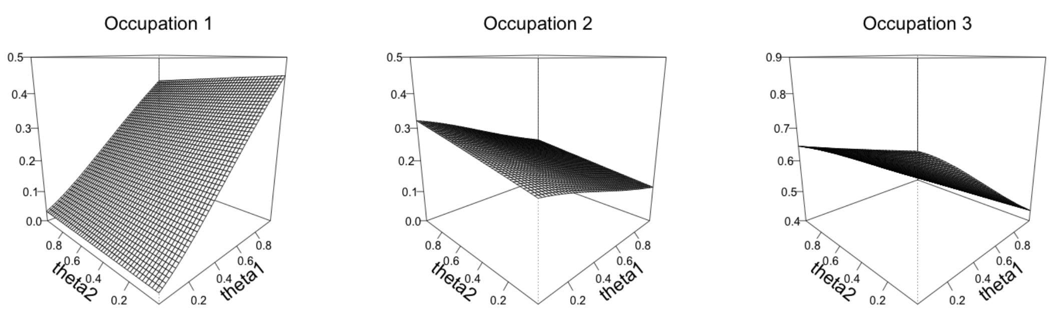


Figure C.3: Predicted Search Effort Allocations in the 41st Quarter – Type B Workers (Continued)

Workers with $(s_1^0, s_2^0) = (4.02, 3.03)$



Workers with $(s_1^0, s_2^0) = (4.02, 2.69)$



Appendix D

Mathematical Proofs for Chapter 4

Proof of Proposition 1

Accounting for the reservation wage property of type 1 immigrant worker's problem, equation (4.4) can be rewritten as

$$rV_m^u = b + \alpha_0\lambda_0 \int_{R_m}^{\overline{w}} (V_m^e(x) - V_m^u)dF(x) + \eta(V_n^u - V_m^u) - \delta_2 V_m^u.$$

Analogously equation (4.5) can be rewritten as

$$\begin{aligned} rV_m^e(w) = w + \alpha_1\lambda_1 \int_{w}^{\overline{w}} (V_m^e(x) - V_m^e(w))dF(x) + \eta I(w \geq R_n)(V_n^e(w) - V_m^e(w)) \\ + \eta I(w < R_n)(V_n^u - V_m^e(w)) + \delta_1(V_m^u - V_m^e(w)) - \delta_2 V_m^e(w). \end{aligned}$$

The above equation yields

$$\frac{d}{dw}V_m^e(w) = \frac{1 + \eta I(w > R_n)\frac{d}{dw}V_n^e(w)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(w)} \tag{D.1}$$

where $V_m^e(w)$ is differentiable.[1] Equation (4.2) yields

$$\frac{d}{dw}V_n^e(w) = \frac{1}{\rho + \lambda_1 \overline{F}(w)}. \tag{D.2}$$

These two equations, together with $V_m^e(R_m) = V_m^u$ and $V_n^e(R_n) = V_n^e$, yield

$$\begin{aligned} R_m =& b + (\alpha_0\lambda_0 - \alpha_1\lambda_1)\int_{R_m}^{\overline{w}} (V_m^e(x) - V_m^e(R_m))dF(x) + \eta I(R_m > R_n)\big(V_n^e(R_n) - V_n^e(R_m)\big) \\ =& b + (\alpha_0\lambda_0 - \alpha_1\lambda_1)\int_{R_m}^{\overline{w}} \overline{F}(x)\frac{d}{dw}V_m^e(x)dx + \eta I(R_m > R_n)\int_{R_m}^{R_n} \frac{d}{dw}V_n^e(x)dx, \end{aligned} \tag{D.3}$$

where the second term is obtained by integration by parts.

Using equation (D.1), the integral in the second term on the right-hand side of equation (D.3) is given by

$$\begin{aligned} \int_{R_m}^{\overline{w}} \overline{F}(x)\frac{d}{dw}V_m^e(x)dx =& \int_{R_m}^{\overline{w}} \frac{\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)}dx \\ &+ I(R_m < R_n)\int_{R_n}^{\overline{w}} \frac{\eta\overline{F}(x)\frac{d}{dw}V_n^e(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)}dx + I(R_m \geq R_n)\int_{R_m}^{\overline{w}} \frac{\eta\overline{F}(x)\frac{d}{dw}V_n^e(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)}dx. \end{aligned} \tag{D.4}$$

The integral in the third term on the right-hand side of equation (D.4) can be split into two terms as follows:

$$\int_{R_m}^{\overline{w}} \frac{\eta\overline{F}(x)\frac{d}{dw}V_n^e(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)} = \int_{R_n}^{\overline{w}} \frac{\eta\overline{F}(x)\frac{d}{dw}V_n^e(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)}dx - \int_{R_n}^{R_m} \frac{\eta\overline{F}(x)\frac{d}{dw}V_n^e(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)}dx.$$

[1] From equation (4.5), it is clear that $V_m^e(w)$ has a kink at R_n.

Thus equation (D.4) can be rearranged to

$$\begin{aligned}\int_{R_m}^{\overline{w}} \overline{F}(x)V_m^e(x)dx &= \int_{R_m}^{\overline{w}} \frac{\overline{F}(x)}{\rho+\eta+\alpha_1\lambda_1\overline{F}(x)}dx + \int_{R_n}^{\overline{w}} \frac{\eta\overline{F}(x)V_n^e(x)}{\rho+\eta+\alpha_1\lambda_1\overline{F}(x)}dx \\ &\quad -I(R_m \geq R_n)\int_{R_n}^{R_m} \frac{\eta\overline{F}(x)V_n^e(x)}{\rho+\eta+\alpha_1\lambda_1\overline{F}(x)}dx\end{aligned}$$

Then substituting the above expression into equation (D.3) and a few algebraic steps yield the desired result. ∎

Proof of Lemma 1

Define function $H(w)$ by

$$\begin{aligned}H(w) = b - w + \int_{w}^{\bar{w}} \frac{(\alpha_0\lambda_0 - \alpha_1\lambda_1)\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(w)}dx + \int_{R_n}^{\overline{w}} \frac{(\alpha_0\lambda_0 - \alpha_1\lambda_1)\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)} \frac{\eta}{\rho + \lambda_1\overline{F}(x)}dx \\ - I(w > R_n)\int_{R_n}^{w} \frac{\rho + \eta + \alpha_0\lambda_0\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)} \frac{\eta}{\rho + \lambda_1\overline{F}(x)}dx. \quad \text{(D.5)}\end{aligned}$$

By Proposition 1, R_m solves equation $H(R_m) = 0$. Function $H(w)$ can be rewritten as

$$\begin{aligned}H(w) = b - \bar{w} + \int_{w}^{\bar{w}} \frac{\rho + \eta + \alpha_0\lambda_0\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)}dx + \int_{R_n}^{\overline{w}} \frac{(\alpha_0\lambda_0 - \alpha_1\lambda_1)\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)} \frac{\eta}{\rho + \lambda_1\overline{F}(x)}dx \\ - I(w > R_n)\int_{R_n}^{w} \frac{\rho + \eta + \alpha_0\lambda_0\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)} \frac{\eta}{\rho + \lambda_1\overline{F}(x)}dx. \quad \text{(D.6)}\end{aligned}$$

Note that $H(w)$ is continuous. It is also decreasing because

$$H'(w) = -\frac{\rho + \eta + \alpha_0\lambda_0\overline{F}(w)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(w)} < 0$$

for $w < R_n$, and

$$H'(w) = -\frac{\rho + \eta + \alpha_0\lambda_0\overline{F}(w)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(w)} - \frac{\rho + \eta + \alpha_0\lambda_0\overline{F}(w)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(w)}\frac{\eta}{\rho + \lambda_1\overline{F}(w)} < 0$$

for $w > R_n$. Therefore $R_n > R_m$ if and only if $H(R_n) < H(R_m) = 0$.

Evaluate $H(w)$ at R_n:

$$\begin{aligned} H(R_n) &= b - \bar{w} + \int_{R_n}^{\bar{w}} \frac{\rho + \eta + \alpha_0\lambda_0\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)}dx + \int_{R_n}^{\overline{w}} \frac{(\alpha_0\lambda_0 - \alpha_1\lambda_1)\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)}\frac{\eta}{\rho + \lambda_1\overline{F}(x)}dx \\ &= b - \bar{w} + \int_{R_n}^{\bar{w}} \frac{\rho + \eta + \alpha_0\lambda_0\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)}\frac{\eta + \rho + \lambda_1\overline{F}(x)}{\rho + \lambda_1\overline{F}(x)}dx - \int_{R_n}^{\bar{w}} \frac{\eta}{\rho + \lambda_1\overline{F}(x)}dx. \end{aligned} \quad \text{(D.7)}$$

Now equation (4.3) can be rewritten as

$$b - \overline{w} = -\int_{R_n}^{\overline{w}} \frac{\rho + \lambda_0\overline{F}(w)}{\rho + \lambda_1\overline{F}(x)}dx. \quad \text{(D.8)}$$

Substituting equation (D.8) into equation (D.7) yields

$$H(R_n) = \int_{R_n}^{\bar{w}} \frac{\rho + \eta + \alpha_0\lambda_0\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)}\frac{\eta + \rho + \lambda_1\overline{F}(x)}{\rho + \lambda_1\overline{F}(x)}dx - \int_{R_n}^{\bar{w}} \frac{\eta + \rho + \lambda_0\overline{F}(x)}{\rho + \lambda_1\overline{F}(x)}dx.$$

Hence $R_n > R_m$ if and only if

$$\int_{R_n}^{\bar{w}} \left[\frac{\eta + \rho + \lambda_0\overline{F}(x)}{\rho + \lambda_1\overline{F}(x)} - \frac{\rho + \eta + \alpha_0\lambda_0\overline{F}(x)}{\rho + \eta + \alpha_1\lambda_1\overline{F}(x)}\frac{\eta + \rho + \lambda_1\overline{F}(x)}{\rho + \lambda_1\overline{F}(x)}\right]dx > 0,$$

which is the desired result. ∎

Proof of Proposition 2

Given $\alpha_1 < 1$ and $\eta > 0$, using equations (4.8) and (4.9), we can show that

$$G_{m1}(w) = \frac{F(w)}{1 + \frac{\alpha_1 \lambda_1}{\delta+\eta}\overline{F}(w)} > \frac{F(w)}{1 + \frac{\lambda_1}{\delta+\eta}\overline{F}(w)} > \frac{F(w)}{1 + \frac{\lambda_1}{\delta}\overline{F}(w)} = G_n(w),$$

which yields $G_{m1}(w) > G_n(w)$ for any $w \in [\underline{w}, \overline{w}]$. This establishes the first-order stochastic dominance of $G_n(w)$ over $G_{m1}(w)$.

Next, using equations (4.8) and (4.10), we can show that

$$G_{m2}(w) = \left(1 - \frac{\delta_2 \alpha_0 (\delta + \lambda_0)\alpha_1 \lambda_1 \overline{F}(w)}{\delta(\alpha_0 \delta_2 + \alpha_0 \lambda_0 + \delta + \eta)(\delta + \eta + \alpha_1 \lambda_1 \overline{F}(w))}\right) G_n(w). \tag{D.9}$$

It is straightforward to show

$$0 \leq \frac{\delta_2 \alpha_0 (\delta + \lambda_0)\alpha_1 \lambda_1 \overline{F}(w)}{\delta(\alpha_0 \delta_2 + \alpha_0 \lambda_0 + \delta + \eta)(\delta + \eta + \alpha_1 \lambda_1 \overline{F}(w))} < 1,$$

which yields $G_{m2}(w) < G_n(w)$ for any $w \in [\underline{w}, \overline{w}]$. This establishes the first-order stochastic dominance of $G_{m2}(w)$ over $G_n(w)$. ∎

Proof of Proposition 3

$G_m(w)$ is given by

$$G_m(w) = \frac{E_{m1} G_{m1}(w) + E_{m2} G_{m2}(w)}{E_{m1} + E_{m2}} \tag{D.10}$$

where E_{m1} and E_{m2} denote the steady-state measures of employed type 1 and type 2 immigrants, respectively. These two variables are determined by the following steady-state flow analysis.

First, define U_{m1} and U_{m2} as the steady-state measures of unemployed type 1 and type 2 immigrants, respectively. Then, E_{m1}, E_{m2}, U_{m1}, and U_{m2} sum up to the measure of all immigrant workers:

$$U_{m1} + U_{m2} + E_{m1} + E_{m2} = \mu. \tag{D.11}$$

Second, type 1 immigrants leave employment at rate δ_1 due to job separation, and at rate δ_2 due to permanent exit from the labour market. They may become type 2 immigrants at rate η. This outflow is balanced by the inflow of unemployed type 1 immigrants becoming employed at rate $\alpha_0\lambda_0$. Therefore,

$$(\delta_1 + \delta_2 + \eta)E_{m1} = \alpha_0\lambda_0 U_{m1}. \tag{D.12}$$

Third, type 2 immigrants leave unemployment at rate λ_0, and leave the labour market permanently at rate δ_2. This outflow is balanced by type 1 unemployed immigrants becoming type 2 at rate η, and type 2 employed immigrants becoming unemployed at rate δ_1. Therefore,

$$(\lambda_0 + \delta_2)U_{m2} = \eta U_{m1} + \delta_1 E_{m2}. \tag{D.13}$$

Fourth, employed type 2 immigrants become unemployed at rate δ_1, or leave the labour market permanently at rate δ_2. This outflow is balanced by the inflow of unemployed type 2 immigrants becoming employed at rate λ_0 and type 1 immigrants becoming type 2 immigrants. Therefore,

$$(\delta_1 + \delta_2)E_{m2} = \lambda_0 U_{m2} + \eta E_{m1}. \tag{D.14}$$

Solving equations (D.11) - (D.14), we obtain E_{m1} and E_{m2}, respectively, as

$$E_{m1} = \frac{\delta_2\alpha_0\lambda_0}{(\eta + \delta_2)(\alpha_0\lambda_0 + \eta + \delta)}\mu, \tag{D.15}$$

and

$$E_{m2} = \frac{\eta\lambda_0(\delta + \eta + \alpha_0\lambda_0 + \alpha_0\delta_2)}{(\eta + \delta_2)(\delta + \lambda_0)(\alpha_0\lambda_0 + \eta + \delta)}\mu. \tag{D.16}$$

Substituting equations (4.9), (4.10), (D.15), and (D.16) into equation (D.10) to obtain the expression for $G_m(w)$, and then comparing the result with $G_n(w)$ in equation (4.8) yields the following relationship between $G_m(w)$ and $G_n(w)$:

$$G_m(w) = \left(1 + \frac{\delta_2(\delta+\lambda_0)(\delta+\eta)\alpha_0\lambda_0(1-\alpha_1)\lambda_1\overline{F}(w)}{\delta((\delta+\eta+\lambda_0)(\lambda_0\delta_2\alpha_0+\eta\lambda_0(\delta+\eta+\alpha_0\lambda_0))(\delta+\eta+\lambda_1\overline{F}(w))}\right)G_n(w).$$

Since $\alpha_1 < 1$, then

$$\frac{\delta_2(\delta+\lambda_0)(\delta+\eta)\alpha_0\lambda_0(1-\alpha_1)\lambda_1\overline{F}(w)}{\delta((\delta+\eta+\lambda_0)(\lambda_0\delta_2\alpha_0+\eta\lambda_0(\delta+\eta+\alpha_0\lambda_0))(\delta+\eta+\lambda_1\overline{F}(w))} > 0,$$

which yields

$$G_m(w) > G_n(w)$$

for any $w \in [\underline{w}, \overline{w}]$. Thus, $G_n(w)$ first-order stochastically dominates $G_m(w)$. ∎

Proof of Proposition 4

In order to establish Proposition 4, we solve a number of differential equations. The outline of this proof is as follows. First, we show the limit of $G_n(w, a)$. Then we do the same for $G_m(w, a)$, and show that it has the same limit.

Step 1

Let $E_n(a)$ and $U_n(a)$ denote the mass of natives of age a who are employed and unemployed, respectively. At any given time, employed workers become unemployed at rate δ_1 and leave the labour market at rate δ_2, and unemployed workers become employed at rate λ_0, and leave the labour market at rate δ_2. Moreover, due to the steady-state assumption, all cohorts have identical aggregate employment and unemployment profiles at all ages. Exploiting the steady-state assumption, $U_n(a)$ and $E_n(a)$ are given by the following system of differential equations

$$\dot{U}_n(a) = -(\lambda_0 + \delta_2)U_n(a) + \delta_1 E_n(a)$$
$$\dot{E}_n(a) = \lambda_0 U_n(a) - (\delta_1 + \delta_2)E_n(a).$$

Furthermore, all workers enter the labour market initially unemployed, implying $E_n(0) = 0$, while the aggregate condition yields $\int (U_n(a) + E_n(a))da = 1 - \mu$. Together with these conditions, the system of the these differential equations yields

$$U_n(a) = \left[\frac{\delta_1\delta_2}{\lambda_0 + \delta_1}e^{-\delta_2 a} + \frac{\lambda_0\delta_2}{\lambda_0 + \delta_1}e^{-(\lambda_0+\delta_1+\delta_2)a}\right](1-\mu), \tag{D.17}$$
$$E_n(a) = \left[\frac{\lambda_0\delta_2}{\lambda_0 + \delta_1}e^{-\delta_2 a} - \frac{\lambda_0\delta_2}{\lambda_0 + \delta_1}e^{-(\lambda_0+\delta_1+\delta_2)a}\right](1-\mu). \tag{D.18}$$

Let $M_n(w, a)$ be the steady-state stock of natives of age a earning wage w or less. Clearly $M_n(w, 0) = 0$. Job-to-job transitions and unemployment-to-job transitions produce the following differential equation regarding $M_n(w, a)$:

$$\frac{dM_n(w,a)}{da} = -(\delta_1 + \delta_2 + \lambda_1\overline{F}(w))M_n(w,a) + \lambda_0 F(w)U_n(a).$$

The differential equation shows that the change in $M_n(w, a)$ with respect to a consists of two parts. The first part is the outflow from $M_n(w, a)$ due to on-the-job search, job destruction and permanent exit from the labour market. The second part is the inflow of unemployed natives finding wage offers of w or less. The general solution to this differential equation is given by

$$M_n(w,a) = e^{-(\delta+\lambda_1\overline{F}(w))a}\left[\int [e^{(\delta+\lambda_1\overline{F}(w))a}\lambda_0 F(w)U_n(a)]da + C\right] \tag{D.19}$$

where C is a constant to be determined by the condition $M_n(w, 0) = 0$. Together with equation (D.17), equation (D.19) gives

$$
\begin{aligned}
M_n(w,a) =&(1-\mu)\lambda_0 F(w)\left[\frac{\delta_1\delta_2}{\lambda_0+\delta_1}\frac{e^{-\delta_2 a}}{\delta_1+\lambda_1\overline{F}(w)}-\frac{\lambda_0\delta_2}{\lambda_0+\delta_1}\frac{e^{-(\lambda_0+\delta)a}}{\lambda_0-\lambda_1\overline{F}(w)}\right] \\
&+e^{-(\delta+\lambda_1\overline{F}(w))a}C.
\end{aligned}
\tag{D.20}
$$

After pinning down C by the condition $M_n(w,0)=0$, equation (D.20) can be rewritten as

$$
\begin{aligned}
M_n(w,a) = (1-\mu)\lambda_0 F(w)\Bigg[&\frac{\delta_1\delta_2}{\lambda_0+\delta_1}\frac{e^{-\delta_2 a}-e^{-(\delta+\lambda_1\overline{F}(w))a}}{\delta_1+\lambda_1\overline{F}(w)} \\
&-\frac{\lambda_0\delta_2}{\lambda_0+\delta_1}\frac{e^{-(\lambda_0+\delta)a}-e^{-(\delta+\lambda_1\overline{F}(w))a}}{\lambda_0-\lambda_1\overline{F}(w)}\Bigg].
\end{aligned}
\tag{D.21}
$$

The age-dependent earnings distribution of natives, given by $G_n(w,a)$, is

$$
G_n(w,a)=\frac{M_n(w,a)}{E_n(a)}.
$$

Using equation (D.18) and (D.21), we obtain

$$
\begin{aligned}
G_n(w,a) =&\frac{\delta_1 F(w)}{\lambda_1\overline{F}(w)+\delta_1}\frac{1-e^{-(\lambda_1\overline{F}(w)+\delta_1)a}}{1-e^{-(\lambda_0+\delta_1)a}} \\
&+\frac{\lambda_0 F(w)}{\lambda_0-\lambda_1\overline{F}(w)}\frac{e^{-(\lambda_1\overline{F}(w)+\delta_1)a}-e^{-(\lambda_0+\delta_1)a}}{1-e^{-(\lambda_0+\delta_1)a}},
\end{aligned}
$$

and therefore

$$
\lim_{a\to\infty} G_n(w,a)=\frac{\delta_1 F(w)}{\delta_1+\lambda_1\overline{F}(w)}.
$$

Step 2

The age profile of the earnings distribution of immigrants can be derived similarly. Let $M_{my}(w,a)$ be the steady-state stock of type y immigrants of age a earning wage w or less. For type 1 immigrants, the change in $M_{m1}(w,a)$ with respect to a comes from the outflow of workers due to job destruction, permanent exit, on-the-job search and type change, and the inflow of unemployed workers accepting wage w or less:

$$\frac{dM_{m1}(w,a)}{da} = -(\delta_1 + \delta_2 + \eta + \alpha_1\lambda_1\overline{F}(w))M_{m1}(w,a) + \alpha_0\lambda_0 F(w)U_{m1}(a). \tag{D.22}$$

The above differential equation can be solved with the condition $M_{m1}(w,a) = 0$. For type 2 immigrants, the change in $M_{m2}(w,a)$ with respect to a comes from the outflow of workers due to job destruction, permanent exit and on-the-job search, the inflow from the pool of employed type 1 immigrants due to type change, and the inflow of unemployed workers accepting wage w or less:

$$\frac{dM_{m2}(w,a)}{da} = -(\delta_1 + \delta_2 + \lambda_1\overline{F}(w))M_{m2}(w,a) + \eta M_{m1}(w,a) + \lambda_0 F(w)U_{m2}(a). \tag{D.23}$$

The above differential equation can be solved with the condition $M_{m2}(w,a) = 0$.

The evolution of unemployment and employment measures of immigrants by age is given by

$$\begin{bmatrix} \dot{U}_{m1}(a) \\ \dot{E}_{m1}(a) \\ \dot{U}_{m2}(a) \\ \dot{E}_{m2}(a) \end{bmatrix} = \begin{bmatrix} -(\alpha_0\lambda_0 + \eta + \delta_2) & \delta_1 & 0 & 0 \\ \alpha_0\lambda_0 & -(\eta+\delta) & 0 & 0 \\ \eta & 0 & -(\lambda_0+\delta_2) & \delta_1 \\ 0 & \eta & \lambda_0 & -\delta \end{bmatrix} \begin{bmatrix} U_{m1}(a) \\ E_{m1}(a) \\ U_{m2}(a) \\ E_{m2}(a) \end{bmatrix}$$

The system of differential equation has the solution

$$\begin{bmatrix} U_{m1}(a) \\ E_{m1}(a) \\ U_{m2}(a) \\ E_{m2}(a) \end{bmatrix} = \begin{bmatrix} 0 & 0 & \delta_1 & 1 \\ 0 & 0 & \alpha_0\lambda_0 & -1 \\ \delta_1 & 1 & \frac{\delta_1(\delta_1+\alpha_0\lambda_0-\eta)}{\eta-\lambda_0-\delta_1} & -\frac{\eta}{\eta+(\alpha_0-1)\lambda_0} \\ \lambda_0 & -1 & \frac{(\delta_1+\alpha_0\lambda_0)\lambda_0-\alpha_0\lambda_0\eta}{\eta-\lambda_0-\delta_1} & \frac{\eta}{\eta+(\alpha_0-1)\lambda_0} \end{bmatrix} \begin{bmatrix} c_1 e^{-\delta_2 a} \\ c_2 e^{-(\lambda_0+\delta)a} \\ c_3 e^{-(\delta_2+\eta)a} \\ c_4 e^{-(\alpha_0\lambda_0+\eta+\delta)a} \end{bmatrix}$$

with constants c_1, c_2, c_3 and c_4 determined by the following conditions:

$$U_{m2}(0) = E_{m1}(0) = E_{m2}(0) = 0,$$

and

$$\int_0^\infty (U_{m1}(a) + U_{m2}(a) + E_{m1}(a) + E_{m2}(a))da = \mu.$$

These conditions yield

$$c_1 = \frac{\delta_2}{\delta_1 + \lambda_0}\mu, \quad c_3 = \frac{\delta_2}{\delta_1 + \alpha_0\lambda_0}\mu, \quad c_4 = \frac{\alpha_0\lambda_0\delta_2}{\delta_1 + \alpha_0\lambda_0}\mu,$$

and

$$c_2 = \frac{\eta\lambda_0(\alpha_0\lambda_0 + \delta_1)(\eta - \lambda_0 - \alpha_0\delta_1)}{(\lambda_0 + \delta_1)(\eta - \lambda_0 - \delta_1)(\eta + (\alpha_0 - 1)\lambda_0)}\mu.$$

The general solution to differential equation (D.22) is given by

$$M_{m1}(w, a) = e^{-(\alpha_1\lambda_1\overline{F}(w))a}\left[\alpha_0\lambda_0 F(w)\int \left[U_{m1}(a)e^{(\alpha_1\lambda_1\overline{F}(w)+\delta+\eta)a}\right]da + C\right]. \tag{D.24}$$

The integral on the right-hand side of equation (D.24) can be rearranged to

$$\int U_{m1}(a)e^{(\alpha_1\lambda_1\overline{F}(w)+\delta+\eta)a}da = \int \left[c_3\delta_1 e^{(\alpha_1\lambda_1\overline{F}(w)+\delta_1)a} + c_4 e^{(\alpha_1\lambda_1\overline{F}(w)-\alpha_0\lambda_0)a}\right]da$$
$$= \frac{c_3\delta_1\, e^{(\alpha_1\lambda_1\overline{F}(w)+\delta_1)a}}{\alpha_1\lambda_1\overline{F}(w)+\delta_1} + \frac{c_4\, e^{(\alpha_1\lambda_1\overline{F}(w)-\alpha_0\lambda_0)a}}{\alpha_1\lambda_1\overline{F}(w)-\alpha_0\lambda_0}.$$

Therefore,

$$M_{m1}(w,a) = \alpha_0\lambda_0 F(w)\left[\frac{c_3\delta_1\, e^{-(\eta+\delta_2)a}}{\alpha_1\lambda_1\overline{F}(w)+\delta_1} + \frac{c_4\, e^{-(\eta+\delta_2)a}}{\alpha_1\lambda_1\overline{F}(w)-\alpha_0\lambda_0}\right] + Ce^{-(\alpha_1\lambda_1\overline{F}(w)+\eta+\delta)a}. \quad \text{(D.25)}$$

Using $M_{m1}(w,0) = 0$ to solve for C, and substituting the result into equation (D.25) yields

$$M_{m1}(w,a) = \alpha_0\lambda_0 F(w)\Bigg[c_3\delta_1\, \frac{e^{-(\eta+\delta_2)a} - e^{-(\alpha_1\lambda_1\overline{F}(w)+\eta+\delta)a}}{\alpha_1\lambda_1\overline{F}(w)+\delta_1}$$
$$+c_4\, \frac{e^{-(\eta+\delta_2)a} - e^{-(\alpha_1\lambda_1\overline{F}(w)+\eta+\delta)a}}{\alpha_1\lambda_1\overline{F}(w)-\alpha_0\lambda_0}\Bigg].$$

The general solution to equation (D.23) then yields

$$\begin{aligned} M_{m2}(w,a) = \quad & e^{-(\lambda_1\overline{F}(w)+\delta)a}\left[\int (\eta M_{m1}(w,a) + \lambda_0 F(w)U_{m2}(a))e^{(\lambda_1\overline{F}(w)+\delta)a}da + C_2\right] \\ = \quad & \eta e^{-(\lambda_1\overline{F}(w)+\delta)a}\int M_{m1}(w,a)e^{(\lambda_1\overline{F}(w)+\delta)a}da \\ & + \lambda_0 F(w)e^{-(\lambda_1\overline{F}(w)+\delta)a}\int U_{m2}(a))e^{(\lambda_1\overline{F}(w)+\delta)a}da + C_2 e^{-(\lambda_1\overline{F}(w)+\delta)a}. \end{aligned} \quad \text{(D.26)}$$

The integrals on the right-hand side of equation (D.26) can be written as

$$\int M_{m1}(w,a)e^{(\lambda_1\overline{F}(w)+\delta)a}da = \frac{\alpha_0\lambda_0 F(w)c_3\delta_1}{\alpha_1\lambda_1\overline{F}(w)+\delta_1}\left[\frac{e^{(\lambda_1\overline{F}(w)+\delta_1-\eta)a}}{\lambda_1\overline{F}(w)+\delta_1-\eta} - \frac{e^{((1-\alpha_1)\lambda_1\overline{F}(w)-\eta)a}}{(1-\alpha_1)\lambda_1\overline{F}(w)-\eta}\right]$$
$$+ \frac{\alpha_0\lambda_0 F(w)c_4}{\alpha_1\lambda_1\overline{F}(w)-\alpha_0\lambda_0}\left[\frac{e^{(\lambda_1\overline{F}(w)-\alpha_0\lambda_0-\eta)a}}{\lambda_1\overline{F}(w)-\alpha_0\lambda_0-\eta} - \frac{e^{((1-\alpha_1)\lambda_1\overline{F}(w)-\eta)a}}{(1-\alpha_1)\lambda_1\overline{F}(w)-\eta}\right],$$

and

$$\int U_{m2}(a)e^{\lambda_1\overline{F}(w)+\delta)a}da = c_1\delta_1\frac{e^{(\lambda_1\overline{F}(w)+\delta_1)a}}{\lambda_1\overline{F}(w)+\delta_1} + c_2\frac{e^{(\lambda_1\overline{F}(w)-\lambda_0)a}}{\lambda_1\overline{F}(w)-\lambda_0}$$
$$+ c_3\frac{\delta_1(\delta_1+\alpha_0\lambda_0-\eta)}{\eta-\lambda_0-\delta_1}\frac{e^{(\lambda_1\overline{F}(w)+\delta_1-\eta)a}}{\lambda_1\overline{F}(w)+\delta_1-\eta} - \frac{c_4\eta}{\eta+(\alpha_0-1)\lambda_0}\frac{e^{(\lambda_1\overline{F}(w)-\alpha_0\lambda_0-\eta)a}}{\lambda_1\overline{F}(w)-\alpha_0\lambda_0-\eta}.$$

Using $M_{m2}(w,0) = 0$ to solve for C_2 in equation (D.26), and substituting the result into the same equation yields

$$\begin{aligned}
&M_{m2}(w,a)\\
&= \frac{\eta\alpha_0\lambda_0F(w)c_3\delta_1}{\alpha_1\lambda_1\overline{F}(w)+\delta_1}\left[\frac{e^{-(\eta+\delta)a}-e^{-(\lambda_1\overline{F}(w)+\delta)a}}{\lambda_1\overline{F}(w)+\delta_1-\eta} - \frac{e^{-(\alpha_1\lambda_1\overline{F}(w)+\eta+\delta)a}-e^{-(\lambda_1\overline{F}(w)+\delta)a}}{(1-\alpha_1)\lambda_1\overline{F}(w)-\eta}\right]\\
&+\frac{\eta\alpha_0\lambda_0F(w)c_4}{\alpha_1\lambda_1\overline{F}(w)-\alpha_0\lambda_0}\left[\frac{e^{-(\alpha_0\lambda_0+\eta+\delta)a}-e^{-(\lambda_1\overline{F}(w)+\delta)a}}{\lambda_1\overline{F}(w)-\alpha_0\lambda_0-\eta} - \frac{e^{-(\alpha_1\lambda_1\overline{F}(w)+\eta+\delta)a}-e^{-(\lambda_1\overline{F}(w)+\delta)a}}{(1-\alpha_1)\lambda_1\overline{F}(w)-\eta}\right]\\
&+\lambda_0F(w)c_1\delta_1\frac{e^{-\delta_2 a}-e^{-(\lambda_1\overline{F}(w)+\delta)a}}{\lambda_1\overline{F}(w)+\delta_1} + \lambda_0F(w)c_2\frac{e^{-(\lambda_0+\delta)a}-e^{-(\lambda_1\overline{F}(w)+\delta)a}}{\lambda_1\overline{F}(w)-\lambda_0}\\
&+\lambda_0F(w)c_3\frac{\delta_1(\delta_1+\alpha_0\lambda_0-\eta)}{\eta-\lambda_0-\delta_1}\frac{e^{-(\delta_2+\eta)a}-e^{-(\lambda_1\overline{F}(w)+\delta)a}}{\lambda_1\overline{F}(w)+\delta_1-\eta}\\
&-\lambda_0F(w)\frac{c_4\eta}{\eta+(\alpha_0-1)\lambda_0}\frac{e^{-(\alpha_0\lambda_0+\eta+\delta)a}-e^{-(\lambda_1\overline{F}(w)+\delta)a}}{\lambda_1\overline{F}(w)-\alpha_0\lambda_0-\eta}.
\end{aligned}$$

The age-dependent earnings distribution of immigrants, given by $G_m(w,a)$, is

$$G_m(w,a) = \frac{M_{m1}(w,a)+M_{m2}(w,a)}{E_{m1}(a)+E_{m2}(a)}. \tag{D.27}$$

To show that $G_m(w,a)$ converges to the same distribution as $G_n(w,a)$, rewrite equation (D.27) as

$$G_m(w,a) = \frac{e^{\delta_2 a}[M_{m1}(w,a)+M_{m2}(w,a)]}{e^{\delta_2 a}[E_{m1}(a)+E_{m2}(a)]} = \frac{e^{\delta_2 a}M_{m1}(w,a)+e^{\delta_2 a}M_{m2}(w,a)}{e^{\delta_2 a}E_{m1}(a)+e^{\delta_2 a}E_{m2}(a)}.$$

It is then straightforward to show the limits of the terms appearing in both the numerator and denominator.

$$\lim_{a\to\infty} e^{\delta_2 a} M_{m1}(w,a) = 0, \quad \lim_{a\to\infty} e^{\delta_2 a} M_{m2}(w,a) = \frac{\lambda_0 c_1 \delta_1 F(w)}{\lambda_1 \overline{F}(w) + \delta_1},$$
$$\lim_{a\to\infty} e^{\delta_2 a} E_{m1}(a) = 0, \quad \lim_{a\to\infty} e^{\delta_2 a} E_{m2}(a) = \lambda_0 c_1.$$

Therefore we obtain the desired result.

$$\lim_{a\to\infty} G_m(w,a) = \frac{\delta_1 F(w)}{\lambda_1 \overline{F}(w) + \delta_1}.$$

∎

Appendix E

Likelihood Contributions for Chapter 4

For all individuals in the data set, the likelihood contributions account for their initially observed employment outcomes, the duration of the initial spell, and if applicable, the wages earned and transition made at the end of the spell. In addition to these pieces of information, the characteristics of the next spell also enter into the likelihood if the initial spell is an unemployment spell, or if it is a job spell that ends with a transition to an unemployment spell. More specifically, a list of variables used in the likelihood can be written as $x_i = (w_1, d_1, c_1, t_1, d_2, c_2)$ for those initially employed, where w_1 and d_1 represent the wage earned and duration of the initial job, c_1 takes on a value of 1 if the initial spell is censored and 0 otherwise, and t_1 takes on a value of 1 if the initial spell ends with a transition to a new job, and 0 if it ends with a transition to unemployment. If the initial spell ends with a transition to unemployment, d_2 represents the duration of the second spell, with c_2 being the indicator for censoring of the second spell. For those initially unemployed, x_i is given as $x_i = (d_1, c_1, w_2, d_2, c_2, t_2)$ with d_1 and c_1 representing the duration and censoring outcome of the unemployment spell, respectively, w_2, d_2 and c_2 representing the wage, duration and censoring outcome of the following job spell, respectively, and t_2 representing the type of transition made if the second spell is complete.

The likelihood contributions $L_n(\theta; x_i)$ and $L_{m2}(\theta; x_i)$ have the familiar structure in the job search literature because the search behaviours of the corresponding groups are standard. In contrast, derivation of $L_{m1}(\theta; x_i)$ involves accounting for possible changes in the search process among workers, and requires more careful presentation. We discuss derivations for these three in turn.

Derivation of $L_n(\theta; x_i)$ and $L_{m2}(\theta; x_i)$

In the steady state, a native worker is employed with probability $\lambda_0/(\lambda_0 + \delta)$, and the distribution of wages earned on that job is given by $G_n(w_1)$. Given the initial wage w, the residual duration of the first job spell follows the exponential distribution with parameter $(\lambda_1\overline{F}(w) + \delta)$. At the end of a job spell, a native worker makes a job-to-job transition with probability $\lambda_1\overline{F}(w)/(\lambda_1\overline{F}(w) + \delta)$, or a job-to-unemployment transition with probability $\delta_1/(\lambda_1\overline{F}(w) + \delta)$. The duration of a new unemployment spell follows the exponential distribution with parameter $(\lambda_0 + \delta_2)$ and ends with a transition to a new job with probability $\lambda_0/(\lambda_0 + \delta_2)$. Gathering all the components together, the likelihood contribution for native-born individuals initially employed is given by

$$\frac{\lambda_0}{\lambda_0 + \delta} g_n(w) e^{-(\lambda_1\overline{F}(w)+\delta)d_1} \left[(\lambda_1\overline{F}(w))^{t_1} (\delta_1 e^{-(\lambda_0+\delta_2)d_2} \lambda_0^{1-c_2})^{1-t_1}\right]^{1-c_1}$$

where $g_n(w)$ denotes the density of $G_n(w)$.

The probability that a native individual is unemployed in the steady-state is $\delta/(\lambda_0 + \delta)$. The residual unemployment duration follows the exponential distribution with parameter $(\lambda_0 + \delta_2)$, and the probability that an unemployment spell ends with a transition to a new job as opposed to a permanent exit from the labour market is given by $\lambda_0/(\lambda_0 + \delta_2)$. The distribution of wages on new jobs is given by $F(w)$. The duration of a new job follows the exponential distribution with parameter $(\lambda_1\overline{F}(w) + \delta)$, and ends with a transition to a job spell with probability $\lambda_1\overline{F}(w)/(\lambda_1\overline{F}(w)+\delta)$ or with a job-to-unemployment transition with probability $\delta_1/(\lambda_1\overline{F}(w)+\delta)$. Therefore, the likelihood contribution for native workers initially unemployed takes the form

$$\frac{\delta}{\lambda_0+\delta}e^{-(\lambda_0+\delta_2)d_1}\left[\lambda_0 f(w)e^{-(\lambda_1\overline{F}(w)+\delta)d_2}\left(\delta_1^{1-t_2}(\lambda_1\overline{F}(w))^{t_2}\right)^{1-c_2}\right]^{1-c_1}.$$

The likelihood contribution for type 2 immigrants is similar to the natives' since they share the same search process. $L_n(\theta, x_i)$ and $L_{m2}(\theta_i, x_i)$ differ due to differences in the probabilities of the initial employment status and earned wage. The probability that a type 2 immigrant worker is employed at any instance is given by

$$\frac{\lambda_0(\delta+\eta+\alpha_0\lambda_0+\alpha_0\delta_2)}{(\lambda_0+\delta)(\delta+\eta+\alpha_0\lambda_0)},$$

and the distribution of the wage earned is given by $G_{m2}(w)$ with its density denoted by $g_{m2}(w)$. The steady-state probability that a type 2 immigrant worker is unemployed is given by

$$\frac{\delta((\delta+\eta)+\delta_1\alpha_0\lambda_0)}{(\lambda_0+\delta)(\delta+\eta+\alpha_0\lambda_0)}.$$

Thus, $L_{m2}(\theta, x_i)$ takes the form

$$\frac{\lambda_0(\delta+\eta+\alpha_0\lambda_0+\alpha_0\delta_2)}{(\lambda_0+\delta)(\delta+\eta+\alpha_0\lambda_0)}g_{m2}(w)e^{-(\lambda_1\overline{F}(w)+\delta)d_1}\left[(\lambda_1\overline{F}(w))^{t_1}(\delta_1 e^{-(\lambda_0+\delta_2)d_2}\lambda_0^{1-c_2})^{1-t_1}\right]^{1-c_1}$$

for those initially employed, or

$$\frac{\delta((\delta+\eta)+\delta_1\alpha_0\lambda_0)}{(\lambda_0+\delta)(\delta+\eta+\alpha_0\lambda_0)}e^{-(\lambda_0+\delta_2)d_1}\left[\lambda_0 f(w)e^{-(\lambda_1\overline{F}(w)+\delta)d_2}\left(\delta_1^{1-t_2}(\lambda_1\overline{F}(w))^{t_2}\right)^{1-c_2}\right]^{1-c_1}$$

for those initially unemployed.

Derivation of $L_{m1}(\theta; x_i)$

Derivation of $L_{m1}(\theta; x_i)$ is more involved because of the possibility that type 1 immigrants experience changes in search parameters. It is therefore helpful to introduce variables reflecting immigrant types upon transitions between spells. The variables, denoted y_1 and y_2, are used to first form the joint probabilities with the observed outcomes, and then integrated out to yield the expression for $L_{m1}(\theta; x_i)$. This process results in the following expression of the likelihood contribution of type 1 immigrants who are initially unemployed:

$$P_U \sum_{y_1=1}^{2} \left[P_{u1}(d_1, y_1, c_1) \left[f(w) \sum_{y_2=y_1}^{2} P_{j2}(d_2, y_2, c_2|w, y_1) P_{tr}(t_2|w, y_2)^{1-c_2} \right]^{1-c_1} \right]$$

where P_U represents the probability that a type 1 immigrant is unemployed at any instant in the steady state, i.e., $P_U = (\delta + \eta)/(\delta + \eta + \alpha_0\lambda_0)$. Component probability $P_{u1}(d_1, y_1, c_1)$ is the joint probability of the residual unemployment duration, censoring indicator and the ending immigrant type of the first spell. $f(w)$ is the distribution of accepted wage offers. $P_{j2}(d_2, y_2, c_2|w, y_1)$ is the joint probability of the duration, censoring outcome and ending immigrant type of the following job spell conditional on the accepted wage and the starting immigrant type on the job. The last factor, $P_{tr}(t_2|y_2, w)$, accounts for the type of transition made at the end of the second spell.

Similarly, for type 1 immigrants seen initially employed, the likelihood contribution takes the form

$$P_E g_{m1}(w) \sum_{y_1=1}^{2} \left[P_{j1}(d_1, y_1, c_1|w) \left[P_{tr}(t_1|w, y_1) \left(\sum_{y_2=y_1}^{2} P_{u2}(d_2, y_2, c_2|y_1) \right)^{1-t_1} \right]^{1-c_1} \right]$$

where P_E denotes the probability that a type 1 immigrant is employed at any instant in the steady state, i.e., $P_E = \alpha_0\lambda_0/(\delta + \eta + \alpha_0\lambda_0)$, and $g_{m1}(w)$ is the density of the steady state earned wage distribution for type 1 immigrants. Component probability $P_{j1}(d_1, y_1, c_1|w)$ is the joint probability of the residual duration, censoring outcome and the ending immigrant type of the first job spell, and $P_{tr}(t_1|w, y_1)$ is the probability of the observed transition from the job spell. For those who transition to unemployment, $P_{u2}(d_2, y_2, c_2|y_2)$ accounts for the joint probability of the duration, ending immigrant type and censoring indicator of the following unemployment spell .

Having presented the overall structure of the likelihood contribution, we now provide the expressions for its components. If an immigrant remains as type 1 during his first observed spell, the residual duration of the spell follows the exponential distribution with parameter $(\alpha_0\lambda_0 + \delta_2)$ if it is an unemployment spell, or with parameter $(\alpha_1\lambda_1\overline{F}(w) + \delta)$ if it is a job spell. Therefore for $y_1 = 1$, $P_{u1}(d_1, c_1, y_1)$ and $P_{j1}(d_1, c_1, y_1|w)$ are given by, respectively,

$$P_{u1}(d_1, c_1, 1) = (\alpha_0\lambda_0 + \delta_2)^{1-c_1} e^{-(\alpha_0\lambda_0+\delta_2+\eta)d_1} \tag{E.1}$$

and

$$P_{j1}(d_1, c_1, 1|w) = (\alpha_1\lambda_1\overline{F}(w) + \delta)^{1-c_1} e^{-(\alpha_1\lambda_1\overline{F}(w)+\delta+\eta)d_1}. \tag{E.2}$$

If immigrants change types during the first spell, i.e., $y_1 = 2$, the duration of the spell can be given as the sum of two independent exponential random variables. If the spell is an unemployment spell, the relevant two variables are exponential with parameters $(\alpha_0\lambda_0 + \eta + \delta_2)$ and $(\lambda_0 + \delta_2)$. Letting s and $(d - s)$ denote the realizations of these two variables, the distribution of their summed value, d, is given by

$$\int_0^d \Big(e^{-(\alpha_0\lambda_0+\eta+\delta_2)s}\eta\Big)\Big(e^{-(\lambda_0+\delta_2)(d-s)}(\lambda_0 + \delta_2)\Big)ds$$
$$= \frac{\eta(\lambda_0 + \delta_2)}{(\alpha_0 - 1)\lambda_0 + \eta}(e^{-(\lambda_0+\delta_2)d} - e^{-(\alpha_0\lambda_0+\eta+\delta_2)d}).$$

The probability that a completed unemployment spell ends with a transition to a job is

$\lambda_0/(\lambda_0+\delta_2)$, therefore for $y_1 = 2$ and $c_1 = 0$, $P_{u1}(d_1, y_1, c_1)$, is given by

$$P_{u1}(d_1, 2, 0) = \frac{\eta\lambda_0}{(\alpha_0-1)\lambda_0+\eta}(e^{-(\lambda_0+\delta_2)d_1} - e^{-(\alpha_0\lambda_0+\eta+\delta_2)d_1}). \tag{E.3}$$

If the spell is censored, i.e., $c_1 = 1$, the relevant expression for $P_{u1}(d_1, y_1, c_1)$ is given by

$$\begin{aligned} P_{u1}(d_1, 2, 1) &= \int_{d_1}^{\infty} \frac{\eta(\lambda_0+\delta_2)}{(\alpha_0-1)\lambda_0+\eta}(e^{-(\lambda_0+\delta_2)\tau} - e^{-(\alpha_0\lambda_0+\eta+\delta_2)\tau})d\tau \\ &= \frac{\eta(\lambda_0+\delta_2)}{(\alpha_0-1)\lambda_0+\eta}\Big[\frac{e^{-(\lambda_0+\delta_2)d_1}}{\lambda_0+\delta_2} - \frac{e^{-(\alpha_0\lambda_0+\eta+\delta_2)d_1}}{\alpha_0\lambda_0+\eta+\delta_2}\Big]. \end{aligned} \tag{E.4}$$

Analogously, if a type 1 immigrant become a type 2 immigrant during a job spell, the spell duration is the sum of two independent exponential random variables with parameters, respectively, $(\alpha_1\lambda_1\overline{F}(w)+\eta+\delta)$ and $(\lambda_1\overline{F}(w)+\delta)$. If it is a completed spell, i.e., $c_1 = 0$, the expression for $P_{j1}(d_1, y_1, c_1|w)$ is given by

$$\begin{aligned} P_{j1}(d_1, 2, 0|w) &= \int_0^{d_1} e^{-(\alpha_1\lambda_1\overline{F}(w)+\delta+\eta)s}\eta e^{-(\lambda_1\overline{F}(w)+\delta)(d_1-s)}(\lambda_1\overline{F}(w)+\delta)ds \\ &= \frac{\eta(\lambda_1\overline{F}(w)+\delta)}{(\alpha_1-1)\lambda_1\overline{F}(w)+\eta}(e^{-(\lambda_1\overline{F}(w)+\delta)d_1} - e^{-(\alpha_1\lambda_1\overline{F}(w)+\delta+\eta)d_1}). \end{aligned} \tag{E.5}$$

If it is censored, it is given by

$$\begin{aligned} P_{j1}(d_1, 2, 1|w) &= \int_{d_1}^{\infty} P_{j1}(\tau, 2, 0|w)d\tau \\ &= \frac{\eta(\lambda_1\overline{F}(w)+\delta)}{(\alpha_1-1)\lambda_1\overline{F}(w)+\eta}\Big[\frac{e^{-(\lambda_1\overline{F}(w)+\delta)d_1}}{\lambda_1\overline{F}(w)+\delta} - \frac{e^{-(\alpha_1\lambda_1\overline{F}(w)+\delta+\eta)d_1}}{\alpha_1\lambda_1\overline{F}(w)+\delta+\eta}\Big]. \end{aligned} \tag{E.6}$$

If an immigrant starts the second spell as a type 1 immigrant, i.e., $y_1 = 1$, the component probabilities given in equations (E.1) – (E.6) apply, so that $P_{u2}(d_2, c_2, y_2|y_1 = 1) = P_{u1}(d_2, c_2, y_2)$ and $P_{j2}(d_2, c_2, y_2|w, y_1 = 1) = P_{j1}(d_2, c_2, y_2|w)$. If an immigrant is of type 2 at the start of a spell, the duration of the spell follows the exponential distribution with parameter $(\lambda_0 + \delta_2)$ if it is an unemployment spell or with parameter $(\lambda_1 \overline{F}(w) + \delta)$ if it is a job spell. Therefore for $y_1 = y_2 = 2$, $P_{u2}(d_2, y_2, c_2|y_1)$ and $P_{j2}(d_2, y_2, c_2|y_1)$ are, respectively,

$$P_{u2}(d_2, 2, c_2|2) = (\lambda_0 + \delta_2)^{1-c_2} e^{-(\lambda_0+\delta_2)d_2},$$

and

$$P_{j2}(d_2, 2, c_2|w, 2) = (\lambda_1 \overline{F}(w) + \delta)^{1-c_2} e^{-(\lambda_1 \overline{F}(w)+\delta)d_2}.$$

For $j \in \{1, 2\}$, $\mathrm{P}_{tr}(t_j|w, y_j)$ is the component probability of the transition outcome from a job spell. Depending on the ending immigrant type, it is given by

$$\mathrm{P}_{tr}(t_j|w, y_j = 1) = \frac{(\alpha_1 \lambda_1 \overline{F}(w))^{t_j} \delta_1^{1-t_j}}{\alpha_1 \lambda_1 \overline{F}(w) + \delta},$$

or

$$\mathrm{P}_{tr}(t_j|w, y_j = 2) = \frac{(\lambda_1 \overline{F}(w))^{t_j} \delta_1^{1-t_j}}{\lambda_1 \overline{F}(w) + \delta}.$$

Curriculum Vitae

Name:	Masashi Miyairi
Post-Secondary Education and Degrees:	Keio University, Tokyo, Japan 1998–2002 B.A.
	University of Michigan, Ann Arbor, MI 2002–2004 M.A.
	New York University, New York, NY 2004–2006 M.A.
	The University of Western Ontario, London, ON 2006–2014 Ph.D.
Honours and Awards:	Western Graduate Research Scholarship The University of Western Ontario 2006–2010
Related Work Experience:	Teaching Assistant The University of Western Ontario 2006–2010
	Research Assistant The University of Western Ontario 2008–2014
	Instructor The University of Western Ontario 2010–2013