ESSAYS ON ECONOMIC CAUSES AND CONSEQUENCES OF MIGRATION

Yury Andrienko

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

Discipline of Economics

Faculty of Economics and Business

The University of Sydney

August 2010
Statement of originality

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.
For my parents who believed in me

but will never know I wrote this

and for Natalia and Nika with love
Acknowledgements

I would like to thank many people without whom this research project, the longest in my life so far, would not have been completed. First and foremost, I am grateful to my supervisor Russell Ross whose guidance, encouragement, advice, and support inspired my research on a regular basis. It would not be fair not to mention other people who helped me in research. I wish to thank Hajime Katayama, my associate supervisor, who guided me in econometric modelling and highly supported the technique I have been using. My work has benefitted from discussions with three other colleagues, Stephen Whelan, Andrew Wait, and Vladimir Smirnov, who made useful comments on various chapters. My special thanks to Michael Paton who served as a mentor at one stage of research and who read a draft of the thesis and made suggestions how to improve the writing and structure. I am obliged to Tigger Wise who did editorial work on almost the entire thesis and noted some inaccuracies.

I also would like to thank Guyonne Kalb, Andrew Leigh, and especially Kostas Mavromaras for excellent discussions of a paper which motivated further research and also other participants of national and international conferences ALMRW 2008 in Wellington, ECOMOD 2009 in Ottawa, LEW 2009 in Brisbane, and 2009 PhD Conference in Economics and Business in Perth. My special appreciation goes the three examiners who made a lot of comments and suggestions.

Last but not least, I am thankful to my wife whom I met in the middle of the deepest and most sacred lake, thanks to God for this, and who has done everything possible to organise my domestic life and provide a working atmosphere for me. A very special gratitude is to our little daughter, our gift from God, whose appearance has radically changed our lives. I am so sorry I could not give more attention to you these years.
This thesis uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) and is managed by the Melbourne Institute of Applied Economic and Social Research (MIAESR). The findings and views reported in this thesis, however, are those of the author and should not be attributed to either FaHCSIA or the MIAESR.
Abstract

Migration is a multidimensional phenomenon requiring an interdisciplinary approach. This thesis studies some economic aspects of the internal migration of labour. A model of migration as investment in human capital is applied throughout the thesis to study economic causes and consequences of internal migration on a micro level. Various predictions from the theory are verified on longitudinal micro data from the Household Income and Labour Dynamics in Australia (HILDA) survey. The thesis is composed of three essays:

1) Causes of migration, the individual level push and pull factors facilitating or hampering mobility and representing both costs and benefits to migration, are studied in Chapter 2. A binary dependent variable model for the likelihood of an individual migration decision is estimated on panel data from the HILDA survey by means of the probit model with individual random effects. The main results are that those not in the labour force, similarly to the unemployed, are more mobile than the employed; and that higher individual wages and greater remoteness from larger urban centres also increase the likelihood of migration.

2) Chapter 3 studies wage returns to internal migration. Evidence is sought for the theoretical predictions of the traditional human capital model of investment in migration about a positive wage premium: positive returns to migration distance and human capital. Using individual-level data from the HILDA Survey and applying a system GMM to a dynamic panel earnings model, it is found that in the short-run there are returns to distance which increase with the level of education and decline with the level of pre-migration wage. The conclusion is that internal migration in Australia is a good strategy only for better educated and lower income individuals.

3) Several theoretical models of migration destination search are presented in Chapter 4. It discusses two models of migration as an outcome of the fixed-sample-size search and the sequential search. A model with endogenous investment in search activity demonstrates that lower initial utility increases chances to participate in search and that the likelihood of migration depends on budget constraints: those of the poor who can afford to buy relatively more information are expected to gain more than others.
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABS</td>
<td>Australian Bureau of Statistics</td>
</tr>
<tr>
<td>ACT</td>
<td>Australian Capital Territory</td>
</tr>
<tr>
<td>ARIA</td>
<td>Accessibility/Remoteness Index of Australia</td>
</tr>
<tr>
<td>ASGC</td>
<td>Australian Standard Geographical Classification</td>
</tr>
<tr>
<td>CD</td>
<td>Census District</td>
</tr>
<tr>
<td>c.d.f.</td>
<td>cumulative distribution function</td>
</tr>
<tr>
<td>FOC</td>
<td>first order condition</td>
</tr>
<tr>
<td>FSS</td>
<td>fixed sample size</td>
</tr>
<tr>
<td>FT</td>
<td>full-time</td>
</tr>
<tr>
<td>GMM</td>
<td>generalized method of moments</td>
</tr>
<tr>
<td>HILDA</td>
<td>Household Income and Labour Dynamics in Australia</td>
</tr>
<tr>
<td>IIA</td>
<td>irrelevance of independent alternatives</td>
</tr>
<tr>
<td>i.i.d.</td>
<td>identically independently distributed</td>
</tr>
<tr>
<td>LHS</td>
<td>left hand side</td>
</tr>
<tr>
<td>NEG</td>
<td>new economic geography</td>
</tr>
<tr>
<td>NEM</td>
<td>new economics of migration</td>
</tr>
<tr>
<td>NILF</td>
<td>not in the labour force</td>
</tr>
<tr>
<td>NIS</td>
<td>New Immigrant Survey</td>
</tr>
<tr>
<td>NLSY</td>
<td>National Longitudinal Survey of Youth</td>
</tr>
<tr>
<td>NP</td>
<td>nondeterministic polynomial time</td>
</tr>
<tr>
<td>NPV</td>
<td>net present value</td>
</tr>
<tr>
<td>NSW</td>
<td>New South Wales</td>
</tr>
<tr>
<td>NT</td>
<td>Northern Territory</td>
</tr>
<tr>
<td>OLS</td>
<td>ordinary least squares</td>
</tr>
</tbody>
</table>
pdf  probability density function
PSID  Panel Study of Income Dynamics
PT  part-time
Q.E.D.  quod erat demonstrandum
Qld  Queensland
RA  remoteness area
RHS  right hand side
RRMA  Rural, Remote and Metropolitan Areas
SA  South Australia
SD  Statistical Division
SLA  Small Local Area
SOC  second order condition
UK  United Kingdom
US  United States
Vic  Victoria
WA  Western Australia
Tas  Tasmania
# Table of contents

1 Introduction: migration theories and previous evidence - 1 -

1.1 Migration as multidiscipline phenomenon - 2 -

1.2 Migration as an economic problem - 4 -

1.2.1 Theories of migration - 4 -

1.2.2 Empirics of migration - 9 -

1.2.2.1 Literature on causes of migration - 13 -

1.2.2.2 Literature on consequences of migration - 20 -

1.3 Focus of thesis - 22 -

2 Migration decisions: Micro level study of internal out-migration - 27 -

2.1 Introduction - 27 -

2.2 Hypotheses - 29 -

2.3 Empirical section - 35 -

2.3.1 Data - 35 -

2.3.2 Data: Remoteness and geography - 38 -

2.3.3 Descriptive statistics - 47 -

2.3.4 Methodology of econometric analysis - 52 -
4 Theoretical models: destination choice and search

4.1 Introduction

4.2 Model of destination choice

4.2.1 Exogenous N: no uncertainty

4.2.2 Exogenous N: uncertainty

4.2.3 Comparative statics

4.2.4 Endogenous information search

4.2.5 Examples: model of migration on circle

4.3 Migration as search

4.3.1 Model of search without recall

4.3.1.1 Case of finite T, \( T = \infty \)

4.3.1.2 Case of infinite T, \( T = \infty \)

4.3.2 Search without recall in general case: optimal sample size

4.3.2.1 Case of finite T, \( T < \infty \)

4.3.2.2 Case of infinite T, \( T = \infty \)

4.3.3 Migration as search with recall, \( N = 1 \)

4.3.4 Search with recall in general case: optimal sample size

4.3.4.1 Case of finite T, \( T < \infty \)
List of Tables and Diagrams

TABLE 1. MIGRATION THEORIES ACROSS DISCIPLINES - 3 -

TABLE 2. CLASSIFICATION OF REMOTENESS AREAS - 39 -

TABLE 3. TRANSITION MATRIX BETWEEN REMOTENESS AREAS FOR LONG DISTANCE MIGRANTS, PERCENTAGE - 41 -

TABLE 4. SUMMARY STATISTICS AND COMPARISON MIGRANTS VS. NON-MIGRANTS - 43 -

TABLE 5. (PREVIOUS TABLE CONTINUED) - 45 -

TABLE 6. PROPORTION OF MIGRANTS, BY WAVE, PERCENTAGE - 47 -

TABLE 7. PROPENSITIES TO MIGRATE, BY AGE AND EDUCATION, PERCENTAGE - 48 -

TABLE 8. PROPENSITIES TO MIGRATE, BY BROADLY DEFINED EMPLOYMENT STATUS, PERCENTAGE - 49 -

TABLE 9. PROPENSITIES TO MIGRATE, BY ANNUAL WAGE QUARTILE, PERCENTAGE - 50 -

TABLE 10. PROPENSITIES TO MIGRATE, BY HOME OWNERSHIP, PERCENTAGE - 51 -

TABLE 11. PROPENSITIES TO MIGRATE, BY REMOTENESS AREA, PERCENTAGE - 52 -

TABLE 12. PROBIT MODEL RESULTS, MARGINAL EFFECTS IN PERCENTS - 59 -

TABLE 13. INCOME AND INCOME INEQUALITY, STATES AND TERRITORIES, 1997-98 - 84 -

TABLE 14. DESCRIPTIVE STATISTICS, MOVERS VS. STAYERS - 109 -

TABLE 15. DISTRIBUTION OF THE SAMPLE BY DISTANCE OF MOVE, PERCENTAGE - 110 -

TABLE 16. DISTRIBUTION OF THE SAMPLE BY WAGE QUARTILES AND DISTANCE OF MOVE, PERCENTAGE - 110 -

TABLE 17. PROPORTION OF REAL WEEKLY WAGE GAINERS, BY QUARTILES AND DISTANCE MOVED, PERCENTAGE - 111 -
TABLE 18. GROWTH OF MEAN REAL WEEKLY WAGE, BY QUARTILES AND DISTANCE MOVED, PERCENTAGE

TABLE 19. GROWTH OF MEAN REAL HOURLY WAGE, BY QUARTILES AND DISTANCE MOVED, PERCENTAGE

TABLE 20. CHANGE OF MEAN WORKING HOURS, BY QUARTILES AND DISTANCE MOVED, PERCENTAGE

TABLE 21. GROWTH OF MEAN REAL ANNUAL WAGE, BY QUARTILES AND DISTANCE MOVED, PERCENTAGE

TABLE 22. SUMMARY OF THEORETICAL RESULTS

TABLE A1.1. (TABLE 12 CONTINUED). PROBIT MODEL RESULTS, MARGINAL EFFECTS IN PERCENTS

TABLE A1.2. PROPENSITIES TO MIGRATE INTERSTATE IN THE US, 1980-1985, BY AGE AND EDUCATION

TABLE A1.3. EVIDENCE FROM PUBLISHED STUDIES ON URBANIZATION OR COUNTERURBANIZATION TRENDS IN EUROPEAN COUNTRIES DURING THE 1970-1990S

TABLE A2.1. DEFINITION OF VARIABLES

TABLE A2.2. SUMMARY STATISTICS

TABLE A2.3. WAGE EQUATION, SYSTEM GMM RESULTS

DIAGRAM 1. DISTRIBUTION OF DISTANCE BY REMOTENESS AREA

DIAGRAM 2. THE EFFECT OF AGE ON THE PROPENSITY TO MIGRATE, %

DIAGRAM 3. THE EFFECT OF TENURE WITH EMPLOYER ON THE PROPENSITY TO MIGRATE, %

DIAGRAM 4. THE EFFECT OF HOME TENURE ON THE PROPENSITY TO MIGRATE, %
Diagram 5. Block diagram of decisions and payoffs, N destinations - 129 -

Diagram 6. Block diagram of decisions and payoffs, sequential search - 131 -

Diagram A1.1. Total commuters by distance commuted, 2001 – SLAs in Illawarra, Hunter, Sydney, and Inner Sydney SRs - 82 -
1. Introduction: migration theories and previous evidence

Human migration, hereafter labour migration or simply migration, is an interesting and intriguing interdisciplinary topic of research. Migration of labour compels the attention of economists no less than investment in capital and changes in total factor productivity. All three are important because they constitute the inputs in the production function. In part this close attention to migration is due to the lower costs of labour reallocation and therefore, the relative mobility of labour and the immobility of capital and technologies. There are a lot of migration theories in economics, which can be found in various surveys such as Ghatak et al. (1996), Greenwood (1997), Lucas (1997), Massey (2004), Molho (1986), and Shields and Shields (1989), and there is substantive empirical literature supporting or contradicting theoretical predictions. In essence almost all of them are built on the assumption of a rational microeconomic agent maximizing satisfaction from life. This satisfaction is represented in algebraic form by utility from consumption of various goods and services, and enjoyment from non-material things such as local amenities/disamenities, job, family, and friends. An example of a different theoretical approach is given in Massey et al. (1993) who consider risk minimization as another motivating factor for migrants. This idea is in line with the new economics of migration (NEM) to be discussed in Section 1.2.1.

The first chapter is an introduction to the thesis. Since economics is not the only social science studying migration, the first chapter outlines the disciplines involved, problems they address, and methodology used. After that it proceeds to economic approaches and then to the specific theoretical and empirical problems studied in this thesis; see section 1.3 on pages 21 to 25. The main body of the thesis, chapters two to four, relies on economic approaches and includes development of both theoretical and empirical models. In the second chapter migration decisions
are studied empirically using micro level data from the Household Income and Labour Dynamics in Australia (HILDA) survey. A panel binary dependent variable model is estimated by the probit model with random individual effects. The third chapter contains an empirical estimation of the economic consequences of migration, also using individual data from the HILDA survey. The dynamic earning equation estimated by the system GMM is presented in this chapter. The fourth chapter proceeds to theoretical economic models of migration. It considers two models of the migration decision based on an information search done either sequentially, similar to the process of a job/lower price search, or as a once off, after all potential destinations are evaluated. The latter is known as the fixed sample size (FSS) approach. A FSS model of migration on a circle with the randomized arrival of a set of destinations is elaborated. The remainder is the concluding part of the thesis.

1.1 Migration as multidiscipline phenomenon

Migration is a phenomenon without clear borders not only for migrants themselves but for scholars in all the social sciences. Not surprisingly different disciplines address similar questions such as who moves, when do they move, where do they move, and why do they move. These types of research questions are common to anthropologists, demographers, economists, historians, and sociologists (Brettell and Hollifield 2000). Although research aims are different, as well as tools and methods of analysis, both qualitative and quantitative, data used across disciplines basically come from the same sources. The source is either Census or other administrative records or individual/household level surveys. The unit of analysis depends on the research question and can vary from micro level (individual or household) to macro level (city, region, national, or international). The various approaches adopted by scholars in international migration are summarised in the following table:
### Table 1. Migration theories across disciplines

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Research Question(s)</th>
<th>Levels/Units of Analysis</th>
<th>Dominant Theories</th>
<th>Sample Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anthropology</td>
<td>How does migration effect cultural change and affect ethnic identity?</td>
<td>More micro/individuals, households, groups</td>
<td>Relational or structuralist and transnational</td>
<td>Social networks help maintain cultural difference</td>
</tr>
<tr>
<td>Demography</td>
<td>How does migration affect population change</td>
<td>More macro/populations</td>
<td>Rationalist (borrows heavily from economics)</td>
<td>Immigration increases birth rate</td>
</tr>
<tr>
<td>Economics</td>
<td>What explains the propensity to migrate and its effects?</td>
<td>More micro/individuals</td>
<td>Rationalist: cost-benefit and push-pull</td>
<td>Incorporation depends on the human capital of immigrants</td>
</tr>
<tr>
<td>History</td>
<td>How we understand the immigrant experience?</td>
<td>More micro/individuals and groups</td>
<td>Eschews theory and hypothesis testing</td>
<td>n/a</td>
</tr>
<tr>
<td>Law</td>
<td>How does the law influence migration?</td>
<td>Macro and micro/the political and legal system</td>
<td>Institutionalist and rationalist (borrows from all the social sciences)</td>
<td>Rights create incentive structures for migrants</td>
</tr>
<tr>
<td>Political Science</td>
<td>Why do states have difficulty controlling migration?</td>
<td>More macro/political and international systems</td>
<td>Institutionalist and rationalist</td>
<td>States are often captured by pro-immigrant interests</td>
</tr>
<tr>
<td>Sociology</td>
<td>What explains immigrant incorporation?</td>
<td>More macro/ethnic groups and social class</td>
<td>Structuralist and/or functionalist</td>
<td>Immigrant incorporation is dependent on social capital</td>
</tr>
</tbody>
</table>

Source: Brettell and Hollifield (2000).

One may note there are intersections in both research questions and dominant theories among disciplines. It is possible to add another discipline to this table, geography, which has a research agenda and methodology very similar to that in economics, demography, and sociology. It can be observed from the table that many other disciplines besides economics share the theory of economic
rationality in international migration. It is only economists who seem to rely mostly on their own
economic approach while other researchers borrow ideas from other disciplines.

Although the focus of this thesis is on internal migration, that is migration within a country, rather
than on international, the discussion above is relevant because the research questions, data, and
methodology are basically the same in both. One can conclude from the table above that migration
in general and internal migration in particular is a multifaceted phenomenon attracting the attention
of all social sciences. Furthermore, the table demonstrates that, simply using a single discipline
approach without coordination of efforts across disciplines, a full picture is unlikely to be
developed.

1.2 Migration as an economic problem

There are a number of models developed by economists in the last five decades. This section
outlines the theoretical models, which have dominated the economics literature at different times,
and empirical verifications for some of them. In no way does the list of models pretend to be
exhaustive. The discussion of the models relies mostly on surveys in Massey (2004), Molho (1986),
Greenwood (1997), and Lucas (1997).

1.2.1 Theories of migration

Neoclassical economics, the oldest theoretical approach, models migration as a response to
geographical differences in labour supply and demand (Lewis 1954, Ranis and Fei 1961). Workers
move from regions with excess supply of labour, which are characterised by lower wages, towards
regions with excess demand and therefore, higher equilibrium wages. At equilibrium, wage
differentials between regions still persist but at lower levels, reflecting the costs of moving between
regions. This model focussing on equilibrium in the national labour market with mobile labour is
very simplistic since it does not consider heterogeneity of workers in and across regions.
Another macro-economic model of migration, the Harris-Todaro model, was developed to explain migration from undeveloped rural to developed urban areas (Harris and Todaro 1970 and Todaro 1976). A rural-urban migration stream is observed when risk-free but lower paid jobs in the rural sector are less attractive than higher paid urban jobs with a risk of unemployment. As in the previous model, migration continues as long as there are gains to it. There will be no migration once equilibrium is achieved, that is when the rural wage is equalized with the expected urban wage less the costs of migration. This model and its modifications are popular among researchers of developing countries where rural-urban migration continues to be significant. This approach, however, has not been very successful in explaining some observed tendencies in lower income countries. Lucas (1997, p. 738) notes that

“... it seems the average duration of initial unemployment may not be very long, that a substantial portion of migrants have identified their job prior to moving, and by no means all informal sector entrants subsequently transfer to the formal sector. This scenario does not fit well with the Todaro or Harris-Todaro models.”

A macroeconomic model based on gravity law is an example of a model which is still popular in applied economic research, not only regarding labour mobility but also transport and international trade. The gravity model which originated in human geography emphasised space as a determinant of migration (Molho 1986). Greenwood (1997) refers to a paper of Princeton astronomer Stewart (1941) who noted that the travel distance to students’ home towns obeys the law of gravitation. Stewart developed a spatial interactions model in which, like the Newtonian universe gravity law, distance between objects plays a negative role. In the gravity model gross migration flows between any pair of countries (also they could be regions, cities, or localities) are modelled as a function of the masses of the two areas, e.g. population size or GDP, with some distance deterrence function. In
the modified gravity model economists add a set of pull and push factors, such as real income, unemployment rate, public goods provision, amenities (climate, terrain, sea costs, and other) for both the source and host regions. In addition, the degree of separation between regions can be approximated not only by geographical distance but also by cultural and linguistic distances (ethno-linguistic fractionalization and language proximity) which are proxies for other barriers to mobility. They may be less important if there is a “friends and relatives” effect as proposed by Greenwood (1969), who found that the stock of migrants from the source to the host regions reduces the negative effect of distance on migration flows. In empirical applications the popular gravity model can also be derived as a consequence of an aggregation of the following micro level approach.

The most widely known approach so far is the microeconomic model of migration as investment in human capital developed in pioneer publications by Becker (1962), who proposed the general idea, and in particular by Sjaastad (1962) with an emphasis on migration. A migration decision is based on the estimated streams of costs and benefits associated with life in a particular destination, which depend also on the source area. The theory offered the explanation for the way migration decisions are made. It is the selection of the optimal destination with the maximum present value of income gain for person. A spontaneous question may arise here, “why is this model linked to human capital formation?” A possible answer can be found in Massey (2004). Migration inevitably causes the learning of a new environment and possibly a different culture and the development of new skills which are conducive to adaptation. A defect of the human capital approach, as noted in Molho (1986), is in the treatment of the information. The process of acquiring information about costs and benefits at potential destinations before a decision can be made is unclear.

The model of migration as the investment decision of a rational agent is the disequilibrium approach, based on the potential of speculative migration. The alternative, the equilibrium approach, is built on the assumption that there are no such opportunities, that is, individual utility is
constant and wages and prices are in equilibrium at any point in time and that any existing variations are compensating differentials for unequal amenities and quality of life. Therefore, migration leading to an improvement in one characteristic is inevitably linked with deterioration in at least one other.

The utility of the disequilibrium approach is demonstrated by empirical applications developed thereafter. When the basic model is extended to include uncertainty, attitudes to risk, and discount rate, the model proposes a variety of economic, demographic, social, and environmental factors influencing individual migration decisions, not only wage and unemployment rate differentials as in the preceding two neoclassical models. This extension gives rise to an empirical model similar to the modified gravity model.

Further theoretical development of microeconomic theory stemmed from the need for a deeper understanding of the migration decision process. The decision making was split up into two or three stages, the decision to migrate and selection of a destination in the first one or two stages and then a process of search. As a result of this approach two types of migration were distinguished in the economic literature: speculative migration, when migrants are not searching for a job prior to a move, and contracted migration, when they find a job and then migrate. This evolution of the theory was mostly due to the development of search models. In the very beginning the approach was too simplistic. In a problem considered in Stigler (1961) duration of search was not a random variable but rather a fixed number of periods during which a buyer would search a fixed number of shops and then go to one offering the minimum price. This approach was called the fixed sample size (FSS) model. Economics of search, or more recently, economics of information has started to flourish in 1970s when the model of sequential search was developed. In terms of its application to job search, the search period or, in other words, the unemployment spell, must be continued as long as the expected gain from an extra job searching period of time exceeds the costs of the extra period.
(Lippman and McCall 1976). The equivalent “optimal stopping rule” of search is derived from an equality of costs and benefits: stop searching and accept the first job offer which is greater or equal to the reservation wage. Later in the 1980s this model was generalized to equilibrium search models and also to a game theoretical model of search and bargaining (Lippman and McCall 2004, Pissarides 2004). The search model and the FSS model in application to migration will be discussed in Chapter 4 in more detail.

Two authors who extended the search model deserve to be mentioned here. The first is David (1974) who proposed a model of speculative migration. In it the migration decision is done before the process of job search which is modelled not traditionally as FSS. Only when all offers are collected, the best one is selected. Migration is shown to be in the direction of the maximum expected variance of wage offers relative to the average offer. Some elements of this model are also used in one of the models presented in Chapter 4 of the thesis. Another extension is by Rogerson (1982), also with a model of speculative migration in which multiple labour markets are presented. Then under some conditions there will be migration to the labour market with the highest reservation wage followed by the standard job search.

Since that time migration research has been considering a more complex structure of migration decisions. In particular, Mincer (1978) elaborated a family level approach in his economics of family migration. According to it, migration of family is discouraged by family ties which are negative externalities usually internalized by families (Greenwood 1997). From the empirical point of view the presence of “tied” family members, that is those whose individual gain is less than the gain of other family members, decreases the family propensity to migrate. Generally, the wife is a tied mover who may even have a loss in income which is compensated by a gain in the husband’s income. This theory generated great interest within the profession and led to a number of hypotheses being tested in the empirical literature, which is still expanding.
Another approach was developed relatively recently in the framework of the new economics of labour migration (Stark 1991). As noted by Ghatak (1996), Stark shifted the focus of migration research from individual independence to mutual interdependence. Involving a collective decision, it contrasts with the neoclassical micro and macro models. Since financial markets are imperfect and incomplete, family-decision making is a risk diversification which results in allocation of family labour resources over geographically dispersed markets. Migration of a family member is perceived as an investment in a new income source which maximises family income streams. In addition to risk diversification, another motivation, proposed in the same book (Stark 1991), is relative deprivation. It implies that the family is seeking a way to improve its relative position in the local community wealth hierarchy. Sending someone to work in another city or country and remit some part of their earnings home may increase the household’s income relative to others and therefore, reduce relative deprivation compared with some reference group.

1.2.2 Empirics of migration

The empirical literature on causes and consequences of internal migration is vast. To date the two surveys by Greenwood (1997) and Lucas (1997) published in the *Handbook of Population and Family Economics* are the most comprehensive overviews of research in the economics of internal migration. The former is based on around 200 theoretical and empirical papers with the focus on developed countries and the latter paper reviews about 270 books and papers on migration research in the context of developing countries.

Empirical problems studied by economics are much wider than those briefly outlined above in section 1.1. Thus, in his papers Greenwood (1997, 2004) names the following traditional questions addressed by empirical research on internal migration:

(1) Who migrates?
(2) Why do they migrate?

(3) Where do they come from and where do they move?

(4) When do they migrate?

(5) What are the consequences for them and for other people in the sending and the receiving areas?

The main focus of the two empirical parts of the thesis, detailed in section 1.3 below, is on questions 2 and 5 about the determinants and consequences of migration decisions on an individual level, but question 1 will also be touched on tangentially because individual characteristics of respondents are to be controlled for in the analysis.

Both micro and macro level approaches are adopted by scholars looking for answers to these questions. In the very beginning most research was on aggregate level data but currently it is probably fair to say that micro level studies have gradually overtaken macro studies in volume. This is due to a two-sided tendency: greater availability of individual data coming from various surveys on the one side and progressing methodology of its analysis on the other side. Some researchers disagree with this observation, for example, as noted in Greenwood (1997, p. 659)

“... even in the presence of many microdata sets, aggregate data are frequently studied today. Not only are aggregate trends and tendencies of interest in their own right, but also for many countries such data are all that is available. Thus, because much still can be learned from studying aggregate data and because they will remain a major source of information concerning migration, some attention to the use of aggregate data is appropriate.”
The main advantage of the macro level approach used by earlier researchers in the migration field is the wide availability of official statistics on an aggregate level by regions and countries, not only for developed by also for less developed countries. Another advantage is in the simplification of the sophisticated models that have arrived in the migration literature due to the micro level theoretical approach discussed above. Two disadvantages of the macro approach are to be noted. The first, aggregate characteristics of the source and host regions do not necessarily reflect the characteristics of migrants between these two regions because migrants are known to be self-selected (e.g. Greenwood 1997, Borjas 1987 and 1991). Positive self-selection occurs when migrants are selected from among more productive workers than the average comparable worker in the origin and destination while negative selection implies that migrants are less productive than the average in both labour markets. Inability to control the average migrant’s characteristics may lead to incorrect conclusions. The second drawback of macro models is the missing effect of individual variables such as age, gender, education, employment and marital statuses. All of these are important factors in individual migration decisions. For example, even though theoretically a higher unemployment rate is a push factor, this was studied mostly at the aggregate level, often with ambiguous conclusions. The problem was not resolved until the micro study of DaVanzo (1978) who found that the unemployment status increases the propensity to migrate.

Empirical studies of the determinants of migration decisions have been progressing from mostly descriptive early studies to sophisticated statistical modelling at present. The first empirical study of this sort was “The Laws of Migration” by Ravenstein (1885) who derived conclusions from a descriptive analysis of Census data in the United Kingdom (UK) and called them “laws”. According to Greenwood’s (1997, pp. 659-660) classification they are as follows:

(1) Most migrants move only a short distance and then typically to major cities.
(2) Rapidly growing cities are populated by migrants from nearby rural areas. In turn, the “gaps” left in the rural population are filled by migrants from more distant areas.

(3) The process of dispersion is the inverse of the process of absorption and exhibits similar features.

(4) Each main current of migration produces a compensating countercurrent.

(5) Long-distance migrants tend to move to major cities.

(6) Rural people have a higher propensity to migrate than urban people.

(7) Women have a higher propensity to migrate than men.

Two other laws, not usually quoted but important for economics, appeared in the second paper of Ravenstein (1889) under the same title “The Laws of Migration” (Lee 1966). Following Lee paper, one can extend the list of laws by incorporating these two:

(8) Technology and migration

“Does migration increase? I believe so! ... Whenever I was able to make a comparison I found that an increase in the means of locomotion and a development of manufacturers and commerce have led to an increase in migration.”

(9) Dominance of the economic motive

“Bad or oppressive laws, heavy taxation, an unattractive climate, uncongenial social surroundings, and even compulsion (slave trade, transportation), all have produced and are still producing currents of migration but none of these
currents can compare in volume with that which arises from the desire inherent in most men to ‘better’ themselves in material respects.”

These laws demonstrated the richness of micro level statistics, from Census data in particular, and also demonstrated that even a simple descriptive study may generate hypotheses that can be tested by multivariate analysis. Some of these nine general observations were later explored by researchers on migration.

The focus of the present thesis is on micro level studies of internal migration since they are more influential nowadays. Hypotheses (1), (5)-(7), and (9) from the list of Ravenstein’s laws are addressed in the empirical Chapters 2 and 3, albeit implicitly, since they are not the main focus. A selected list of both macro and micro level empirical economic studies of migration will be reviewed in the following two sections.

1.2.2.1 Literature on causes of migration

Relying mostly on the disequilibrium approach economic analysis has demonstrated both theoretically and empirically the significance of various demographic, human capital and other individual and aggregate level characteristics. Due to the multifaceted nature of the phenomenon, the number of factors affecting migration decisions is enormous and, therefore, comprehensive micro level data are strongly demanded in such studies. In spite of the micro level approach being the core of the theory, until recently migration studies were based on aggregate level data for regions and countries mostly due to their wide availability.

There is unsurprising unanimity across disciplines in the migration literature that migration probability declines with age starting from the twenties and grows with educational level (Greenwood 1997). The negative effect of age is explained by economists as a shorter time interval for investment to bring returns (Borjas 2004). The positive effect of educational attainment on
mobility is linked to the better ability to collect and evaluate information about potential destinations by the better educated, which brings about higher returns (Greenwood 1997). Borjas (2004) also noted that education broadens the borders of the labour market, e.g. college professors have almost fully transferable skills in a national or even an international labour market.

Other personal demographic characteristics such as marital status, family size, number and age of children reflect the costs of migration and therefore should serve as impediments to mobility. A family migration decision is a combined strategy with outcomes often very different from those obtained in simpler individual decisions. Analysis becomes more complex by the presence of “tied” movers, who may be unemployed or not in the labour force (NILF) after migration or lose some part of their income and therefore, are worse off but the entire family is better off (Mincer 1978). For simplicity of the empirical analysis family migration decisions are not studied in the thesis.

Both Greenwood and Lucas underline the importance of income and employment in the migration decision. The first theoretical model of economically rational migration, the Sjaastad (1962) model described above, was later extended by Todaro (1969), who modelled rural-urban migration and replaced Sjaastad’s known urban (but not rural) income by the expected value of income.

The macro migration equation specified in Lucas (1997, p.740) and in only slightly different form in Greenwood (1997, p. 663) is a modified gravity model widely used in migration and international trade research

\[ m_{ij} = m(w_i, w_j, d_{ij}, A_i, A_j, \epsilon) \]

where \( m_{ij} \) is the number of population in region \( i \) migrated to region \( j \), \( w_i \) and \( w_j \) are the mean wages in regions \( i \) and \( j \), \( d_{ij} \) is the distance between \( i \) and \( j \), \( A_i \) and \( A_j \) are the vectors of region \( i \) and \( j \) characteristics, and \( \epsilon \) is a stochastic disturbance. Regional characteristics commonly included in
this model are population, average wage, unemployment rate, degree of urbanization, climate, topology, fiscal variables, and others.

This model has been widely estimated using macro level data in many countries. Sjaastad’s model of migration as investment in human capital begets the hypothesis that in the disequilibrium case, when wages are not homogeneous across regions, the mean wage in the source region is expected to have the negative sign and the mean wage in the destination is expected to have the positive sign in the model shown in equation (1). In line with this hypothesis is an earlier conjecture of Hicks, that the main cause of migration is difference in wages (Hicks 1932, p. 76). As Greenwood (1997, p. 670) concludes after review of many studies

“Based on aggregate data, empirical findings associated with income, earnings, and wage variables in modified gravity models have not been uniformly strong, although it is probably fair to conclude that the weight of available evidence favours Hicks's expected disequilibrium results …”

The gravity model has identified in particular a negative effect of distance on migration flows (Andrienko and Guriev 2004). Distance elasticity of migration typically varies in the range between -0.1 to well over -2.0 depending upon the population subgroup, type of migration, time period, geographical area, and other explanatory variables in the gravity model but this elasticity is similar in terms of order of magnitude across countries (Greenwood 1997). The proposed explanation of the negative impact stems from the relationship between information and psychic costs and physical distance between locations. The psychic cost of moving depends positively on distance to the destination as far as it could be offset by more frequent trips back (Schwartz 1973). Availability of information diminishes with distance and costs of information search increase with distance (Greenwood 1997).
A common finding from the modified gravity model with respect to economic variables is that income and job opportunities in receiving regions are important in contrast to the sending region’s characteristics (Shaw 1985). In a previous paper of mine the modified gravity model was estimated on a matrix of interregional migration flows between Russian regions from 1992 to 1999 (Andrienko and Guriev 2004). Results are in line with other migration studies, in particular with regard to the negative effect of distance. Elasticity of migration with respect to distance is equal to -0.9, very close to estimates obtained from the modified gravity model in the US, Spain, and China. Results also confirm the importance of a large number of economic, social, and demographic variables in sending and receiving regions such as real income, unemployment, poverty, education, climate, geography, education, demographic and ethnic structure, urbanization, resource potential, reform indicators, conflict, and public good provision. In particular, one interesting and unusual finding showed that liquidity constraints impede interregional migration and therefore, the population of the poorest regions is in a poverty trap being unable to finance the cost of moving.

The shortcoming of the modified gravity model is that it cannot be used in micro level analysis of individual migration decisions because wage is observed only in the current region at a given time. However, if one can predict individual wages across regions based on both observed and unobserved individual characteristics, it is possible to estimate a destination choice model. There is an advantage in this model anyway. A matrix of flows on the left hand side (LHS) of the model in equation 1 gives many observations (the number of regions squared) even for small countries. This distinction, rarely met in the official statistics, favours economics of migration and some other areas. Other examples are transportation flows, international or internal trade flows, but these are not so compact and not so readily collectable and suitable for analysis. Because of its vastness and the potential to study the impact of any aggregate level variable and to receive statistically
significant results, the modified gravity model has been widely used in migration modelling up to date and presumably will continue to be used in the future.

Much more insight into the causes of migration became possible with the availability of individual level data. Personal unemployment is one example of such causes. It was shown to be one of the main reasons for migration. A review of ten micro level studies, in the USA, UK, and Netherlands demonstrated that personal unemployment increases migration likelihood in almost all nine of the studies (Herzog et al. 1993). The effect of non-participation in the labour force on mobility is not so well investigated. However, a study of Eliasson et al. (2003) has shown that for individuals with the NILF status the probability is significantly higher than for the employed.

Micro level studies using income are surprisingly infrequent, possibly due to missing values and difficulties with individual income measurement in the origin and absent information about potential income across destinations. Nevertheless research has demonstrated that income differential is important in explaining migration decisions. DaVanzo (1978) used data on 1,609 couples from the Panel Study of Income Dynamics (PSID) and showed that for families with an employed husband looking for another job (but not unemployed or not looking for another job) the propensity to migrate between 1971 and 1972 depended positively on potential wage returns to family migration.

Individual mobility has been found to decline significantly with income at the origin (Goss and Paul 1986). The authors estimated a logit model of migration between 1974 and 1975 on a sample of heads of households who were in the labour force in 1974. However, another micro level migration study by Goss and Schoening (1984) on almost the same sample of individuals from PSID did not confirm this result. Real income from wages in 1977 was found to be negative and insignificant. Surprisingly, the first paper did not cite the second one and did not use the unemployed status and
duration of unemployment, both found to be significant in the previous study, but instead added tenure and work experience. Estimated coefficients for the income variable in both papers look inconsistent, perhaps, because they used lognormally distributed nominal wage instead of logarithm of wage. Thus, Goss and Schoening showed two insignificant coefficients for the real income in 1977 different by factor 10, \(-0.013 \times 10^{-3}\) and \(-0.128 \times 10^{-3}\), in two slightly different regressions with other coefficients being very close. In contrast, Goss and Paul found the intermediate value for income in 1974 coefficient, \(-0.049 \times 10^{-3}\), to be highly significant. Finally, there were two more recent papers by these authors (Goss and Paul 1990a and 1990b). They estimated, among other things, the effect of real income on migration between 1981 and 1982 in the former and between 1982 and 1983 in the latter on the sample of heads of household from PSID who were in the labour force in 1981 and 1982 in the former and between 1982 and 1983 and unemployed at some time in 1982 in the latter. This time the effect was significantly positive with the coefficient \(0.016 \times 10^{-3}\) in the former paper and insignificantly positive \(0.016 \times 10^{-3}\) in the latter. Sadly, in three of the four papers the authors did not provide the summary statistics and only in one reported the mean income but not the minimum. Therefore, the reader does not know whether individuals with zero income are excluded from the sample in which case a sample selection problem arises.

Results of DaVanzo (1978) concerning individual wage in the current area are mixed. She did a number of regressions with a set of income variables and different definitions of migration and demonstrated that the family income effect is negative. The effect of a husband’s wage on family migration is unclear. For an employed husband the effect is not stable but for an unemployed one the effect of the imputed wage is positive and significant. The compensated or ‘pure’ effect of the husband’s wage, which takes into account that the family income includes the husband’s wage, was not calculated by DaVanzo but seems to be negative in her regressions.
In the analysis of migration in the sociological literature one may find contextual variables, constructed on the community level. They have not been studied in the economics literature so widely. However, many of them allow economic interpretation, e.g. commercialization and community development are measures of tightness and risks in the local labour market (Lucas 1997, p.754). A notable exception is the new economics of migration with, for example, studies on relative deprivation (Stark 1991).

The most serious empirical shortcoming in both equilibrium and disequilibrium approaches acknowledged by researchers is that there are too many unobserved variables, including motivation and abilities. Economists simply model migration as life-time utility maximization. In many instances the potential migrant knows neither the exact wage and costs of living, nor the stability of a found job. Even if one imagines someone higher educated, like a person with a doctorate in economics or mathematics, who knows the distribution of these variables (as economists imply in their models) in the considered destinations, no one seems to possess a clear comprehension of all possible (types of) costs associated with migration, not to speak of their distribution. This is what researchers relegate to a generalized notion of “uncertainty”, which is a part of the world and is unavoidable.

The only empirical study of internal migration decisions in Australia is a working paper by Mitchell (2008) exploring the first five waves of the HILDA survey. Among other equations he estimated the migration decision equation but did not explore the richness of the panel survey. In particular, individual unobserved characteristics such as skills and ability were not controlled in his analysis. He only applied an individual clustering correction to the standard errors, which is only a weak substitute for the individual effects. Two definitions of migration were used in his paper. A migrant is a person who is living at a new address in the current wave and either: (1) changed Small Local Area (SLA), a relatively small administrative unit, or (2) moved a distance over 30 km. Results of
the probit model (though only described in words without tabulation) suggest that lower educated (in the author’s terminology, lower skilled, defined as individuals with less than 12 years of school and no further qualification), older, and employed people are less mobile as are people from households with an employed spouse and more children, owning a house, living in non-metropolitan areas, relatively advantaged regions, and with a better social network in their neighbourhood.

1.2.2.2 Literature on consequences of migration

Consequences of internal migration are studied mostly on a macro level in both developed and developing countries. These studies mostly focus on labour market characteristics such as average wages, expected wages, and unemployment rates at the origin and destination. Greenwood et al. (1981) found inter-state in-migration increases wages in Mexico but out-migration has no effect, explaining that enhanced demand for local non-tradable goods dominates the negative impact of increased labour supply. Similar results were obtained in a study by Garcia-Ferrer (1980) of inter-provincial migration in Spain but the out-migration rate was found to decline wages. Differences in unemployment rates and wages between South and North Italy were shown to decline with migration (Salvatore 1980).

There is some empirical evidence at the micro level though it is mostly restricted to individual returns to migration. Empirical studies have shown a positive effect of migration on earnings consistent at least in the long-run. This effect depends on a comparison group. Earnings of internal migrants in the US were shown to be lower than earnings of non-migrants with the same level of education residing in the receiving areas (Lansing and Morgan 1967). However, when migrants from a given area are compared with similar people left behind, Lansing and Morgan found that movers off the farm and from the Deep South were able to earn a wage premium. A similar finding about positive returns to migration was obtained from tabulation constructed on another individual
data set in the US (Gallaway 1969). Gallaway’s conclusion was shown to be not dependent on characteristics of migrants, in particular, their age groups (Cox 1971). Another American study found that black migrants from South to North cities earn more than non-migrants in both the sending and receiving cities, although recent migrants are worse off (Masters 1972). Wertheimer (1970) has estimated returns to South to North and rural to urban migration in the short-run and in the long-run. His findings say that gains do not exist in the first five years but they are significant thereafter. Estimates obtained for individual returns to internal migration in other countries such as Canada, Turkey, and Venezuela are in line with these conclusions (Greenwood 1997). All of these studies employed cross-sectional data, with a single observation per individual. Initial studies on longitudinal micro level surveys, known also as “panel” data, confirmed results from cross-sectional studies, e.g. in Canada (Grant and Vanderkamp 1980) and in the US (Borjas et al. 1992). Perhaps somewhat surprisingly, recent studies on micro level panel data found a positive wage premium for internal migration for men even in the short-run (in the USA, Yankow 1999 and 2003; in Britain, Bőheim and Taylor 2007). In particular Yankow (2003) has shown the increasing in time effect of migration on returns with a maximum 11 percent wage premium after over five post-migration years. Similarly, in Britain a positive premium is found in the two-year period following migration but with an earlier peak at 3 percent at the first year.

Another effect of migration which was studied in the empirical economics literature is the impact on employment post-migration. Herzog et al. (1993) in their survey provide ambiguous conclusions equally distributed across positive, negative, and neutral findings. In particular, they found that migration decreases (Herzog and Schlottmann 1984; Bailey 1991; and Shumway 1991), increases (Vandijk et al. 1989; and Goss et al. 1993) or does not change (Herzog and Schlottmann 1984; Vandijk et al. 1989; and Podgursky and Swaim 1990) the re-employment likelihood of the unemployed (Herzog et al. 1993, p. 335). A simple descriptive analysis in Mitchell (2008, p. 27) on
Australian data is mostly in line with a similar analysis of the US micro level data (Saben 1964; Herzog and Schlottmann 1984; and Herz 1991). Mitchell showed that migration increases chances to be employed one year post-migration for both the unemployed and NILF. More sophisticated regression analysis in this paper demonstrated that internal migrants are in general less likely to be employed post-migration, apparently due to lower re-employability of those employed pre-migration (Mitchell 2008, p. 31).

A relatively recent stream of migration research treats the migration decision as undertaken not by individuals but rather by families. A household head and tied person may have opposite effects on their earnings. Indeed, the author of the traditional earning equation, Mincer (1978), reports a positive effect of migration on earnings of men but a negative effect on the earnings of women. Conclusions obtained from panel data analysis in Sandell (1977) indicate similar outcomes, that wage of the husband increases post-migration but even though wage of the wife temporary decreases, family earnings are higher. In another study, Long (1974) found that migration reduces the wife’s labour force participation.

1.3 Focus of thesis

In the empirical perspective, although the reasons and consequences of migration are well studied in many countries, especially on an aggregate level, a number of white spots remain unknown to researchers. These basically include the economic causes and consequences of migration at an individual level. In particular, labour economists know only a little about individual and household employment behaviour, which includes the dynamics of labour force status and income streams, both prior to migration and after it, in the short/medium/long run. They also do not know whether migration is a simple speculative move from a lower paid to a better paid area or a more complex human capital investment, in line with educational and career decisions, maximising
individual/household present value of utility. It can also be family income risk management, that is, maximization of the expected value of NPV of utility of the whole family. Although many problems remain to be further investigated, researchers already have tools and data to rely on. Availability of rich micro level data stimulates such research. This thesis hopes to shed some light on some of these problems.

Furthermore, migration economic literature is missing research based on Australian microdata. The HILDA survey, the sole longitudinal survey of households in Australia, supplies data necessary for migration research because it tracks individuals in the case they move within the country. The Census is another good source of such data because individuals are asked about their place of residence twelve months ago though it is not a longitudinal survey and has only limited information.

The focus of the empirical part of the thesis, which comprises Chapters 2 and 3, is on some aspects of the reasons for, and outcomes of, migration. Availability of rich micro level data allows us to test a number of hypotheses about the individual migration process.

In Chapter 2, the causes of migration decisions in Australia, derived from the existing theoretical and empirical evidence, are studied. The potential causes cover several categories: demographic characteristics (age, gender, marital status, and number of children), economic variables (employment status, occupation, tenure, and annual wage), human capital (educational attainment), migration experience (years at current address and age left home for the first time), home ownership, and regional variables (remoteness area and state/territory). It is advisable that all of the determinants of migration decision are included in the model in order to avoid missing variable bias. However, some variables important for certain potential migrants such as conditions of land and housing markets, as well as time invariant topological, climatological, and environmental amenities, are not studied due to difficulties in obtaining them. On the other hand, the differences in
amenities are partially capitalized in the labour market. A binary model of migration decision is estimated on a micro level panel from a longitudinal survey of Australian households, comprising seven waves of the HILDA survey. The probit model can only be estimated with individual random effects which incorporate unobserved individual characteristics such as abilities, risk aversion, inclination to travel, ties to the current location, and others. The random effect is assumed to be uncorrelated with observed individual characteristics. The alternative, the fixed effects probit model, is not applied since it is known to give inconsistent estimates on small samples due to an incidental parameters problem. Since the HILDA survey tracks internal migrants and because not all respondents are interviewed each wave, an unbalanced panel is used for the model construction. The model is estimated on the entire sample of adults, excluding only full time students and individuals with omitted core variables.

The novelty of the study in Chapter 2 is in regards to its new results, in particular, the significant effect of such economic variables as NILF and annual wage and in the positive effect of geographical area remoteness from larger cities which serves as a proxy for accessibility to services. This is the first such comprehensive economics study of internal migration decisions in Australia to the author’s knowledge.

Wage returns to migration are studied in Chapter 3. Economic theory suggests expected wage differentials are among the reasons driving migration. But the study in Chapter 2 finds a positive effect of wage on outmigration. Workers who are unemployed or lower paid are more likely to move than others but these hypotheses are partially confirmed in Chapter 2 regarding unemployment only. A hypothesis studied in Chapter 3 is that wage increases as an outcome of migration. The chapter tests this hypothesis by estimating the earnings equation on the same individual panel data as in the previous chapter but with a reduced number of observations due to missing data on wages.
The main econometric problem in migration studies, perhaps, is that the migration dummy variable in the wage equation can be endogenous. This sort of drawback can be found in many if not all papers on migration outcomes. The thesis addresses the endogeneity problem considering a continuous variable, the distance of migration, instead of using a migration dummy. Distance of move is correlated with costs of migration and, therefore, workers moving larger distances are expected to have a higher wage premium in order to cover these extra costs. Distance is modelled to be either an exogenous or predetermined variable in the context of the dynamic panel model estimated by a system Generalised Method of Moments (GMM). The main finding is that in the short-run there are returns to distance for better educated and lower income workers.

In a more general theoretical perspective, one does not know how people collect and analyse information which is necessary to inform their migration decisions. Theories of migration tend only to explain how decisions are made on available information but mostly lack models which may help to understand these processes. Understanding of information accumulation seems to be crucial in the development of comprehensive models.

Some aspects of the economic theory of migration, different from those widely studied by scholars, are considered in Chapter 4. A model of migration as an outcome of search is studied in a more general form than simply as a search for job. It does not assume that only the job is important but takes into consideration all other components such as quality of life, amenities, price level, income, and others. In other words the model considers search for maximum utility. In the models presented in this thesis information about migration destinations is modelled to arrive in two different ways.

A first model, using so called fixed sample search, assumes that information about a fixed number of random destinations is collected during the period of search and the best destination in terms of utility is selected at the end of the period. The best destination may coincide with the current place
and, therefore, the search may not result in migration. In another case there will be migration to the best destination. The number of destinations is modelled to be either exogenously given or endogenously determined by a searcher. In the former the distribution of utilities is unknown to the searcher and therefore, the participation decision is not investigated. Comparative statics of the probability to migrate is studied in a special case with a normally distributed utility and costs of migration. For the latter case the distribution of utilities is given. An important result obtained is that there are only corner solutions to the expected utility maximization problem in the endogenous case when an individual needs to decide how much information to search. Then, in an Appendix to Chapter 4, a model of migration on a circle with costless information and uniform distribution of utilities and cost of migration is developed for an exogenously given number of random destinations. It identifies the proportion of agents who migrate, average distance they move, average utility of stayers, and average utility of migrants before and after move.

A second theoretical model considered in Chapter 4 assumes that information about the level of utility in a random destination arrives sequentially with known distribution. It cannot be rational for a potential migrant to wait until the end of the period, the assumption made in the first model. Thus, the sample size is modelled as a random variable. Then the chapter proceeds to a discussion of different assumptions about the search process, whether it is finite or infinite, with recall or without it, and shows their effect on individual migration decisions.

In brief, the structure of the remaining parts of this thesis is as follows. It presents the empirical model of individual migration decisions in Chapter 2. The wage returns to migration on individual level are estimated in Chapter 3. Chapter 4, discussing two theoretical models of migration, completes the main body of the thesis. Finally, Chapter 5 concludes and presents ideas on future directions of research.
2 Migration decisions: Micro level study of internal out-migration

2.1 Introduction

Individuals move geographically, or as economists often call this process, “vote with their feet”, for a better quality of life which depends not least on wages, provision of public goods, and amenities. Better prospects in other places are the main forces behind the migration decision. These prospects are determined by both the general economic environment and individual circumstances. There are many non-economic reasons for migration which are closely linked to life-cycle events such as the start/completion of higher education, marriage/divorce, birth/raising/educating children, retirement, death of partner, and others. It would be quite naive to think that in all of these life events pecuniary motives are irrelevant. This chapter outlines the theoretically justified determinants of migration decisions and evaluates their quantitative effects which can be derived from micro-level data, particularly in relation to individual employment and well-being characteristics.

Economic theory assumes that economic motives are the most important in migration decisions and that the expected income of a person and/or entire household is maximised in a single destination. In general the migration decision consists of a chain of destinations and timing of migrations if an agent finds this chain is optimal. Hence, it is quite likely that a single search or chain of migrations may end up at the origin (say, the place of birth) or even generate a decision not to depart from the original location. Ideally, in their analysis researchers would prefer to have in their hands entire life-course observations of the individuals and all household members. This would include a calendar and details of events such as primary/secondary/tertiary education, job and its characteristics, household composition, home address, and health. Such comprehensive data collection seems to be attainable in the near future.
The economics approach to the migration decision applied here is the traditional investment in the human capital approach which seeks an optimal behaviour based on life-time costs and returns. One important aspect of this approach that seems to be overlooked is the role of financial constraints. Investment in a person or family relocation could be very costly and sometimes unaffordable for relatively lower income individuals and even for rich families due to financial constraints. These constraints are binding for presumably many individuals and households because of significant observed and unobserved costs in the search for potential destinations and a new job search, psychic and transportation costs of moving, costs of adaptation in a new location, and others. Both the presence of uncertainty and the absence of full information are in theory costly for risk averse agents, though clearly the modern informational era and technological progress diminish their effect. Thus, an optimal decision on migration relies on individual and household circumstances, certainty of a job in the current location and in all considered destinations, and the expected wage levels across destinations included in an individual’s search process.

This chapter considers various push and pull factors surrounding migration decisions at the micro and macro, i.e. individual and community, levels. Individual wage and employment characteristics such as employment and NILF status, tenure, employment and unemployment spells are the main focus of the analysis in this chapter.

This chapter is organized in five parts. First, it derives hypotheses from both theoretical and empirical evidence. Second, it describes the data and methodology of the empirical analysis. In the third part an empirical model of migration decisions is estimated on the individual level data from the HILDA survey and verified for robustness. Fourth section discusses the results obtained in the previous part. Then the fifth part shows directions of future research. Finally, it summarises the main findings.
2.2 Hypotheses

The job search literature discusses differences in job search between an employed person, who is searching for a job, and an unemployed person. Theoretically, in the framework of the sequential job search model a currently employed person has at least the same reservation wage as an unemployed person, *ceteris paribus*. Under the assumption of a zero job termination rate the reservation wage for the employed is equal to his/her current wage. As a result of a global job search, that is the search is not restricted by a local labour market, the propensity to migrate for the employed is lower than when he/she is unemployed. Also, their intensity of search must be different in favour of the unemployed due to different time resources for search. These two factors, reservation wage and intensity of search, affect the duration of search which should be shorter for the unemployed. Interestingly, theoretically the option of an on the job search decreases the reservation wage for an unemployed person (Combes *et al.* 2008). This further raises their probability of migration.

It is likely that three states, non-participation, job-seeking, and employment, could be endogenous to the overall (local and global) labour market and personal conditions, and to government assistance to the unemployed. For some unemployed it could be rational to search neither locally nor globally before migration and either not to search or search either locally or globally after migration (Molho 2001). Other unemployed may find that it is beneficial to search prior to migration. A non-participant will be a participant after migration only if their reservation wage in the host region exceeds the unemployment benefits plus other components of opportunity costs.

In order to understand the difference in migration behaviour of an employed person and NILF person assume for the latter that wage is never in their objective function. Then the agent’s problem is not a job search but rather a destination choice, in which consumption and personal motives prevail. As will be discussed in the theoretical Chapter 4, destination choice is different from job
search in one important aspect. It can be not random in nature and does not depend on the (job offer) arrival rate and (job) destruction rate. In algebraic form, the destination choice is a simple discrete maximization problem:

\[
U(x) \rightarrow \max \text{ s.t. } p(x) \cdot z(x) + cm(x) \leq I
\]

where each destination \( x \) from an exogenous set \( X \) is presented by a deterministic vector of characteristics \( z(x) = (z_1(x), \ldots, z_n(x)) \), utility is a function of the vector of these characteristics, \( U(x) = U(z_1(x), \ldots, z_n(x)) \), and \( cm(x) \) is the cost of migration to \( x \). \( I \) is a budget constraint, and \( p(x) \cdot z(x) \) is a scalar product of two vectors, the vector of prices \( p(x) \) and the vector of consumption \( z(x) \). Some items in the consumption vector are good, that is with a positive marginal utility, e.g. consumer goods and services, ocean coast, and warm climate. Others are bad, with a negative marginal utility, such as air pollution, density of population, and remoteness from larger cities.

In contrast to the destination choice, job search is a two parameter search problem in which the pair \((x, y)\), consisting of a destination \( x \) and wage offer \( y \), is stochastic. In both problems the set of destinations \( X \) is the same and all characteristics of destinations in \( X \) are known.¹ In the job search problem, the optimal solution is well known: to continue to search an extra period as long as it is worthwhile (see e.g. Lippman and McCall 1976). The optimal is a sequential search with the following stopping rule:

---

¹ In general to collect information about characteristics across all destinations in some set is costly. This immediately implies that for job search either the set of destinations is smaller or fewer characteristics are collected, or both.
**Stopping rule.** Accept the pair \((x, y)\) if \(y \geq R(x)\), where \(R(x)\) is the reservation wage at destination \(x\).

The reservation wage \(R(x)\) depends on the distribution of offers in \(x, F_x(\cdot)\) which are c.d.f. of independently distributed random variables, costs of living in \(x, p(x)\), costs of migration to \(x, cm(x)\), and constant costs of continuing search one extra period, \(c\), and determined by the equation

\[
\int_{R(x)}^{\infty} (1-F_x(z))\,dz = c - \Delta u
\]

where the new item \(\Delta u\), a monetary equivalent to a change in utility after migration from the source area to the destination \(x\), is added to the traditional equation for the reservation wage.

This reservation wage is an increasing function of the search costs, costs of living in \(x\), costs of migration (and the distance from the current location to the destination \(x\)), and the mean and variance of wage distribution. The optimal destination, i.e. the solution to the destination choice problem (2), is likely to be in close proximity to the source location and/or have lower costs of living but the job search process, described by the stopping rule and condition (3), selects a destination which is less likely to have this property. Instead, it is likely to have a higher average wage and/or wage inequality than the optimal destination. As a result, a destination from which the offer is accepted in the second problem is generally not the optimal destination. Therefore, the theory predicts that the unemployed and NILF individual from the same origin can behave differently, e.g. migrate to different destinations. Moreover, the NILF can be more mobile than the unemployed because they are not constrained by the job search. Lastly, the unemployed are more mobile than the employed due to a lower reservation wage as shown above.
In the empirical literature, the unemployed are repeatedly found to have a higher probability of migration than the employed due to lower opportunity costs. Herzog et al. (1993) reviewed ten microdata studies, eight on geographical mobility within the US and other two for the UK and Netherlands, and found unemployment status significantly augments migration likelihood in nine studies. For individuals NILF the effect is also known to be significantly higher than for the employed (Eliasson et al. 2003).

Another economic factor, income, is recognized in the theoretical literature as having a negative effect on mobility. This is evident in the human capital approach, the job search approach, and the traditional macro level gravity model of migration. The gravity model predicts larger flows of migrants between any two areas when income at the origin is lower or income at the destination is higher. Therefore, the propensity to move decreases with income. Alternatively, the liquidity constraint hypothesis implies a lower probability of migration among lower income labour due to inability to finance a move (e.g. Andrienko and Guriev 2004). As a consequence of this if financial constraints are nonbinding, the negative income impact is reversed. In support of this hypothesis, two (unpublished) studies on micro level data from the US National Longitudinal Survey of Youth (NLSY) have found extremely large fixed costs of migration. Bishop (2007) estimates that the cost of a one-mile move by a twenty-year old individual would be $336,000 in US dollars (year 2000 value). The corresponding figure from another paper (Kennan and Walker 2008) is even higher at $363,000. The authors show, however, that the present value of the income gain for some movers can exceed these costs. These estimated costs are very high. Other studies have found lower costs of migration. Davies et al. (2001) report that costs of migration are from $170,000 to $240,000 in four US states which they arbitrarily chose for their study. In a recent study of interstate migration in the US migration costs were found to be much lower, about $18,500, less than one-half of the average annual household income (Bayer and Juessen 2008).
In addition, there is the effect of housing, including homeownership and duration of home tenure. Both factors are theoretically impediments to migration if they represent costs and accumulated local, not easily transmitted, physical and social capital. Home owners are less mobile than those who rent a house because of the costs associated with the sale of the old home and the purchase of the new home. Duration of home occupancy reflects the accumulated social capital and stock of information in the current location with a negative impact on out-migration. Migration requires expenditures on the search for shops, schools, doctors, parks, restaurants, and other places people regularly visit. Homeownership is found to be detrimental to mobility in the empirical literature (e.g. Lansing and Mueller 1967, Goss and Paul 1986, a survey of more recent empirical evidence Dietz and Haurin 2003).

Another barrier to geographical mobility, distance, first appeared in the *Laws of Migration* (Ravenstein 1885). Distance was found from the macro level gravity model to be, perhaps, the most serious and inevitable impediment to mobility, though with gradually diminishing importance in the course of time. Distance between two populated areas decreases their mutual migration flows because costs of information search and migration increase with distance and also because intervening opportunities, a smaller distance from the origin, are potentially attractive destinations (Schwartz 1973, Greenwood 1997). Empirical evidence received in studies employing various methods suggests that intervening opportunities are important determinants in spatial interactions (Sheppard 1979, Fotheringham 1981, Raphael 1998). Results can be interpreted in the following way: that for a given distance maximum migration flows between two areas can be observed only if there are no other destinations within this distance and that every additional location diverts some part of the flows.

In theory, geographical remoteness may have either a decreasing or increasing effect on out-migration depending on the dominating force. Keeping income in both the origin and the
destination constant, the negative effect of remoteness on migration is due to the costly distance which causes higher costs of migration. The positive effect of remoteness is, perhaps surprisingly, also due to the costly distance if a consumer has already been paying for distance after a previous migration. In contrast to the traditional role of distance there is a force which may counteract the “tyranny of distance” and reverse the effect on migration. This force is accessibility/proximity to goods and services demanded by the consumer. If monetary and time costs of isolation measured by the costs of “consumption” in broad terms (which may include shopping, going to school, receiving various services, say, visiting a family doctor or lawyer, attending places of social activity) approximated by travelled distance, are relatively high then it could be rational to reduce the costs by moving house. In that case, assuming budget constraints are non-binding, geographical remoteness increases out-migration.

Another negative influence of distance is a feeling of loneliness due to living far from parents, children, and friends. This may result in higher psychic costs and a larger number of circular trips which are required to compensate for these feelings. Out-migration eliminates a part of these costs. The positive impact of remoteness, apart from the abovementioned reason, could also be explained by a wage premium shown to be linked to the distance from the closest labour markets. Recently the literature on New Economic Geography (NEG) has considered the hierarchical structure of cities and towns in which they are sorted according to population size. In the US the distance discount for wages is 5-9 percent and for housing costs it is 12-17 percent on average depending on the area’s proximity to the closest cities from each of four levels of hierarchy (Partridge et al. 2009). On the other hand, there could be a counter-balancing force, such as a preference for lower density population or lower rent areas, other things being constant, and therefore, one may observe a tendency for de-urbanization (see Table A1.3 in Appendix 1).

2 I use the phrase “tyranny of distance” to indicate the effect of geographical proximity to metropolitan areas, though the term is often applied to Australia’s geographical remoteness from the rest of the world.
Summarising the hypotheses proposed, this chapter seeks to show that

1. Propensity to migrate depends on individual labour force status and wage. While unemployment and nonparticipation may increase mobility, the impact of wage is ambiguous and depends not only on premium to migration but also on affordability of migration.

2. Propensity to migrate is lower with age and family size and higher with educational attainment as the human capital approach predicts.

3. Remoteness from larger cities has an impact on out-migration. It can be either negative, if distance represents mostly costs of migration, or positive if migration costs are compensated by eliminated costs of remoteness such as psychic costs and isolation costs.

In the empirical section the panel version of the probit model is applied in order to estimate the likelihood of the decision to migrate on an individual level as a function of labour force status and earnings, human capital, and other individual characteristics.

2.3 Empirical section

2.3.1 Data

In order to test the hypotheses, seven waves of the Household Income and Labour Dynamics in Australia (HILDA) survey were used to model migration decisions.\(^3\) All explanatory variables of the migration decision were taken from the wave preceding migration in order to exclude a potential

---

\(^3\) Limited information from wave 7 is derived only for construction of the dependent variable, namely the fact of migration and distance moved from the previous wave 6.
problem with endogeneity. A sample of working age adults was used with only a small proportion omitted due to missing observations thereby avoiding a sample selection problem.

A very important first step is the identification of migrants. Migration is defined as a residential address move a distance of at least 30 kilometres after a current wave and before the next wave. The HILDA field survey company has recently geocoded all addresses from detailed maps. Before that the survey linked each address to a geographical centroid of post code. Data report “as the crow flies” distance between the previous and current wave addresses if a person moved and assign zero distance otherwise. It seems to be quite likely that in most cases a move of over 30 kilometres causes individuals to change their labour market. However, since the commuting distance can be over 200 km to a largest city like Sydney (see Diagram A1.1 in Appendix 1), a mover of a distance of 400 kilometres theoretically can remain in the same labour market, though this is highly unlikely for someone from the HILDA survey. Such a small and arbitrary circle as 30 kilometres in diameter does not preclude the possibility that some proportion of migrants, especially in big cities, is commuting regularly to the previous location. Considering a larger minimum distance, one is more likely to exclude such commuters but at the same time can miss migrants who move a shorter distance.

In the literature the threshold used for the distance based measure of migration varies from 30 kilometres (Mitchell 2008) to 50 miles, that is, about 80 km. In a recent study Lemistre and Moreau (2009) considered a set of four threshold distances 0, 20, 50, and 100 km and demonstrated that wage gains generally increase with the threshold. It would be perhaps a good idea to apply a combined approach of a minimum distance and change of job. This approach is not used here because many even very long distance migrants said that they had the same employer in the interview prior to the migration.

---

4 Also see the discussion in the third chapter.
The purpose of the following exercise is to investigate how accurate is the 30 kilometre threshold for the definition of migration and to compare various types of mobility for migrants with the rest of the working age population. In order to do that, first, proportions of migrants and non-migrants who changed statistical division (which is a proxy for the labour market in this study), employer, and two-digit occupation were found. The sample for this exercise was restricted to 25,391 working age individuals who were employed at least two waves in a row. These three proportions are respectively 78, 55, and 41 percent for migrants at distances of at least 30 kilometres. The three numbers are 0.2, 12, and 14 percent respectively for non-migrants, that is stayers and movers of distances less than 30 km. In contrast to non-migrants, migrants are much more mobile geographically, between jobs and between occupations. Note that an increased minimum migration distance of up to 200 kilometres would cover almost all interregional migrants who changed statistical division, i.e. 98 percent, but not dramatically increase the proportion of migrants who change employer, only 64 percent. Although this implies that many long distance migrants continue to work for the same employer, only a small number of them, less than 2 percent, in the question about the reasons for move, mentioned the reason as “work transfer”.

Migrants are also more mobile in all three of these aspects in other periods. An extra exercise based on the same sample shows that in the past, current migrants were more mobile than current non-migrants; for the wave preceding the migration, the three proportions are 14, 26, and 24 for migrants as compared to 3, 14, and 15 percent for non-migrants. Finally, migrants are also more mobile than non-migrants in the consecutive wave, with nearly the same proportions observed for the wave preceding the migration for both groups (on a sample of 25,082 observations).

In this study, emigration from Australia is not included as migration because these migrants were not followed up. Moreover, the study does not distinguish between different types of internal migration: primary, circular, repeat, temporary, and permanent. As Greenwood (1997, p. 690)
warns, different types of migration are affected differently by various determinants of migration. Since it is impossible to determine the type of migration from the data, the study ignores this problem. However, the data do allow modelling of the effect of the migration experience to some extent, supplying variables for previous immobility (immobility dummy for those who never left home, however, a person could have moved frequently with parents) and recent moves (the number of homes lived in during the last ten years).

2.3.2 Data: Remoteness and geography

The only community level variable considered in the analysis is the Remoteness Aria. It is a purely geographical classification representing an area’s ARIA+ index (Area Remoteness Index Australia) calculated from the shortest road distances from a population place to the nearest city or service centre for each level of a five-level hierarchy of cities of different sizes. The idea of service centres was originally based on the accessibility of goods, services, and places of social interaction. All services are available in the largest cities while in small towns there are only limited services. The categorisation of a service centre depends only on the population size of the city. ARIA+ Index is constructed as a sum of five subindexes, representing (minimum) road distance to the closest city from a particular category. A subindex varies from 0 to 3 and ARIA+ is between 0 and 15. Category A, which includes major cities, is defined as urban centres with a population of over 250,000. Zero, the lowest level of remoteness, is assigned to this category, accommodating two thirds of the population in Australia according to Table 2. Category B includes urban centres with a population of between 48,000 and 250,000 and category C consists of cities with a population of between 18,000 and 48,000. Census Districts (CD), the smallest unit for which ARIA+ was constructed, with a score below 0.2 are located in close proximity to the major cities. No towns of a population above 25,000 have an ARIA+ score above 3.9 and only three towns with a population

---

5 Plus indicates that the original four-level ARIA is extended to the five-level hierarchy.
above 10,000 have a score exceeding that value (ABS 2001). The last two categories are towns with a population of between 5,000 and 18,000, category D, and those of between 1,000 and 5,000, category E. In the latter in general only a few services are available.

Remote and very remote areas have the highest ARIA+ scores but they represent only a minor proportion of the population, 1.8 and 1.1 percent respectively. It was found for ARIA, which is equivalent to ARIA+ but developed for a four-level hierarchy, that the average distance migrated declines with the size of the city: 413 km for category A, 239 km for B, 139 km for C, and 88 km for D (DHAC 2001). The corresponding value for E has not yet been ascertained by the author.

**Table 2. Classification of remoteness areas**

<table>
<thead>
<tr>
<th>Remoteness Area Class</th>
<th>Average ARIA+ Score Across CD’s</th>
<th>Proportion of Total Population (2001 Census), %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major Cities of Australia</td>
<td>0 to 0.2</td>
<td>65.92</td>
</tr>
<tr>
<td>Inner Regional Australia</td>
<td>0.2 to 2.4</td>
<td>20.63</td>
</tr>
<tr>
<td>Outer Regional Australia</td>
<td>2.4 to 5.92</td>
<td>10.54</td>
</tr>
<tr>
<td>Remote</td>
<td>5.92 to 10.53</td>
<td>1.78</td>
</tr>
<tr>
<td>Very Remote</td>
<td>10.53 to 15</td>
<td>1.07</td>
</tr>
<tr>
<td>Migratory *)</td>
<td>n/a</td>
<td>0.05</td>
</tr>
</tbody>
</table>

*) Migratory Areas composed of off-shore, shipping and migratory CDs


The following histograms show the distribution of the distances people move for each category of remoteness. Only long distance migrants are shown, and the minimum distance moved is set to be 30 kilometres. Altogether, the five diagrams reflect the gravity law, showing a decreasing number
of migrants with larger distances. As one would anticipate, relatively short distances of move prevail in larger cities and areas close to them: more than a half of the sample move distances less than 500 kilometres. In contrast, more than a half of migrants from remote and very remote areas move over 500 kilometres.

Finally, the matrix of transitions in Table 3, showing where all migrants over 30 kilometres from the pooled sample move to, gives an idea about the relative attractiveness of remoteness categories. Long distance migration shifts the category of remoteness in most cases; most notably this happens in remote (but not in very remote areas), at 93 percent of cases. Stable preferences are noted only for major city inhabitants: half of the long distance migrants from major cities are still living in big cities. In absolute values, major cities are the most attractive followed by inner regional and very remote areas. However, the observation from this matrix and the last five rows in Table 4 below regarding population distribution is that the majority of long distance migrants move mostly to a destination in the same category or one tier higher (one tier lower only for major cities). Remote and especially very remote areas seem to gain the most from internal migration but since there are not enough observations this hypothesis needs to be tested on Census data.
Diagram 1. Distribution of distance by remoteness area

Table 3. Transition matrix between remoteness areas for long distance migrants, percentage

<table>
<thead>
<tr>
<th></th>
<th>Major City</th>
<th>Inner Regional</th>
<th>Outer Regional</th>
<th>Remote</th>
<th>Very Remote</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major City</td>
<td>53</td>
<td>33</td>
<td>11</td>
<td>3</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>Inner Regional</td>
<td>43</td>
<td>36</td>
<td>18</td>
<td>2</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>Outer Regional</td>
<td>28</td>
<td>31</td>
<td>34</td>
<td>5</td>
<td>2</td>
<td>100</td>
</tr>
<tr>
<td>Remote</td>
<td>21</td>
<td>22</td>
<td>44</td>
<td>7</td>
<td>6</td>
<td>100</td>
</tr>
<tr>
<td>Very Remote</td>
<td>38</td>
<td>6</td>
<td>16</td>
<td>9</td>
<td>31</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>44</td>
<td>33</td>
<td>19</td>
<td>3</td>
<td>2</td>
<td>100</td>
</tr>
</tbody>
</table>
Both alternative definitions of migration, a change of SD or SLA, used below for a check of robustness in section 2.4, and “as the crow flies” distances, proposed in section 2.6, were developed from a classification in Australian Standard Geographical Classification (ASGC). There are four geographical classifications widely used by ABS in its official statistics. The primary classification is based on the following disaggregation level (the number of units according to ASGC 2001 is shown in parenthesis): States and Territories (9), Statistical Divisions (66), Statistical Subdivisions (207), and Small Local Areas (1353).\(^6\) Fifty six Statistical Divisions (SD), presented in the HILDA survey, serve in this research as an approximation for the local labour markets. A SD is defined by ASGC as representing a large, general purpose, regional type geographic area. It represents relatively homogeneous regions characterised by identifiable social and economic links between the inhabitants and between the economic units within the region, under the unifying influence of one or more major towns or cities. A SD consists of one or more Statistical Subdivisions and covers, in aggregate, the whole of Australia without gaps or overlaps. In addition, SDs do not cross state or territory boundaries and are the largest statistical building blocks of states and territories (ABS 2006).

Assuming that a SD forms a circle, the “diameter” of each SD was found using data from the ABS. The smallest SD is Other territories (not presented in the HILDA sample), which has a diameter of 16 km, the largest (balance of NT) 1311 km, the mean 304 km, and the median 202 km. A median size area is a good approximation of one labour market if borders of the labour market are defined so as to allow the possibility for labour to commute.

\(^6\) There is another, the lowest level of disaggregation, Census Districts, total over 37,000, used by ABS only in Census. From Census 2012 there will be even further disaggregation on mesh blocks, a new micro-level geographical classification developed by ABS.
Table 4. Summary statistics and comparison migrants vs. non-migrants*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>Non-migrant</th>
<th>Migrant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Migration: distance moved between current and next wave is over 30 km</td>
<td>0.047</td>
<td>0.211</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Age left home: Age when first moved out of home as a young person</td>
<td>18.48</td>
<td>6.06</td>
<td>0</td>
<td>60</td>
<td>18.53</td>
<td>17.32</td>
</tr>
<tr>
<td>Still living at home</td>
<td>0.065</td>
<td>0.246</td>
<td>0</td>
<td>1</td>
<td>0.065</td>
<td>0.074</td>
</tr>
<tr>
<td><strong>Home ownership:</strong> Own/pay mortgage</td>
<td>0.705</td>
<td>0.456</td>
<td>0</td>
<td>1</td>
<td>0.718</td>
<td>0.429</td>
</tr>
<tr>
<td>Rent or pay board</td>
<td>0.273</td>
<td>0.446</td>
<td>0</td>
<td>1</td>
<td>0.261</td>
<td>0.523</td>
</tr>
<tr>
<td>Live rent free</td>
<td>0.022</td>
<td>0.146</td>
<td>0</td>
<td>1</td>
<td>0.021</td>
<td>0.048</td>
</tr>
<tr>
<td>Home tenure: Years at current address</td>
<td>7.74</td>
<td>8.69</td>
<td>0</td>
<td>80.35</td>
<td>7.95</td>
<td>3.54</td>
</tr>
<tr>
<td><strong>Employment status:</strong> Employed FT</td>
<td>0.565</td>
<td>0.496</td>
<td>0</td>
<td>1</td>
<td>0.566</td>
<td>0.542</td>
</tr>
<tr>
<td>Employed PT</td>
<td>0.208</td>
<td>0.406</td>
<td>0</td>
<td>1</td>
<td>0.209</td>
<td>0.185</td>
</tr>
<tr>
<td>Unemployed, looking for FT work</td>
<td>0.028</td>
<td>0.164</td>
<td>0</td>
<td>1</td>
<td>0.026</td>
<td>0.053</td>
</tr>
<tr>
<td>Unemployed, looking for PT work</td>
<td>0.008</td>
<td>0.090</td>
<td>0</td>
<td>1</td>
<td>0.008</td>
<td>0.014</td>
</tr>
<tr>
<td>NILF, marginally</td>
<td>0.059</td>
<td>0.235</td>
<td>0</td>
<td>1</td>
<td>0.058</td>
<td>0.080</td>
</tr>
<tr>
<td>NILF, not marginally</td>
<td>0.133</td>
<td>0.339</td>
<td>0</td>
<td>1</td>
<td>0.133</td>
<td>0.127</td>
</tr>
<tr>
<td>Tenure in current/latest occupation (years)</td>
<td>7.93</td>
<td>9.64</td>
<td>0</td>
<td>55</td>
<td>8.06</td>
<td>5.17</td>
</tr>
<tr>
<td>Tenure with current employer (years)</td>
<td>5.36</td>
<td>7.62</td>
<td>0</td>
<td>52</td>
<td>5.47</td>
<td>2.92</td>
</tr>
<tr>
<td>Hours of work: Hours per week usually worked in all jobs</td>
<td>30.28</td>
<td>21.12</td>
<td>0</td>
<td>130</td>
<td>30.36</td>
<td>28.52</td>
</tr>
<tr>
<td>Wage per annum: Last financial year gross wages &amp; salary, includes</td>
<td>7.29</td>
<td>4.70</td>
<td>0</td>
<td>13.43</td>
<td>7.29</td>
<td>7.37</td>
</tr>
<tr>
<td>estimated from net, log</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dummy for non-positive wage per annum</td>
<td>0.287</td>
<td>0.452</td>
<td>0</td>
<td>1</td>
<td>0.288</td>
<td>0.270</td>
</tr>
<tr>
<td>Time spent in jobs in last financial year, %</td>
<td>77.06</td>
<td>39.38</td>
<td>0</td>
<td>100</td>
<td>77.22</td>
<td>73.87</td>
</tr>
<tr>
<td>Category</td>
<td>Minimum</td>
<td>Mean</td>
<td>Median</td>
<td>Maximum</td>
<td>Minimum</td>
<td>Mean</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>---------</td>
<td>------</td>
<td>--------</td>
<td>---------</td>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>Time spent unemployed in last financial year</td>
<td>4.17</td>
<td>16.91</td>
<td>0</td>
<td>100</td>
<td>4.05</td>
<td>6.65</td>
</tr>
<tr>
<td>Time NILF in last financial year</td>
<td>18.76</td>
<td>36.99</td>
<td>0</td>
<td>100</td>
<td>18.73</td>
<td>19.47</td>
</tr>
<tr>
<td>Time in FT education in last financial year</td>
<td>2.40</td>
<td>13.23</td>
<td>0</td>
<td>100</td>
<td>2.27</td>
<td>5.02</td>
</tr>
<tr>
<td>Time in PT education in last financial year</td>
<td>6.80</td>
<td>22.59</td>
<td>0</td>
<td>100</td>
<td>6.73</td>
<td>8.19</td>
</tr>
<tr>
<td>Years of education</td>
<td>12.02</td>
<td>2.42</td>
<td>9</td>
<td>17</td>
<td>12.01</td>
<td>12.05</td>
</tr>
<tr>
<td>Age</td>
<td>40.23</td>
<td>11.84</td>
<td>15</td>
<td>64</td>
<td>40.51</td>
<td>34.46</td>
</tr>
<tr>
<td>Sex</td>
<td>0.496</td>
<td>0.500</td>
<td>0</td>
<td>1</td>
<td>0.496</td>
<td>0.496</td>
</tr>
<tr>
<td>Marital status: married legally or de facto</td>
<td>0.697</td>
<td>0.460</td>
<td>0</td>
<td>1</td>
<td>0.702</td>
<td>0.583</td>
</tr>
<tr>
<td>Number of persons below 15 years of age at June 30</td>
<td>0.799</td>
<td>1.120</td>
<td>0</td>
<td>9</td>
<td>0.803</td>
<td>0.708</td>
</tr>
<tr>
<td>Variable</td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>Min</td>
<td>Max</td>
<td>Non-migrant</td>
<td>Migrant</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>-------</td>
<td>-----------</td>
<td>-----</td>
<td>-----</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td><strong>Occupation</strong>: Managers</td>
<td>0.127</td>
<td>0.333</td>
<td>0</td>
<td>1</td>
<td>0.128</td>
<td>0.110</td>
</tr>
<tr>
<td>Professionals</td>
<td>0.209</td>
<td>0.406</td>
<td>0</td>
<td>1</td>
<td>0.210</td>
<td>0.192</td>
</tr>
<tr>
<td>Technicians and Trades Workers</td>
<td>0.129</td>
<td>0.335</td>
<td>0</td>
<td>1</td>
<td>0.129</td>
<td>0.123</td>
</tr>
<tr>
<td>Community and Personal Service Work</td>
<td>0.093</td>
<td>0.291</td>
<td>0</td>
<td>1</td>
<td>0.092</td>
<td>0.110</td>
</tr>
<tr>
<td>Clerical and Administrative Workers</td>
<td>0.157</td>
<td>0.364</td>
<td>0</td>
<td>1</td>
<td>0.158</td>
<td>0.131</td>
</tr>
<tr>
<td>Sales Workers</td>
<td>0.080</td>
<td>0.271</td>
<td>0</td>
<td>1</td>
<td>0.079</td>
<td>0.099</td>
</tr>
<tr>
<td>Machinery Operators and Drivers</td>
<td>0.064</td>
<td>0.245</td>
<td>0</td>
<td>1</td>
<td>0.064</td>
<td>0.057</td>
</tr>
<tr>
<td>Labourers</td>
<td>0.130</td>
<td>0.336</td>
<td>0</td>
<td>1</td>
<td>0.128</td>
<td>0.166</td>
</tr>
<tr>
<td>Unknown and not asked</td>
<td>0.012</td>
<td>0.111</td>
<td>0</td>
<td>1</td>
<td>0.012</td>
<td>0.013</td>
</tr>
<tr>
<td><strong>Industry</strong>: Agriculture, Forestry and Fishing</td>
<td>0.042</td>
<td>0.201</td>
<td>0</td>
<td>1</td>
<td>0.042</td>
<td>0.045</td>
</tr>
<tr>
<td>Mining</td>
<td>0.013</td>
<td>0.115</td>
<td>0</td>
<td>1</td>
<td>0.013</td>
<td>0.016</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>0.102</td>
<td>0.302</td>
<td>0</td>
<td>1</td>
<td>0.102</td>
<td>0.085</td>
</tr>
<tr>
<td>Electricity, Gas, Water and Waste Services</td>
<td>0.009</td>
<td>0.095</td>
<td>0</td>
<td>1</td>
<td>0.009</td>
<td>0.008</td>
</tr>
<tr>
<td>Construction</td>
<td>0.066</td>
<td>0.248</td>
<td>0</td>
<td>1</td>
<td>0.066</td>
<td>0.061</td>
</tr>
<tr>
<td>Wholesale Trade</td>
<td>0.032</td>
<td>0.177</td>
<td>0</td>
<td>1</td>
<td>0.033</td>
<td>0.029</td>
</tr>
<tr>
<td>Retail Trade</td>
<td>0.086</td>
<td>0.281</td>
<td>0</td>
<td>1</td>
<td>0.085</td>
<td>0.101</td>
</tr>
<tr>
<td>Accommodation and Food Services</td>
<td>0.055</td>
<td>0.227</td>
<td>0</td>
<td>1</td>
<td>0.053</td>
<td>0.090</td>
</tr>
<tr>
<td>Transport, Postal and Warehousing</td>
<td>0.047</td>
<td>0.211</td>
<td>0</td>
<td>1</td>
<td>0.047</td>
<td>0.040</td>
</tr>
<tr>
<td>Information Media and Telecommunication</td>
<td>0.024</td>
<td>0.154</td>
<td>0</td>
<td>1</td>
<td>0.024</td>
<td>0.029</td>
</tr>
<tr>
<td>Financial and Insurance Services</td>
<td>0.033</td>
<td>0.179</td>
<td>0</td>
<td>1</td>
<td>0.033</td>
<td>0.027</td>
</tr>
<tr>
<td>Rental, Hiring and Real Estate Services</td>
<td>0.012</td>
<td>0.110</td>
<td>0</td>
<td>1</td>
<td>0.012</td>
<td>0.008</td>
</tr>
<tr>
<td>Professional, Scientific and Technological</td>
<td>0.067</td>
<td>0.251</td>
<td>0</td>
<td>1</td>
<td>0.068</td>
<td>0.065</td>
</tr>
<tr>
<td>Category</td>
<td>Wave 1</td>
<td>Wave 2</td>
<td>Wave 3</td>
<td>Wave 4</td>
<td>Wave 5</td>
<td>Wave 6</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>Administrative and Support Service</td>
<td>0.030</td>
<td>0.060</td>
<td>0.091</td>
<td>0.109</td>
<td>0.017</td>
<td>0.041</td>
</tr>
<tr>
<td>Public Administration and Safety</td>
<td>0.171</td>
<td>0.237</td>
<td>0.288</td>
<td>0.312</td>
<td>0.130</td>
<td>0.198</td>
</tr>
<tr>
<td>Education and Training</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Care and Social Assistance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arts and Recreation Services</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Services</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown and not asked</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remoteness area: Major City</td>
<td>0.619</td>
<td>0.242</td>
<td>0.117</td>
<td>0.019</td>
<td>0.004</td>
<td>0.177</td>
</tr>
<tr>
<td>Inner Regional</td>
<td>0.486</td>
<td>0.428</td>
<td>0.321</td>
<td>0.136</td>
<td>0.064</td>
<td>0.381</td>
</tr>
<tr>
<td>Outer Regional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remote</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very Remote</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave: 1</td>
<td>0.177</td>
<td>0.166</td>
<td>0.164</td>
<td>0.162</td>
<td>0.166</td>
<td>0.166</td>
</tr>
<tr>
<td>Wave 2</td>
<td>0.166</td>
<td>0.166</td>
<td>0.164</td>
<td>0.162</td>
<td>0.166</td>
<td>0.166</td>
</tr>
<tr>
<td>Wave 3</td>
<td>0.164</td>
<td>0.166</td>
<td>0.164</td>
<td>0.162</td>
<td>0.166</td>
<td>0.166</td>
</tr>
<tr>
<td>Wave 4</td>
<td>0.162</td>
<td>0.166</td>
<td>0.164</td>
<td>0.162</td>
<td>0.166</td>
<td>0.166</td>
</tr>
<tr>
<td>Wave 5</td>
<td>0.166</td>
<td>0.166</td>
<td>0.164</td>
<td>0.162</td>
<td>0.166</td>
<td>0.166</td>
</tr>
<tr>
<td>Wave 6</td>
<td>0.166</td>
<td>0.166</td>
<td>0.164</td>
<td>0.162</td>
<td>0.166</td>
<td>0.166</td>
</tr>
</tbody>
</table>

*) Number of observations is 51,111 for each variable
2.3.1 Descriptive statistics

Summary statistics are shown in Table 4 constructed on a sample of 51,111 observations used in the empirical analysis in the following sections (note longitudinal weights are not used in this Chapter). From the entire sample of around 70 thousand observations of working age individuals who are not full time students, all 16 thousand observations from Wave 7 were excluded. This was done in order to avoid endogeneity problems because the model uses individual characteristics from a previous survey, that is in the beginning of the period in which migration can occur. In the constructed sample migrants are almost evenly distributed between waves with some variation from 4.2 percent in 2004 to 5.3 percent in 2007, on average 4.6 for 6 waves of observations, as one can see from Table 5.

Table 6. Proportion of migrants, by wave, percentage

<table>
<thead>
<tr>
<th>Wave</th>
<th>Migrants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.43</td>
</tr>
<tr>
<td>2</td>
<td>4.62</td>
</tr>
<tr>
<td>3</td>
<td>4.19</td>
</tr>
<tr>
<td>4</td>
<td>4.83</td>
</tr>
<tr>
<td>5</td>
<td>4.53</td>
</tr>
<tr>
<td>6</td>
<td>5.33</td>
</tr>
<tr>
<td>Total</td>
<td>4.65</td>
</tr>
</tbody>
</table>

The association between theoretically important factors of migration and propensity to migrate long distance is tabulated below. The row Total in Table 6 shows that the proportion of people who
migrate a distance of over 30 km declines with age. However, the general observation in the literature that the risk of migration increases with the educational level is not supported. Propensity to migrate rises with education only in 20-24 years of age group.

Table 7. Propensities to migrate, by age and education, percentage

<table>
<thead>
<tr>
<th>Age group</th>
<th>Year11-, Certificate I-II</th>
<th>Year12, Dip, Certificate III-IV</th>
<th>Bachelor+</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-19</td>
<td>10.47</td>
<td>10.46</td>
<td>n.a. 1)</td>
<td>10.56</td>
</tr>
<tr>
<td>20-24</td>
<td>8.10</td>
<td>9.98</td>
<td>12.24</td>
<td>9.25</td>
</tr>
<tr>
<td>25-29</td>
<td>8.43</td>
<td>5.63</td>
<td>7.34</td>
<td>7.57</td>
</tr>
<tr>
<td>30-34</td>
<td>6.26</td>
<td>5.72</td>
<td>7.13</td>
<td>6.38</td>
</tr>
<tr>
<td>35-44</td>
<td>3.57</td>
<td>3.66</td>
<td>3.30</td>
<td>3.53</td>
</tr>
<tr>
<td>45-64</td>
<td>2.55</td>
<td>2.80</td>
<td>2.63</td>
<td>2.66</td>
</tr>
<tr>
<td>Total</td>
<td>4.84</td>
<td>4.38</td>
<td>4.66</td>
<td>4.65</td>
</tr>
</tbody>
</table>

1) 3 observations, in other cells the number of observations varies from 686 to 7,960.

Another interesting fact is that for each educational group there is an extended period with a stable migration propensity lasting about 10 years, with somewhat lower than a maximum propensity observed around completion of education but still at a high level. The same observation can be made from Table A1.2 in Appendix 1 calculated from the USA 1985 Census by Greenwood (1997).

The highest probability of migration between the current wave and the subsequent wave is for the unemployed. This is another common observation in the migration literature. Table 7 shows that unemployed persons, who are looking for a part time or full time job, have double the mobility of

---

7 The current wave here and thereafter implies any particular wave.
those employed being 8-9 percent versus 4 percent respectively. Somewhat surprisingly, the status of NILF is connected with some moderate mobility, with a propensity to migrate of 4.5 and 6.4 percent for those marginally and not marginally attached respectively. This does not seem to be a low mobility rate if one takes into account that people in this group are on average older than the unemployed and especially older than the employed. Among nonparticipants, the marginally attached are those NILF who currently are not looking for job but can start work.

Table 8. Propensities to migrate, by broadly defined employment status, percentage

<table>
<thead>
<tr>
<th>Employment status</th>
<th>Migrants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employed FT</td>
<td>4.46</td>
</tr>
<tr>
<td>Employed PT</td>
<td>4.13</td>
</tr>
<tr>
<td>Unemployed, looking for FT work</td>
<td>8.88</td>
</tr>
<tr>
<td>Unemployed, looking for PT work</td>
<td>7.66</td>
</tr>
<tr>
<td>NILF, marginally</td>
<td>6.35</td>
</tr>
<tr>
<td>NILF, not marginally</td>
<td>4.47</td>
</tr>
</tbody>
</table>

Another piece of evidence on the intermediate mobility of persons NILF is documented by Newbold (2001). The author constructed means on a sample of people aged five and over from the

---

8 “Marginal attachment to the labour force is determined by firstly establishing whether a person not in the labour force has a desire to work, and then by whether they have been actively seeking work or are available to start work within a short period of time. Persons who are marginally attached may satisfy some, but not all, of the criteria required to be classified as unemployed. Persons not in the labour force are considered to be marginally attached to the labour force if they: (i) want to work and are actively looking for work but not available to start work in the reference week; or (ii) want to work and are not actively looking for work but are available to start work within four weeks. Persons not in the labour force are not marginally attached if they: (i) do not want to work; or (ii) want to work but are not actively looking for work and are not available to start work within four weeks” (HILDA Data Dictionary 2010).
1996 Canadian Census, which showed that the employed have lower return migration but higher onward migration than those NILF. However, the sample used is too wide and includes those not of working age, who are probably less likely to have onward migration relatively to a return move.

Propensity to migrate may be affected by income prior to migration. An income quartile for each individual is constructed on their recorded gross wages and salaries for the last financial year. Those who did not have or simply did not report the annual wage, 29 percent of the sample according to Table 4, were assigned to the zero wage quartile. Looking at their annual wages reported in the previous interview, one can find that ¾ of individuals studied had zero wages, but the remainders are not equally divided between the remainder income quartiles, with a lower probability of transition from a higher quartile. The highest probability to move to the zero wage in the subsequent wave is found for the first income quartile, at 20 percent. The first income quartile has the highest propensity to move, at 6 percent, while the top quartile has the lowest, at 4 percent (see Table 8).

**Table 9. Propensities to migrate, by annual wage quartile, percentage**

<table>
<thead>
<tr>
<th>Wage quartile</th>
<th>Migrants</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 *)</td>
<td>4.37</td>
</tr>
<tr>
<td>1</td>
<td>5.92</td>
</tr>
<tr>
<td>2</td>
<td>4.85</td>
</tr>
<tr>
<td>3</td>
<td>4.26</td>
</tr>
<tr>
<td>4</td>
<td>4.05</td>
</tr>
</tbody>
</table>

*) zero quartile for those who do not report or report zero last financial year wages & salaries.

Those who did not report an annual wage have neither high nor low levels of mobility. Their mobility, 4.4 percent, approximately corresponds to the mobility of the median earner. Again one
should take numbers in this table with some caution since income quartiles are composed of different categories, such as younger, disabled, and currently unemployed persons, who tend to be in the lower quartiles.

Mobility on both long and short distances is traditionally closely linked to home ownership status. On the one hand home owners have a higher fixed costs of moving, on the other they are generally richer than non-owners. These two forces have opposite impact on mobility. The proportion of owners and mortgage holders who migrate is three times lower, only at 3 percent, than the proportion of migrating non-owners, see Table 9.

**Table 10. Propensities to migrate, by home ownership, percentage**

<table>
<thead>
<tr>
<th>Type of ownership</th>
<th>Migrants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own/currently pay mortgage</td>
<td>2.83</td>
</tr>
<tr>
<td>Rent or pay board</td>
<td>8.90</td>
</tr>
<tr>
<td>Live rent free</td>
<td>10.22</td>
</tr>
</tbody>
</table>

Finally, people who are living far from big cities may be more prone to migration due to restricted access to services. The remoteness area variable is used to construct migration probabilities. As documented in Table 10 the propensity to migrate is the lowest for the inhabitants of major cities, state or territory capital, and cities close to them, at the level of 3.5 percent. It gradually increases with distance from these cities, approaching 10 percent for remote areas and 15 percent for very remote areas.\(^9\) Remarkably, the lower long distance mobility in big cities is compensated by a

---

\(^9\) In the latter case calculations were based on 213 observations.
higher residential mobility within cities. This is apparently due to agglomeration, which is also true for smaller cities but not to such a large degree.

**Table 11. Propensities to migrate, by remoteness area, percentage**

<table>
<thead>
<tr>
<th>Remoteness area</th>
<th>Migrants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major City</td>
<td>3.52</td>
</tr>
<tr>
<td>Inner Regional</td>
<td>5.63</td>
</tr>
<tr>
<td>Outer Regional</td>
<td>7.34</td>
</tr>
<tr>
<td>Remote</td>
<td>10.34</td>
</tr>
<tr>
<td>Very Remote</td>
<td>15.02</td>
</tr>
</tbody>
</table>

In addition to the simple tabulation done in the tables above it was found that recent migrants are more likely to be repeat migrants. Among those who migrated long distances between the previous and current wave, the probability to move long distances before the next wave is 19 percent, which is fourfold of the average probability for other respondents. Data from other countries indicate that this number can be larger but any comparison is problematic mostly due to the different size of administrative units. A higher propensity to move is observed in HILDA not only one year past migration but for the following years as well. The probability is equal to 15 percent during the second year and 16 percent during the third year.

**2.3.2 Methodology of econometric analysis**

A number of models are traditionally used to estimate binary dependent variable models. This includes logit and probit models. Both models when estimated on cross-sectional data generally give close results. The probit model was selected because it is able to estimate a model on panel
data. The random effects probit model is preferred because probit model with the fixed effects gives inconsistent estimates on small samples such as that used in this chapter due to the incidental parameters problem. The presentation below follows the model setup from Greene (2002).

Denote $y_{it}^*$ as a latent variable for an individual $i$ gain from (optimal) migration at year $t$. Then a model for a latent gain has the following form:

$$y_{it}^* = \beta'x_{it} + \lambda_i + \theta_t + \epsilon_{it}; \ t = 1, ..., T; \ i = 1, ..., N$$

where $x_{it}$ is a vector of individual level regressors, $\lambda_i$ is the individual random effect, $\theta_t$ is the time effect, and $\epsilon_{it}$ is the error term. The panel is balanced for simplicity of presentation. However, in the following empirical section an unbalanced panel is used.

All independent variables are assumed to be strictly exogenous. This implies there is no correlation of any independent variable with any error term, i.e.:

$$\text{Cov}(x_{it}, u_{js}) = 0$$

where $u_{it} = \lambda_i + \theta_t + \epsilon_{it}$, across all individuals $i$ and $j$ and all periods $t$ and $s$, ruling out state persistence, i.e. the presence of the lagged dependent variable.

The latent variable is not observed. What is observed is the realization of a positive gain from migration. Denote $m_{it}$ as a binary variable for migration decision of an individual $i$ between years $t$ and $t+1$:

$$m_{it} = \begin{cases} 0 \text{ if } y_{it}^* \leq 0 \\ 1 \text{ if } y_{it}^* > 0 \end{cases}; \ t = 1, ..., T; \ i = 1, ..., N$$
The probit model is based on the probability of migration having the form:

\[ P\{m_{it} = 1|x_{it}\} = \Phi(\beta'x_{it} + \lambda_i + \theta_i) \]

The vector of parameters \( \beta', \theta_i \) is estimated by maximum likelihood. Full maximum likelihood estimates of the parameters with the homoscedasticity assumption are obtained by maximizing the log likelihood function,

\[ \text{LogL} = \sum_{i=1}^{N} \log \Phi(a_{t1}, ..., a_{tT}|\Sigma') \]

with respect to the unknown elements in \( \beta \) and \( \Sigma \), where \( \Sigma \) is variance-covariance matrix with normalised diagonal elements,

\[ \sigma_u = 1, \quad a_u = (2y_u - 1)x_u'\beta, \quad \sigma^*_u = (2y_u - 1)(2y_u - 1)\sigma_u, \quad t \neq s, \]

and \( \Phi \) denotes the c.d.f. of the T-variate normal distribution. When \( T > 2 \) a computation of this function requires a multidimensional integration which is not readily done in formulas and should be approximated. In general the maximum likelihood estimation is computationally intensive unless one assumes that \( \lambda_i, \theta_i \), and \( \varepsilon \) are independent. Denoting \( \lambda_i \sim N(0, \sigma_{\lambda}), \theta_i \sim N(0, \sigma_{\theta}), \varepsilon \sim N(0, \sigma_{\varepsilon}) \), the correlation across \( i \) is constant, i.e.:

\[ \sigma_{\pi} = \text{Cov}(u_{it}, u_{is}) = \rho = \frac{\sigma^2_{\lambda}}{\sigma_{\lambda}^2 + \sigma_{\theta}^2 + \sigma_{\varepsilon}^2}, \quad t \neq s. \]

Then the estimation can be done by classical simulation methods such as the Gauss-Hermite quadrature method which is used here.\(^{10}\) The econometric software package STATA was used in

\(^{10}\) It should be said that it is quite a slow method. Computational time varied from less than a minute to a few hours depending on the number of regressors.
the following empirical analysis, in particular a command xtprobit for the estimation of the probit model with the random effects.

### 2.4 Analysis

The model of migration decisions to be estimated in this section has a specification:

\[ m_{it} = \beta'x_{it} + \lambda_{it} + \theta_{it} + \epsilon_{it} \]

Notations are similar to models (4) and (5). The dependent variable is a binary variable of the migration decision between the current wave and the subsequent wave of the HILDA survey. The random effects probit model described above is applied for the estimation of this equation.

A vector of individual level variables \( x_{it} \) includes, among others, aggregate level characteristics representing amenities (e.g. climate, terrain, and proximity to an ocean) and public goods, considered to be regional invariant variables and, therefore, modelled by the regional fixed effect. These amenity differentials represent the relative attractiveness of destinations and, as recognized in the economics of migration, can be capitalized in wages and house price differentials. A less than full capitalization does not prevent us from use of regional dummies in the analysis because they can still incorporate other differentials such as amenities/disamenities. Region is a state/territory in the primary model and a statistical division in a check of robustness.

An estimation strategy is to introduce independent variables gradually, one by one, in broad categories. Nine categories, from 1 to 9, which correspond to the number of columns in Table 11 with regression results, are added into the model one by one. They are arranged in the following groups: (1) human capital and demography: educational attainment, sex, age, marital status, and the number of children; (2) employment characteristics: labour force status, occupational/employer tenure; (3) work and unemployment experience; (4) home ownership and tenure (5) annual wage
(weekly and hourly in the robustness check); (6) migration experience: age left home and never left home, (7) remoteness area, (8) wave and region, and (9) occupation and industry (one-digit classifications in order to reduce computation time). Note, only in the very end were wave and regional dummies, and finally, occupation and industry categories added before the F-test of joint significance is done.

Table 11 presents the core results of the econometric model estimation. It contains not the estimated regression coefficients commonly met in linear regression analysis but the marginal effects constructed from these regression coefficients under the assumption of zero random effect. The built-in procedure mfx in STATA was applied after the regression estimation. For a continuous variable the marginal effect is the derivative of the probability of migration with respect to this variable. For a binary variable it is a change of probability when the binary variable changes the value from 0 to 1. Assuming zero random effect, the probability of migration is equal to 2.0 percent.\textsuperscript{11} The following text discusses results obtained for each category of the independent variables.

(i) Demography

In the model specification with a linear effect of age, age is found to have a declining effect on mobility (not reported in Table 11). However, generally the migration decision is modelled as a non-linear function of age representing a higher propensity to move around retirement. In the specification of all nine models this non-linearity is modelled by means of a quadratic function. Both parameters for the linear and quadratic terms are found to be highly significant with the negative and positive sign respectively. Coefficients allow us to estimate an age at which there will be minimum effect on mobility. The first model in column 1 implies that the minimum occurs

\textsuperscript{11} It is quite an arbitrary assumption if one keeps in mind that those prone to migration can be self-selected and, therefore, have a large random effect
outside the working age, at \( \frac{100 \times 0.229}{2 \times 0.159} = 72 \) years. The models in columns 3 to 8, however, unanimously find that the minimum mobility is at age 50 as shown in Diagram 2 based on the parabola coefficients from column 9.

**Diagram 2. The effect of age on the propensity to migrate, %**

One can note that there is no difference in the migration decision between males and females as indicated by the insignificant coefficient for sex in all models. Being male has a positive effect on mobility only in the most comprehensive model, but only on 10 percent confidence level.

Married (legally or de facto) respondents are found to be less mobile than unmarried by 0.2 percent but this result is at a higher level of significance only in the simplest specifications. As was pointed out in the introduction, family migration decisions (and outcomes) of migration are different from individual decisions due to the presence of tied movers, generally wives. Marital status may have a negative effect for males but not for females. This hypothesis is not tested here, however. The
model estimated with dummies for six marital status categories: married, de facto married, separated, divorced, widowed, and never married or de facto married, has found significant
<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
<th>Model 7</th>
<th>Model 8</th>
<th>Model 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of education</td>
<td>0.032</td>
<td>0.067***</td>
<td>0.065**</td>
<td>0.067**</td>
<td>0.057*</td>
<td>0.092***</td>
<td>0.140***</td>
<td>0.140***</td>
<td>0.118***</td>
</tr>
<tr>
<td>Age</td>
<td>-0.229***</td>
<td>-0.170***</td>
<td>-0.168***</td>
<td>-0.238***</td>
<td>-0.247***</td>
<td>-0.245***</td>
<td>-0.229***</td>
<td>-0.225***</td>
<td>-0.231***</td>
</tr>
<tr>
<td>Age^2/100</td>
<td>0.158***</td>
<td>0.102**</td>
<td>0.102**</td>
<td>0.240***</td>
<td>0.251***</td>
<td>0.250***</td>
<td>0.230***</td>
<td>0.224***</td>
<td>0.231***</td>
</tr>
<tr>
<td>Sex</td>
<td>0.044</td>
<td>0.213</td>
<td>0.195</td>
<td>0.152</td>
<td>0.130</td>
<td>0.216</td>
<td>0.181</td>
<td>0.187</td>
<td>0.245*</td>
</tr>
<tr>
<td>Marital status</td>
<td>-0.421***</td>
<td>-0.319**</td>
<td>-0.321**</td>
<td>-0.190</td>
<td>-0.201</td>
<td>-0.210</td>
<td>-0.240*</td>
<td>-0.246*</td>
<td>-0.245*</td>
</tr>
<tr>
<td>Number of children</td>
<td>-0.201***</td>
<td>-0.228***</td>
<td>-0.217***</td>
<td>-0.111*</td>
<td>-0.104*</td>
<td>-0.124**</td>
<td>-0.160***</td>
<td>-0.166***</td>
<td>-0.158***</td>
</tr>
<tr>
<td>Employment status: Employed FT</td>
<td></td>
<td>base</td>
<td></td>
<td>base</td>
<td></td>
<td>base</td>
<td></td>
<td>base</td>
<td>Base</td>
</tr>
<tr>
<td>Employed PT</td>
<td>-0.102</td>
<td>-0.097</td>
<td>0.069</td>
<td>0.161</td>
<td>0.136</td>
<td>0.118</td>
<td>0.140</td>
<td>0.094</td>
<td></td>
</tr>
<tr>
<td>Unemployed, looking for FT work</td>
<td>2.291***</td>
<td>2.199***</td>
<td>1.897***</td>
<td>1.992***</td>
<td>1.862***</td>
<td>1.888***</td>
<td>1.892***</td>
<td>1.652***</td>
<td></td>
</tr>
<tr>
<td>Unemployed, looking for PT work</td>
<td>2.103**</td>
<td>2.208**</td>
<td>2.083**</td>
<td>2.229**</td>
<td>2.056**</td>
<td>2.239**</td>
<td>2.273**</td>
<td>2.025**</td>
<td></td>
</tr>
<tr>
<td>NILF, marginally</td>
<td>1.980***</td>
<td>2.340***</td>
<td>2.118***</td>
<td>2.235***</td>
<td>2.108***</td>
<td>2.156***</td>
<td>2.196***</td>
<td>2.019***</td>
<td></td>
</tr>
<tr>
<td>NILF, not marginally</td>
<td>1.450***</td>
<td>1.893***</td>
<td>1.819***</td>
<td>1.913***</td>
<td>1.864***</td>
<td>1.879***</td>
<td>1.906***</td>
<td>1.865***</td>
<td></td>
</tr>
<tr>
<td>Tenure in occupation /10</td>
<td>-0.153</td>
<td>-0.174</td>
<td>-0.075</td>
<td>-0.091</td>
<td>-0.052</td>
<td>-0.037</td>
<td>-0.047</td>
<td>0.039</td>
<td></td>
</tr>
<tr>
<td>Tenure in occupation ^2/100</td>
<td>0.024</td>
<td>0.028</td>
<td>0.009</td>
<td>0.013</td>
<td>-0.0003</td>
<td>-0.013</td>
<td>-0.010</td>
<td>-0.020</td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>Column 1</td>
<td>Column 2</td>
<td>Column 3</td>
<td>Column 4</td>
<td>Column 5</td>
<td>Column 6</td>
<td>Column 7</td>
<td>Column 8</td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>Tenure in occupation not observed</td>
<td>-1.230</td>
<td>-1.210</td>
<td>-1.264</td>
<td>-1.348</td>
<td>-1.209</td>
<td>-0.791</td>
<td>-0.918</td>
<td>-0.711</td>
<td></td>
</tr>
<tr>
<td>Tenure with employer /10</td>
<td>-1.321***</td>
<td>-1.369***</td>
<td>-0.825***</td>
<td>-0.854***</td>
<td>-0.798***</td>
<td>-0.863***</td>
<td>-0.828***</td>
<td>-0.950***</td>
<td></td>
</tr>
<tr>
<td>Tenure with employer ^2/100</td>
<td>0.232**</td>
<td>0.245**</td>
<td>0.169</td>
<td>0.176*</td>
<td>0.164</td>
<td>0.174*</td>
<td>0.165</td>
<td>0.188*</td>
<td></td>
</tr>
<tr>
<td>Tenure with employer not observed</td>
<td>0.226</td>
<td>0.333</td>
<td>0.599</td>
<td>0.725</td>
<td>0.540</td>
<td>-0.086</td>
<td>0.088</td>
<td>0.124</td>
<td></td>
</tr>
<tr>
<td>Time unemployed</td>
<td>0.044</td>
<td>-0.249</td>
<td>-0.050</td>
<td>-0.064</td>
<td>-0.056</td>
<td>-0.050</td>
<td>-0.088</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time NILF</td>
<td>-0.561**</td>
<td>-0.679**</td>
<td>-0.523*</td>
<td>-0.517*</td>
<td>-0.453</td>
<td>-0.454</td>
<td>-0.377</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time in FT education</td>
<td>0.279</td>
<td>0.719**</td>
<td>0.787**</td>
<td>0.808**</td>
<td>0.891**</td>
<td>0.940***</td>
<td>0.899**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time in PT education</td>
<td>-0.179</td>
<td>-0.122</td>
<td>-0.119</td>
<td>-0.099</td>
<td>-0.093</td>
<td>-0.062</td>
<td>-0.061</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home ownership: Own/mortgage</td>
<td>base</td>
<td>base</td>
<td>base</td>
<td>base</td>
<td>base</td>
<td>base</td>
<td>Base</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rent/pay board</td>
<td>1.993***</td>
<td>2.000***</td>
<td>1.842***</td>
<td>1.972***</td>
<td>1.948***</td>
<td>1.930***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Live rent free</td>
<td>3.440***</td>
<td>3.477***</td>
<td>3.294***</td>
<td>2.38***</td>
<td>2.347***</td>
<td>2.338***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>Model</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>Home tenure /10</td>
<td></td>
<td>-1.738***</td>
<td>-1.725***</td>
<td>-1.658***</td>
<td>-1.612***</td>
<td>-1.572***</td>
<td>-1.541***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home tenure ^2/100</td>
<td></td>
<td>0.210***</td>
<td>0.208***</td>
<td>0.207***</td>
<td>0.204***</td>
<td>0.200***</td>
<td>0.195***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage per annum</td>
<td></td>
<td>0.147*</td>
<td>0.139*</td>
<td>0.216***</td>
<td>0.205***</td>
<td>0.206***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage per annum not observed</td>
<td></td>
<td>1.461</td>
<td>1.390</td>
<td>2.238**</td>
<td>2.106*</td>
<td>2.242**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age left home</td>
<td></td>
<td>-0.185***</td>
<td>-0.148***</td>
<td>-0.143***</td>
<td>-0.140***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Still living at home</td>
<td></td>
<td>-1.971***</td>
<td>-1.754***</td>
<td>-1.727***</td>
<td>-1.681***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remoteness area: Major City</td>
<td></td>
<td>base</td>
<td>base</td>
<td>base</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inner Regional</td>
<td></td>
<td>1.451***</td>
<td>1.450***</td>
<td>1.431***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outer Regional</td>
<td></td>
<td>2.707***</td>
<td>2.660***</td>
<td>2.629***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remote</td>
<td></td>
<td>5.288***</td>
<td>4.899***</td>
<td>4.939***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very Remote</td>
<td></td>
<td>7.583***</td>
<td>8.407***</td>
<td>8.208***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave 1</td>
<td></td>
<td>base</td>
<td>base</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.045</td>
<td>0.051</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>-0.300*</td>
<td>-0.292*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-0.014</td>
<td>-0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.204</td>
<td>-0.185</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.180</td>
<td>0.198</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

State/territory: NSW

| Vic | -0.444*** | -0.435*** |
| Qld | 0.176 | 0.186 |
| SA | -0.615*** | -0.595*** |
| WA | -0.182 | -0.192 |
| Tas | -0.782*** | -0.777*** |
| NT | 0.486 | 0.314 |
| CTA | 0.244 | -0.021 |

*) No. of observations is 51063. No. of respondents is 12261. Occupation and industry dummies are included only in model 9.
If the married compose a reference group, then the widowed do not have statistical
different mobility but other groups are more mobile. Surprisingly, never married/de facto married
have a lower propensity to migrate than the de facto married. The former have a marginal effect on
the propensity equal to 0.6 percent while for the latter it is 1.1 percent. The most mobile groups are
the divorced and separated, with a higher risk of migration by 1.4 and 2.0 percent respectively.

Finally, the number of children with age below 15 is a significantly negative variable in all
regressions. The marginal effect of an extra child is estimated to be about -0.2 percent. This implies
that if a family has two children then for a married working age adult from this household the
probability of migration is lower by 0.6 percent (from column 9 the total effect is
\(-0.245 - 2 \times 0.168 = 0.581\)) than in the case this adult is unmarried and has no children.

(ii) Education

In the economic literature human capital was found to have a positive impact on geographical
mobility. In this study the number of years of education is used as a measure of human capital. It is
assumed that the risk of migration depends linearly on this number. Depending on the specification
of the model, the marginal effect of education is generally significant. It is small and insignificant in
the simplest specification with six regressors. It is relatively large and highly significant in the more
comprehensive models. Each extra year of education increases the probability of move by 0.14
percent. Thus, the difference in mobility between a year 12 school graduate and a university
graduate with bachelor degree will be \(3 \times 0.140 = 0.420\) percent.

Further, perhaps even more interesting results are obtained when dummies for nine educational
categories are substituted for the number of years of education. All persons with 9 or fewer years of
schooling are taken as a base group. A hypothesis that all the other eight dummies included in the

\(\text{\textsuperscript{12}}\) This and other checks of robustness with categorical dummies are not shown in Table 11.
regression are statistically zero is rejected on 1 percent level by means of F-test, because 
\[ \chi^2(8) = 25.46, \text{Prob}>\chi^2 = 0.0013. \] Results show that while year 12 graduates are not different in migration propensity from the base group, the higher levels of educational attainment are found to have a rising probability of migration, statistically different from that of the base group. Certificate level III or IV, the most common forms of vocational training in Australia, have the risk of migration higher by 0.34 percent than year 9 and below. Those who have received a diploma or advanced diploma, are 0.55 percent more mobile. The highest mobility is observed among the highest levels of educational attainment. The holders of a graduate diploma or graduate certificate are 0.93 percent more mobile than year 9 graduates, master or doctorate are also more mobile by 0.94 percent, as well as bachelor or honours, by 0.96 percent. These results can be interpreted as a sort of educational ladder. The higher is the stair, the higher the mobility and very likely not only geographical mobility.

Those who recently completed education are hypothetically at a higher risk of migration. Two variables identifying recent students were studied in addition to the educational attainment variable: proportion of time spend in full-time (FT) education and in part-time (PT) education in the last financial year. More time in FT education is found to have a statistically positive influence on migration in the short-term while time in PT education is not significant. The marginal effect is large, e.g. 100 percent time spend in FT education last year raises the migration propensity by 0.9 percent as compared with no FT education.

(iii) Employment

Labour force status. The labour force status is broadly defined with classification on six categories: FT and PT employed, unemployed looking for FT and PT jobs, marginally NILF and not
marginally attached to the labour market. Recall from the construction of the sample that current FT students are excluded from the analysis.

The estimation results are unanimous across all estimated models. There are four categories of not employed: unemployed irrespective of whether they are looking for FT or PT job and persons NILF marginally and not marginally have a higher propensity to move than the employed. The marginal effects are estimated 2.0 percent for the unemployed looking for PT job and those marginally NILF. Increase in the probability to migrate for those not marginally out of the labour force is 1.9 percent and for unemployed eager to find FT job is 1.7 percent. F-test on equal coefficients for all four categories is not rejected.

These results are at odds with Table 7. Mean propensities to migrate in Table 7 clearly say that the unemployed looking for FT jobs are the most mobile and that those not marginally attached to the labour force are among the least mobile as well as the FT employed. However, neither observation is confirmed after demographic and other individual variables were controlled in the analysis.

Tenure. Two measures of tenure were probed. The first is tenure in occupation and the second is tenure with employer. For people currently not working, tenure in occupation and with employer is taken to be equal to the most recent observation, when they were employed. After this retrospective projection there is still a large proportion of respondents with either one or both undefined tenures. In this case the unobserved tenure is set to be zero. In addition, two dummy variables, “tenure in occupation is not observed” and “tenure with employer is not observed”, are introduced into the model. When either of the two measures of tenure is considered in the model, tenure is negative and its square is positive and both estimated coefficients are significant (not reported in Table 11). However, when both tenures are studied together, only tenure with employer and squared is statistically different from zero, though the quadratic term is barely significant. Those who do not
have tenure are statistically not different from the remainders as the two insignificant dummies point out.

Therefore, the tenure-migration relationship is U-shaped in the same way as the age-migration pair. This conclusion is demonstrated by Diagram 3 constructed on column 9 coefficients. According to the graph, minimum mobility is observed around 25 years of tenure with employer and maximum with nil tenure and then only over 50 years, which is, perhaps, a very rare event.

**Diagram 3. The effect of tenure with employer on the propensity to migrate, %**

*Duration of unemployment.* People who are the long-time unemployed or out of the labour force may have a different propensity to migrate than the short-term unemployed and employed. Recent unemployment and NILF spells were approximated by the proportion of time unemployed and time NILF, respectively, in the last financial year. These two variables, though not perfect measures of the two spells, are mostly insignificant in the regression results. Nevertheless, a recent long NILF
spell is somewhat significantly negative in half of the models, though these are the less comprehensive models.

(iv) Home ownership and tenure

*Home ownership.* Home owners are known to be less mobile than non-owners. This is a widely acknowledged finding which is also confirmed in the analysis here. Three home ownership categories: (1) own or pay mortgage; (2) rent; and (3) live rent free, were modelled in the form of dummy variables. The first category was excluded from the regression. The other two are always found to be statistically positive. The marginal effect of renting and living rent free on the probability to migrate is found to be at 1.9 and 2.3 percent respectively.

*Home tenure.* Another variable which might negatively affect outmigration is home tenure. Similar to the variables for age and tenure, the number of years lived at the current address and its quadratic value were included in the list of regressors. Both variables are highly significant with the conventional U-shaped form. This relationship is also demonstrated in Diagram 4 below. Minimum mobility is found for tenure about 40 years.

(v) *Wage.* A study of income effect is often complicated by the presence of missing observations. Not all of the respondents report weekly and even annual wages. Those who are unemployed or out of the labour force in general do not have wages but can still migrate for different reasons including job related ones. The exclusion of these two non-random categories from the analysis of migration decisions would lead to a sample selection bias in the estimates. By keeping all respondents in the sample and considering the binary variable for non-earners in addition to the wage variable, this problem was overcome.
Three different variables representing wages were explored in the model, one at a time. Table 11 reports the results of only one of them, wages and salaries for the last financial year. They show that both the annual wage and the dummy for zero annual wage have a positive effect on migration, at 1 and 5 percent level of statistical significance. Wage has a relatively small effect on mobility: a 10 percent increase in the annual wage leads only to 0.02 percent higher chances of out-migration. If wage for the previous financial year is absent, the marginal effect on the propensity to migrate is large at 2.2 percent. The discontinuity of the income effect found in zero is a puzzle. Theoretically, non-earners should not be different from lower income earners unless there are unobserved characteristics not controlled in the model which are highly correlated with the non-earner status (e.g. total household income or wealth) but not so with wage.

The analysis performed with either of the alternative measures of wage, weekly and hourly, did confirm the same positive finding showing highly significant coefficients for both wage and non-earner dummy, at this time at 1 percent level (not reported in Table 11). Wage increases the
propensity to migrate but absent wage has a positive effect. A question about the size of the effect of unobserved wage can be addressed by calculation of the ratio of the coefficients for the binary variable and wage. This ratio gives us information about “equivalent income” for non-earners. Those who do not report the annual wage are mobile as if they have approximately 10\textsuperscript{th} percentile wage (from the distribution of positive annual wages), that is in essence not very mobile. However, the same exercise with either weekly or hourly wages led to a different conclusion: people who did not report wage for the previous week will migrate before the subsequent wave of the survey with approximately the same probability as a median wage earner.

As a check of robustness instead of the continuous annual wage variable, five binary variables were constructed for each of the four wage quartiles and in addition for non-earners, which is a reference group. Findings from the model with the four dummy variables (one is always excluded) do not diverge from the abovementioned results. The top wage quartile has a statistically larger mobility than non-earners while the other three wage quartiles have statistically the same propensity to move as non-earners.

(vi) Mobility experience. Mobility experience has a positive effect on future geographical mobility due to lower costs of repeat migration. Past migrants accumulate knowledge and skills in job and house search, and in the process of adjustment to a new area, which includes search for schools, shops, services, recreational facilities, new friends, and other places.

Household surveys rarely contain data on migration history. In such a survey respondents are asked about all locations and dates they lived there. Unfortunately, there is no such historical section in the HILDA yet.

In this study the age a person left home for the first time as a young person serves as a proxy for migration experience. The earlier a person left home, the more migration experience is
accumulated, *ceteris paribus*. In addition to this continuous variable a dummy for those who have never left home is also explored. There could be some shortcomings in these two proxies for migration experience. Say, if parents were mobile in a person’s childhood (which is less likely according to our results though) then this person has already accumulated some migration experience without leaving the parents’ home.

Data show that most young people leave home shortly after they finish school. Thus, the maximum probability of leaving home occurs at age 18 and is equal to 16 percent, and the total probability to leave home within the five year period from 17 to 22 years of age is 57 percent.

The regression analysis demonstrated that the proxy representing the migration experience is very influential. People who left home earlier are more likely to be migrants. The estimated marginal effect shows that an additional year lived with parents reduces the probability of migration by 0.1 percent, the same effect as from one year of education less. The dummy for those who never left home has the negative sign as anticipated. The marginal effect of this dummy is quite big, at -1.7 percent. This effect is equivalent to the period of 12 years longer stay with parents. However, one concern should be mentioned here. This dummy variable is unlikely to be a good proxy for immobility in young ages, especially shortly after the completion of school. If needed, this can be studied by another dummy variable or, in a more general case, by the interaction of the never left home dummy with age.

(vii) **Remoteness.** For each individual the HILDA reports a category of remoteness assigned according to ARIA+. Five categories of remoteness are studied. The *Major City* category is a base. The other four have statistically significant marginal effects. Thus, the above conclusion from Table 10 is supported. Respondents living in close proximity to major cities, *Inner Regional Australia*, are 1.4 percent more likely to migrate than major city inhabitants. In relatively more remote areas,
**Outer Regional Australia**, the propensity to move is higher by 2.6 percent. In the **Remote Australia** category population is even more mobile, with 4.9 percent increment, but **Very Remote Australia** has the highest mobility, the marginal effect is 8.2 percent. It is worthy of note that these differences are much smaller than the means calculated in Table 10, respectively 2.1, 3.8, 6.8, and 12.5 percent, which is apparently due to differences in terms of demographic and socio-economic factors across the categories of remoteness areas.

(viii) **Time and region.** The time effect is studied here because there could be macroeconomic shock common to all areas in the national economy, e.g. interest rate and inflation. The time effect is modelled by a binary variable for each wave, from one to six, in total six dummies. The regional effect represents the attractiveness of a region, which is assumed to be stable in time. In particular it may include amenities and disamenities of a region such as terrain, climate, and precipitation and some aggregate economic and other indicators, e.g. unemployment rate, crime rate, price level, and public goods provision. Dummies for the eight states and territories were chosen to represent the regional unobserved effect. The first wave and NSW are excluded from the model by a STATA programme as usual. A model, estimated with the time and regional dummies, demonstrated that previous results hold. A hypothesis about joint insignificance of time effects is not rejected. However, the regional effects are not jointly equal to zero according to F-test, since $\chi^2(7)=28.20$, Prob$>\chi^2=0.0002$. Three regions were found to have a significantly different out-migration risk from that in NSW. None of the seven regions has a statistically higher migration propensity than NSW. **Ceteris paribus**, Australians living in Victoria, South Australia, and Tasmania have a lower out-migration by 0.4, 0.6, and 0.8 percent respectively.

In a check of robustness migration was defined as a change of SD or SLA. In the sample there are 56 SDs out of 60 in Australia, where the four omitted being very small in population. SDs, perhaps, represent more accurately the boundaries of the local labour markets than States and Territories,
though in a few cases they are very large and exclude commuting between SDs. The mean diameter of an SD is about 260 km, if the three SDs with the largest areas in NT, SA, and WA are excluded; without the ten largest and not densely populated SDs the mean diameter is 197 km. These diameters are roughly in accordance with our understanding of labour market boundaries if commuting is considered. Diagram A1.1 in Appendix 1 shows a distribution of commuting distance in Sydney. It is probably fair to conclude from the graph that a miserable proportion of employed in Sydney, less than 1 percent, commutes further than 100 kilometres. The definition of migration, used in this chapter, is not unique, for example, Mitchell (2008) defines migration in two similar ways, a change of SLA and a move over 30 km.

The model was re-estimated with regional dummies for SD. Results of the estimated model do not change in many instances. In the light of the effect of remoteness found above it is not surprising that 30 SDs have a statistically higher probability of out-migration than Sydney SD while another 25 are only insignificantly different. However, in this regression three remoteness area categories lose their significance, only Very Remote is still significantly positive. This is not a surprising result because many SDs are relatively homogeneous in the sense of remoteness. The homogeneity is shown by the average standard deviation of remoteness areas (which in the data base have integer values between 0 and 4) within one SD, which is equal to 0.31. Thus, the double standard deviation is 0.62, which implies that in general a SD comprises remoteness areas mostly of the same category.

(ix) Occupation and industry. Finally, one-digit classifications of occupations and industries were used as extra controls in the regression analysis. For each person currently not employed the most recent occupation and industry were taken. If this search process in the data was unsuccessful the person was assigned to “unknown” industry or occupation. The total number of occupations and industries amounts to 10 and 20 respectively. The estimated coefficients for occupation and industry are placed in a separate Table A1.1 in Appendix 1, for other variables see Table 11, column 9. With
respect to these two categories results demonstrated that for 19 occupations coefficients were jointly equal to zero but for 9 industries they were not. None of the occupations, including not observed occupation, is significantly different from Managers, the base group. Among industries, when Agriculture, Forestry and Fishing are the reference group, only Public Administration and Safety and Arts and Recreation Services are significant with the marginal effects at 1.5 and 1.3 percent respectively. Only not observed industry leads to a lower migration risk, by 0.8 percent.

A comparison of columns 8 and 9 in Table 11 reveals that other variables change neither their significance nor remarkably the magnitude of their marginal effect.

### 2.5 Discussion

The analysis done in the preceding part of the chapter clearly demonstrated the presence of migration costs, the attachment to the current locality, home, and work, the importance of financial constraints and labour force status. This is not a surprising result in the economic literature, just additional evidence of the fruitfulness of an economic approach applied to the study of migration.

Several broad categories of factors which potentially can be either productive or counterproductive in migration decisions were studied. Current labour force status is found to be very important: unemployment drives people to search for a job elsewhere. What is surprising is that people NILF are no less active in migration which is further evidence that not only economic factors are considered in the migration decision. This requires an additional investigation, however, since NILF individuals can still be tied movers. This can be done by means of a few additional variables in the model, such as dummy for a tied person (say, the spouse of the household head), a dummy for a single earner household, and their interaction.

Financial constraints, indicated by the positive sign of individual income, seem to be a serious impediment to mobility. Higher income people, even with the same level of education, are more
mobile even after controlling unobserved characteristics, which are assumed to be uncorrelated with other variables. Somewhat surprisingly, people with zero annual wage are much more mobile than individuals with a very low wage. This conclusion, again, may require extra controls in the model in addition to the abovementioned three variables for tied persons, such as total household income or wealth, in order to reduce the effect of other family members. Although the proposed variables may shed light on the relationship between income, labour force status, and migration, there are some doubts that the main results derived from such an already comprehensive model can change. However, this type of analysis will be in another area of migration research. It is a family migration decision, a more complicated approach than from an individual perspective. This approach was not intended to be investigated in this relatively simple study.

The positive role of remoteness in out-migration found here presents a new interesting direction of migration research. The propensity to move longer distances is shown to increase with the degree of remoteness. It looks like physical distance or time of travel from larger urban centres has a positive effect on geographical mobility. The nature of this phenomenon should be further investigated. New economic geography has found a penalty of remoteness on wages and rents (Partridge et al. 2009). The distance penalty does not explain the phenomenon because wages are found to have a positive impact on migration. A question one may ask is whether it is the proximity to various services available only in larger cities which determines the positive effect of remoteness. The remoteness index was originally designed to measure such proximity. A guess is that it is due to proximity but only partially and there are more things here. Quality of services is also important. Say better teachers, doctors, and lawyers are more likely to live in larger cities. Presumably it is even more important that the remoteness implies higher psychic costs because parents, children, and friends are generally more likely to live in other areas. Higher psychic costs are especially more likely for people living in larger cities which may obviously deter them from leaving cities. Therefore, it is
the feeling of isolation, what Australians call the “tyranny of distance”, which pushes people from remote areas.

Age, marital status, and family size are all factors related to costs of migration which may prevent an individual from investment in migration. On the other hand, higher human capital and previous migration decisions facilitate mobility reducing uncertainty, and costs of search and migration.

Attachment to place, occupation, and employer, measured in this study by various tenure variables, presents a negative role of accumulated human capital and social network, and higher opportunity costs of migration to other locations where this capital may not be fully transferred.

2.6 Directions of future research

There are a number of variables that can be added to the model specification. However, they need some preliminary work. The list includes unemployment duration showing to be important by Goss and Paul (1984), commuting time (Eliasson et al. 2003), and some income variables other than individual wage such as annual family income, and income and employment status of the spouse (DaVanzo 1978).

What is intriguing in migration studies is the effect of (expected) wages across destinations. The task, how to measure these wages for each individual and then to add this measure into the analysis, is a challenge. A possible solution is to predict wages elsewhere using the earnings equation estimated on a panel data with individual fixed effects and regional dummies. SD, the number of which total 60, is probably the minimum possible aggregation level for which this prediction can be made because of the relatively small sample size of the HILDA survey. In a similar way, from the employment equation one can predict the risk of employment. Then, the expected wages across destinations can be estimated. Theoretically to realise migration, it should be higher at least in one destination than in the origin. The idea is to consider only SDs with a higher predicted wage for
each individual and to construct three variables, their count, average wage, and the average distance to them. Alternatively, instead of these three variables one can consider two vectors, with a count of SDs and the average wage in concentric circles of various diameters.

With respect to geographical and demographic characteristics attached to each locality, it would be desirable to add some structure to the model, finding the total population and/or the number of cities of different population size/hierarchical level inside concentric rings of various radiuses. Alternatively, it is possible to calculate “as the crow flies” distances from a respondent’s current CD to the closest urban centres for each category, similar to ARIA. Unfortunately, road distances used for ARIA calculation are not available. However, “as the crow flies” distances can be constructed relying on another ABS standard geographical classification based on Urban Centres and Localities. In this case the remoteness methodology will be very close to the RRMA remoteness classification which prevailed in Australian statistics in 80s and 90s before the superior classifications, ARIA and ARIA+, were developed.

A model of destination choice is another very interesting direction of research. This model is especially breathtaking because it allows a researcher to study not a simple binary choice, migrate or stay, as this chapter has done, but another type of decision, choice of one among a potentially very large set of alternatives. There is a variety of statistical models which can be applied to this model; all are based on the logit model: conditional logit model, nested logit, and mixed logit. The nested logit model with individual fixed effects, a method first described in Falaris (1987), would be the most desirable because it is suitable for panel data modelling. However, there is a property, *Irrelevance of Independent Alternatives (IIA)*, which imposes serious limitation on the model applicability. One can test whether this property holds.
It would not be a bad idea to use different samples. A sample of young people who have completed higher education and have just started their career is an example. Also, one could model the migration decision for primary migrants only. The sample can be restricted to those who have never left home and are not FT students. However, many of the primary migrants will be excluded from the sample because they might have moved to start tertiary education. This type of study would escape ambiguity of results arising from return migration. Note there will be either none or a maximum of one migration for each individual in this sample. Another idea, perhaps, is that a small sample can be constructed of families all of whose members are tracked, e.g. two parent households with all children living together, may be also with grandparents, in the beginning of the survey. In this sample each family member location would be observed every wave and therefore, one could model the migration decision of children as a function of distances from previous addresses, from parents, and from siblings. The hypothesis to be tested is whether psychic costs, measured by distances to family members and to home, are important in migration decisions. Previous evidence from research in Denmark suggests that they are (Dahl and Sorenson 2010).

Another attractive direction is to introduce a spatial dimension since there is a spatial component in the migration decision. This extension has not been widely elaborated yet, surprisingly. That geocoded data has only been made available recently is a primary reason. Another no less important reason, there are no good methods of analysis and no good models for spatial data due to the complexity of the phenomenon. A number of theories which might be helpful can be mentioned. Spatial job search theory: under some conditions one searches globally (otherwise locally) and accepts an offer above some threshold which depends on the region’s characteristics, e.g. distance from current region, wage distribution, costs of living, and others (Maier 1995). New economic geography might be useful here, especially its idea of market potential for a region, equal to the sum of the ratios of a regional GDP over the distance from the base region (Combes et al. 2008). It
implies both capital and labour benefit and therefore both are expected to move to, and be concentrated in, regions with the largest market potentials.

Finally, Tobler’s first law of geography “Everything is related to everything else, but near things are more related than distant things” (Tobler 1970). Spatial econometrics explores a spatial lag very similar to the market potential notion applying it to other variables. It is the average characteristic across nearby regions or, in the general case, the weighted average variable across all regions with some distance decay. A widely used weight is the negative exponential function of distance between a pair of regions (Fingleton 2009). A research question is whether the migration decision depends negatively on the market potential of a sending region which is a proxy for wage minus costs associated with migration.

The final suggestion: instead of the binary dependent variable in the migration decision one may study continuous dependent variable. A panel version of the Tobit model applied to migration distance seems to be very suitable. Distance in the HILDA is a left censored variable, since according to the definition of migration here it is zero in 96 percent of cases. Under a different, broader definition of movers as people moving house, the proportion is still very large, at 82 percent. Preliminary results for this model have provided estimates consistent with the results for the model of migration decisions and the economic theory stated in this study.

2.7 Conclusions

Economic theory provides a rationale behind migration decisions. The role of costs and its association with human capital, previous migration experience, and demographic variables such as age, marital status, and family size has been widely studied in the literature. Aggregate level studies, still popular in the literature due to the vast availability of macro data, have demonstrated the importance of income and employment differentials for direction of migration flows.
This chapter contributes to the literature by considering micro level data and modelling migration decisions on an individual level. The binary dependent variable model of the migration decision was estimated on the panel data from the HILDA survey by means of a probit model with individual random effects.

The results confirmed previous findings about the role of demographic variables. They demonstrated the U-shaped effect of age, the positive impact of educational attainment and the negative of marital status and number of children. New results are found with respect to economic variables. The unemployed, traditionally viewed as the most mobile, are found to be not alone. Those who are not in the labour force are no less mobile. Here the explanation could be sociological rather than economic.

In contrast to aggregate level studies results, the positive effect of income on migration, found in this micro level study, implies that financial constraints are important in migration decisions. The discontinuous pattern of this impact was found in zero income. People without wages are much more likely to migrate than those with low wages. Various tenure measures were studied and the negative effect of attachment to employer and home is confirmed.

Finally, an interesting phenomenon was identified with respect to the measure of remoteness from larger urban centres. Geographical mobility is strongly determined by remoteness. People are more likely to out-migrate from more remote areas. This phenomenon is at odds with the traditional view that distance deters migration. It was suggested in this chapter that not only proximity to services of wider range and better quality but also the elimination of psychic costs of living far from relatives and friends is a possible explanation of the phenomenon. This conjecture needs to be empirically verified.
### Appendix 1

**Table A1.1. (Table 12 continued). Probit model results, marginal effects in percent**

<table>
<thead>
<tr>
<th>Occupation: Managers</th>
<th>base</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professionals</td>
<td>-0.294</td>
</tr>
<tr>
<td>Technicians and Trades Workers</td>
<td>-0.325</td>
</tr>
<tr>
<td>Community and Personal Service Work</td>
<td>-0.205</td>
</tr>
<tr>
<td>Clerical and Administrative Workers</td>
<td>-0.238</td>
</tr>
<tr>
<td>Sales Workers</td>
<td>0.255</td>
</tr>
<tr>
<td>Machinery Operators and Drivers</td>
<td>-0.112</td>
</tr>
<tr>
<td>Labourers</td>
<td>-0.002</td>
</tr>
<tr>
<td>Unknown and not asked</td>
<td>-0.516</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Industry: Agriculture, Forestry and Fishing</th>
<th>base</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mining</td>
<td>0.127</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>-0.009</td>
</tr>
<tr>
<td>Electricity, Gas, Water and Waste Services</td>
<td>0.313</td>
</tr>
<tr>
<td>Construction</td>
<td>0.126</td>
</tr>
<tr>
<td>Wholesale Trade</td>
<td>0.041</td>
</tr>
<tr>
<td>Retail Trade</td>
<td>0.015</td>
</tr>
<tr>
<td>Accommodation and Food Services</td>
<td>0.393</td>
</tr>
<tr>
<td>Transport, Postal and Warehousing</td>
<td>0.175</td>
</tr>
<tr>
<td>Information Media and Telecommunication</td>
<td>0.930</td>
</tr>
</tbody>
</table>
Table A1.2. Propensities to migrate interstate in the US, 1980-1985, by age and education, percentage

<table>
<thead>
<tr>
<th>Education</th>
<th>Agea</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18–24</td>
</tr>
<tr>
<td>Elementary</td>
<td></td>
</tr>
<tr>
<td>0–8 years</td>
<td>8.21</td>
</tr>
<tr>
<td>High school</td>
<td></td>
</tr>
<tr>
<td>1–3 years</td>
<td>9.33</td>
</tr>
<tr>
<td>4 years</td>
<td>11.31</td>
</tr>
<tr>
<td>College</td>
<td></td>
</tr>
<tr>
<td>1–3 years</td>
<td>10.12</td>
</tr>
<tr>
<td>4 years</td>
<td>24.13</td>
</tr>
<tr>
<td>5 years or more</td>
<td>29.04</td>
</tr>
</tbody>
</table>

*The base population is the relevant number of nonmovers over the 1980–85 period, plus out-migrants. Age is defined as of 1985.

Source: Greenwood (1997, p.656)
Table A1.3. Evidence from published studies on urbanization or counterurbanization trends in European countries during the 1970-1990s

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Czechoslovakia</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany (E)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Portugal</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Both processes</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Belgium</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Iceland</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Counterurbanization</td>
<td>Denmark</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Germany (W)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Greece</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Netherlands</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Sweden</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Switzerland</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>


Diagram A1.1. Total commuters by distance commuted, 2001 – SLAs in Illawarra, Hunter, Sydney, and Inner Sydney SRs

3 Who benefits from internal migration

3.1 Introduction

The impact of geographical mobility on income in Australia is of interest due to the unique geographical structure of the country. Australia comprises a small number of large cities and very few second-tier cities. These cities are located unevenly across a continent comparable in size with the main continental land mass of the United States but with a population which is only 7 percent of the USA population size. Since costs of migration are not directly observed the gains from internal migration cannot be measured accurately. However, the returns to geographical mobility can be approximated by the wage premium from migration. Calculated as the difference in an individual’s wage in the destination post-migration and their origin immediately pre-migration, this wage premium will be greater than costs of migrating for an economically rational migrant.

Table 12 gives an idea about the potential gains for intra- and inter-regional internal migrants in Australia. The wage premium for living in capital cities is on average 6 percent (compared to the national average) while the wage discount for living outside the capital cities is 11 percent. Adjusting for demographical factors and differences in housing rents reduces these two figures albeit marginally, to 5 and 9 percent respectively. In spite of the different demographic, skills, and educational compositions of the population across the six states, income inequality, calculated for disposable income in the form of Gini coefficient, does not vary remarkably, lying between 0.29 and 0.32. It is anticipated that the average premium for internal migration in the same occupation is less than the premiums shown in Table 12 and that for the skilled occupations wage premium is

---

13 The words geographical mobility and migration are used interchangeably in this chapter.
14 Australia is a federation of 6 states and 2 territories. The states are New South Wales (NSW), Victoria (Vic), Queensland (Qld), South Australia (SA), Western Australia (WA), Tasmania (Tas); the territories are the Northern Territory (NT) and the Australian Capital Territory (ACT).
larger than for the unskilled but these hypotheses are not of primary interest in this chapter and remain to be studied.

**Table 13. Income and income inequality, states and territories, 1997-98**

<table>
<thead>
<tr>
<th></th>
<th>NSW</th>
<th>Vic</th>
<th>Qld</th>
<th>SA</th>
<th>WA</th>
<th>Tas</th>
<th>NT</th>
<th>ACT</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative disposable income</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capital city</td>
<td>110.3</td>
<td>104.6</td>
<td>110.5</td>
<td>91.8</td>
<td>101.9</td>
<td>89.4</td>
<td>119.4</td>
<td>120.2</td>
<td>106.1</td>
</tr>
<tr>
<td>Balance of State</td>
<td>86.1</td>
<td>88.0</td>
<td>92.8</td>
<td>69.6</td>
<td>108.9</td>
<td>86.1</td>
<td>n.a.</td>
<td>n.a.</td>
<td>89.0</td>
</tr>
<tr>
<td>Total</td>
<td>101.5</td>
<td>100.2</td>
<td>100.6</td>
<td>85.7</td>
<td>103.6</td>
<td>87.5</td>
<td>119.4</td>
<td>120.2</td>
<td>100.0</td>
</tr>
<tr>
<td>Gini coefficient</td>
<td>0.316</td>
<td>0.286</td>
<td>0.303</td>
<td>0.293</td>
<td>0.305</td>
<td>0.288</td>
<td>0.279</td>
<td>0.303</td>
<td></td>
</tr>
</tbody>
</table>

Relative Henderson equivalent income*

<table>
<thead>
<tr>
<th></th>
<th>NSW</th>
<th>Vic</th>
<th>Qld</th>
<th>SA</th>
<th>WA</th>
<th>Tas</th>
<th>NT</th>
<th>ACT</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital city</td>
<td>110.3</td>
<td>104.5</td>
<td>103.3</td>
<td>94</td>
<td>101.7</td>
<td>91.7</td>
<td>107.5</td>
<td>110.2</td>
<td>105.0</td>
</tr>
<tr>
<td>Balance of State</td>
<td>87.6</td>
<td>90</td>
<td>94.6</td>
<td>76.2</td>
<td>107.6</td>
<td>88</td>
<td>n.a.</td>
<td>n.a.</td>
<td>90.8</td>
</tr>
<tr>
<td>Total</td>
<td>102.1</td>
<td>100.6</td>
<td>98.5</td>
<td>89.2</td>
<td>103.2</td>
<td>89.6</td>
<td>107.5</td>
<td>110.2</td>
<td>100.0</td>
</tr>
</tbody>
</table>

*The Henderson equivalence scale is a uniquely Australian measure based on household disposable income after housing costs, and adjusted by household size.


Table 12 indicates that migration between some of the capital cities presents substantial pecuniary gains. The two largest potential wage premiums are the twenty percent for those moving from Hobart (capital of Tasmania) to Sydney (NSW) and the sixteen percent for those moving from Adelaide (SA) to Sydney. These wage premiums may not be sufficiently high to motivate many workers to move a distance in excess of one thousand kilometres. Away from the capital cities there
are even higher potential wage gains. For example, a person moving the equivalent distance of Adelaide to Sydney but moving from elsewhere in South Australia to Sydney could potential increase their income by a third.\(^\text{15}\) An interesting question which arises from the Table is, ‘to what extent is it distance that prevents people from such speculative moves?’.

This question is of interest as distance is known to be a deterrent factor which reduces the incidence of migration between two locations (Greenwood 1997). Income is expected to be increasing with distance people move. This is due to the trade-off between distance moved by a migrant in a destination and the wage s/he requires in order to move to that destination. Conditional on migration, generally speaking, wage should include a reward for every additional kilometre moved farther away from home.

A job search theoretical framework such as that in, for example Cahuc and Zylberberg (2004), lends itself to this analysis. First, returns to distance are positive because the costs associated with a move are increasing with distance. Costs include not only transport expenditures and the opportunity costs of time spent on a move but other costs such as search for information about potential destinations, psychic costs of missing family and friends, communication costs such as telephone calls, and others.\(^\text{16}\) A job offer is accepted only if the wage post-migration at the destination exceeds the reservation wage at least by the costs associated with moving. The reservation wage is measured by the wage pre-migration at the origin or, in general, by a wage, any local job offer above which is accepted. The further an individual moves the greater the costs s/he incurs and therefore s/he will demand a higher income (costs hypothesis).

\(^{15}\) Balance of State is the aggregation of all Statistical Divisions within a state or territory other than its Capital City Statistical Divisions.

\(^{16}\) Following an idea of Schwartz (1973) psychic costs are proportional to the transportation costs times to the number of visits to an old location. Since psychic costs represent multiple visits, they could be much larger than costs of a single move.
Second, there are more opportunities to improve wage with the greater distance. This idea is known as the distance opportunity hypothesis, as proposed by Sjaastad (1962). If it is assumed that the variance of wages does not change with the radius of search then it can be concluded that the reservation wage is increasing with the radius of search. By extending the radius of the job search a person sequentially revises their set of searching destinations (and therefore, the distribution of the net wages). In this process, promising distant labour markets are included and inferior ones excluded. As a result of this optimization process the optimal stopping rule of search determines the reservation wage, above which any offer (minus the costs of migration) is accepted. The reservation wage is higher with radius due to the higher mean of wage distribution (minus the costs of migration) and due to the larger job offer arrival rate. It is noteworthy that the distance elasticity of wage is not only positive but can potentially be an increasing function of distance. This is due to the spatial distribution of cities since the number of destinations within a circle is roughly a quadratic function of the radius.

Third, spatial job search outcomes could be opposite for migrants with different initial conditions. Lower paid workers, *ceteris paribus*, seem to benefit more from job search and ultimately from both migration and distance than higher income earners because their current wage is more likely to be below their reservation wage. The higher income, *ceteris paribus*, brings about a lower wage premium which could be even negative since the initial income is likely to exceed the reservation wage. This income effect hypothesis implies the positive effect of distance on wage is decreasing with initial wage and may disappear after some threshold level of initial wage is reached.

Fourth, there is better ability to collect and evaluate information about potential destinations by the better educated, and this potentially brings about higher returns. As Greenwood states “education increases the benefits of migration while it decreases the costs by improving information about
alternative destinations and decreasing the risk associated with movement over greater distance” (Greenwood 1997, p. 673).

In summary, these four hypotheses suggest that the returns to internal migration are expected to be a positive function of migration distance and education but with a decreasing rate for higher income workers. It is the empirical testing of this hypothesis that is the focus of the remainder of this chapter.

This chapter is organized into the following sections. First, the economic literature on migration is reviewed with an emphasis on the returns to migration. Second, a dynamic earnings equation is estimated as a function of distance and other explanatory variables using longitudinal data from the sole Australian panel data set, the HILDA survey. The elasticity of wage with respect to distance is estimated by means of a system GMM estimator controlling not only for observed but also unobserved individual and regional characteristics. In the last part of the chapter some concluding remarks are offered.

### 3.2 Literature

Traditionally economists model migration as an investment decision of a rational agent (Becker 1962, Sjaastad 1962). Migration of an individual occurs if the income gain exceeds the costs of move is in the life-time income maximization problem. This model is based on the disequilibrium theoretical perspective since there are opportunities for utility gain that can be arbitraged by migration. On the contrary, there are no such opportunities in the equilibrium perspective developed later in the economic literature such as the equilibrium model developed in Greenwood (1997). This approach proposed the idea that wages and prices are in equilibrium at any point in time and observed variations in space are just compensating differentials for unequal amenities, public goods, quality of life, and environmental quality. Accordingly migration results in a lower income only if
there is compensation in the form of income stability and certainty and/or higher level of amenities and quality of life.

The gravity model, traditionally used in macro level analysis of migration, international trade and transportation, found distance to be impediment to migration and trade, e.g. Andrienko and Guriev (2004). The explanation of the negative effect of distance on migration comes from the relation between information costs, psychic costs and distance because (i) psychic costs of moving are a positive function of distance to a destination as far as they could be offset by more frequent trips back to the source region, (ii) available information is diminishing with distance, and (iii) costs of information search are an increasing function of distance (Schwartz 1973, Greenwood 1997).

Empirical evidence about returns to migration is relatively scarce. Until recently migration studies were mostly based on aggregate level data for developed countries. A lack of detailed statistics led to the development of intensive migration research at the macro level. The effect of mobility on income both at the micro and macro levels is not unambiguous. One of the common findings obtained from the estimation of Mincer-type earning equations on micro level data in the US, Canada, and other countries until recently demonstrated a short-term negative effect of migration on earnings (Greenwood 1997). This was shown not only for internal migrants but also for international migrants. For example, data from the HILDA supports this finding for international migrants (Doiron and Guttmann 2009). There are methodological difficulties in the approach used in this type of study since the migrants are usually compared with the reference group at the destination but not with any similar group in the source region or country. On the other hand, there are positive findings for international migration, e.g. the NIS survey in the US demonstrated that new legal immigrants have better education than natives and that there are economic gains from migration for most of them (Jasso et al. 2002).
Recent work based on micro level panel data show a positive wage premium for internal migration for men (in the USA, Yankow 1999 and 2003; in Britain, Bőheim and Taylor 2007). While Yankow (2003) finds increasing in time effect of migration on returns with maximum 11 percent wage premium after five plus post-migration years, Bőheim and Taylor observe the short-run positive effect for two-year period following migration with peak 3 percent at year of migration reducing to zero in three years though they use a shorter panel which does not allow to study longer effect.

The propensity of migrants to move farther away if they can get extra pecuniary benefits, the so-called income distance trade-off, was quantitatively estimated on aggregate data. The trade-off is measured by the ratio of distance elasticity of migration flows to income elasticity of migration flows. Both elasticities are obtained from a modified gravity model. This trade-off value in Canada varied from 3.5 in 1952 to 1.5 in 1967, implying that in time migrants are willing to travel farther for the same growth in income (Courchene 1970). Also, Courchene found that the trade-off value for better educated (i.e. high-school graduates or above) 25-34 years of age persons is 4.4, much larger than the value of 2.9 for persons with elementary school education or below.

Magrini (2006) studied wage returns to spatial mobility for young French workers by estimating a Mincer earning equation on cross-sectional data. She found the income elasticity of distance to be 0.007 on average, ranging from 0.004 for workers with at least five years of education after bachelor degree to 0.009 for bachelor plus 2 to 4 years of additional education.

In some papers the estimation sample is restricted to young males in order to minimise problems with tied movers or stayers and restrict the effect of non-economics reasons of the move (Yankow 1999 and 2003, Bőheim and Taylor 2007). It has been shown that returns to migration vary significantly across groups of population, including age categories and the levels of educational attainment (Greenwood 1975). In the present study the robustness of the results can be verified by
using a number of sub samples, including sub samples for genders, different earning groups, and age categories.

Selectivity of migrants is a potential problem recognized by scholars. Greenwood (1997) mentions four sources of selection:

1. sampling design,

2. panel attrition,

3. time-dependent disturbances, and

4. differential behavioural responses.

The most common source is the last. Statistically, the set of migrants is not a random sample of the general population. Therefore, migrants are self-selected. There are a number of demographic factors such as age, marital status, family size, migration history, and others which characterize a population at risk of migration. Some of these variables, which successfully identify migrants, may have a significant effect in the selection equation but without that effect in the earning equation. Avoiding the selectivity problem may cause a selection bias in the empirical estimation. For example, Détang-Dessendre et al. (2004) studied the impact of migration on wages for young French males using two surveys. They found a positive selection effect for highly educated males and no selection effect for lower educated. Unfortunately for this chapter, the dynamic panel approach applied here cannot address the sample selection problems outlined above. However, the effect of at least two sources of selection, (1) and (2), is reduced by using a balanced panel of respondents who participated in all waves of the HILDA survey.
3.3 Theoretical model

The model constructed in this part is developed to describe a relationship between wage and distance moved by the migrant. It identifies a trade-off between wage in the destination and the costs of moving. The latter are assumed to be a positive function of distance. The trade-off means higher costs of moving, i.e. the longer distance moved, are associated with a higher wage at the destination. The model shows the trade-off originates from spatial maximization of returns, which are wage at the destination less the costs of moving from the origin to the destination and less wage at the origin. Financial constraints and higher costs of move per unit of distance impose restrictions on a feasible set of destinations. These restrictions affect negatively the optimal wage, costs, and benefit. As a result, the greater distance an individual is able to move the higher the optimal wage, costs, and returns. Returns to distance are shown to be generally higher for a financially restricted agent than for an unrestricted.

3.3.1 Model of migration

Let us consider population of agents living in point $x$ which belongs to $X$, a continuous set in some space $S$. Population is heterogeneous only in costs of moving per unit of distance and budget available for a moving between points. Population residing at the origin $x$ is homogeneous in human capital, implying that everyone from the origin $x$ is equally productive and has the same wage as the local population at any point of $X$. However, wage varies across different points of $X$. Wage and costs are exogenous in the model. Wage of an agent from the origin $x$ at the destination $y$, $y \in X$, is defined by a function $W(x, y)$. Assume each agent from the origin $x$ is free to choose a point $y$, migration destination, entertaining profit $\pi(x, y)$, and has a positive budget/financial constraint $B$, available to finance the moving from $x$ to $y$ in the range $(0, +\infty)$. Returns to investment in the move
are the wage at the destination $y$ less the wage at the origin $x$ and the costs of migration from the origin $x$ to the destination $y$, $C(x,y)$. The profit function is:

$$\pi(x,y) = W(x,y) - W(x,x) - C(x,y)$$

Profit must be non-negative in order to migrate. A feasible set is a search set $Z = \{ y \in X : C(x,y) \leq B \}$, a subset of $X, Z \subset X$, which depends on the point $x$, the costs function $C(x,x)$, and the budget $B$, $Z = Z(x,C(x,x),B)$. For simplicity, there are no fixed costs of moving, $C(x,x) = 0$. Verbally, an agent moves to a point with maximum profit from investment in migration choosing among points feasible by the budget constraint. The profit maximization problem for any agent residing at point $x \in X$:

$$\pi(x,y) \rightarrow \max \text{ s.t. } y \in Z$$

This is equivalent to the following problem of benefit maximization:

$$W(x,y) - C(x,y) \rightarrow \max \text{ s.t. } y \in Z$$

When there is no such economically attractive point $y$, the agent remains at the point $x$. Therefore, the minimum possible set $Z$ consists of the only point $x$, $Z = \{ x \}$, in which case the optimal profit is equal to zero. However, in general it could be possible to move to another point $y$ even if the

17 Note, under the assumption of non-zero fixed costs of migration, $C(x,x) > 0$, the maximization problem is equivalent to the problem without fixed costs and with budget restricted by $B - C(x,x)$. For very poor agents who cannot pay fixed costs the search set is empty, $Z = \{ \emptyset \}$. In a case of full information about wages and costs of moving to any point in $X$, agents with wage below some threshold above $C(x,x)$ will be worse off paying fixed costs of search in a non-empty set, $Z = \{ \emptyset \}$, but will not be able to find a better destination than the current point $x$. Then it is a rational decision for such agents not to search.
optimal profit is equal to zero. In this case the agent is indifferent between living in \( x \) and moving to \( y \).

Denote \( y^* \) as a solution of problem (6) which is a function of the initial point \( x \) and budget \( B \), \( y^*(x, B) \). For simplicity it is assumed to be a unique solution of the optimization problem. The maximization condition implies that the following inequality holds:

\[
W(x, y^*(x, B)) - C(x, y^*(x, B)) \geq W(x, y) - C(x, y) \quad \text{for any } y \in Z.
\]

One can define the optimal profit as a function

\[
\pi^*(x, B) = \pi(x, y^*(x, B)).
\]

The next step is to study properties of the solution. Two propositions below show that profit is a monotonic function of budget and costs.

**Proposition 1.** The optimal profit is a non-decreasing function of budget. That is if two agents living at origin \( x \) have different budgets such that \( B_1 > B_2 \), then profit of the higher budget agent is at least as large as that of another agent, \( \pi^*(x, B_1) \geq \pi^*(x, B_2) \).

Proof is in Appendix 2.

In order to study the behaviour of the solution with respect to costs one can consider different types of agents living at a same point. Assume there is a parameter \( \theta \) from the interval \( (0, +\infty) \), the type
of an agent, which helps us to distinguish between high and low cost individuals. Costs are an increasing function of the parameter $\theta$:

$$\frac{\partial C(x,y,\theta)}{\partial \theta} > 0$$

For this costs function the set of feasible points, optimal set, and optimal profit are functions of the parameter $\theta$ in addition to other parameters, $Z = Z(x,B,\theta)$, $y^* = y^*(x,B,\theta)$, $\pi^* = \pi^*(x,B,\theta)$.

**Proposition 2.** The optimal profit is a non-increasing function of costs. If two agents, living at origin $x$, have different types of costs such that $\theta_1 < \theta_2$, then profit of the lower costs agent is at least as much as that of another agent, $\pi^*(x,B,\theta) \geq \pi^*(x,B,\theta_2)$.

See proof also in Appendix 2.

The optimal destination $y^*(x,B,\theta)$, defined above, is a point of either a local/global maximum or some point on the ascending part of the profit function for a given point $x$. The global maximum is a point $Y^*(x)$ in which maximum profit for all parameters $(B,\theta)$ is achieved. The condition for the global maximum is

$$W(x,Y^*(x)) - C(x,Y^*(x)) \geq W(x,y) - C(x,y) \text{ for any } y \in Z$$

It is clear that for each origin $x$ and type of costs $\theta$ there could be a minimum budget $B^*(x,\theta)$ sufficient to reach the global maximum $Y^*(x)$. In the formulary view it means:
$y^*(x, B, \theta) = Y^*(x)$ for any $B \geq B^*(x, \theta)$.

Because of the trade-off between parameters $B$ and $\theta$, $B^*(x, \theta)$ is an increasing function of $\theta$. In the same manner there could be maximum costs $\theta^*(x, B)$ sufficient to move to the global maximum $Y^*(x)$, that is $y^*(x, B, \theta) = Y^*(x)$ for any $\theta \leq \theta^*(x, B)$, where $\theta^*(x, B)$ is a decreasing function of $B$.

For the concluding proposition in this section let us introduce a space metrics on the set $X$ with distance between two elements of $(x, y) \in X$ denoted as $\text{Dist}(x, y)$. Also, assume that optimal profit is a differentiable function with respect to distance. Then one can define the marginal profit as a change of profit when moving an extra unit of distance. In other words it is the derivative of the profit function with respect to distance. The marginal costs are costs of moving an extra unit of distance.

**Proposition 3.** The marginal profit is equal to zero for those agents who move to the global maximum while the marginal profit is nonnegative for those who move to other points.

Proof with some additional assumptions is in Appendix 2.

This proposition means that for an unrestricted agent, who can afford to move to the global maximum, $B \geq B^*(x, \theta)$, the marginal wage is equal to the marginal costs while for a restricted agent, who is not able to move to the global maximum, $B < B^*(x, \theta)$, the marginal wage is larger.
than the marginal costs unless it is a local maximum. But even in a local maximum a sufficiently large distance of move leads to a positive marginal profit.

### 3.3.2 Empirical implementation

It follows from the two properties of the solution of the maximization problem, proved in Propositions 1 and 2, that the optimal profit from migration is a positive function of budget $B$ and a negative function of higher costs of moving, $\theta$:

$$W(x, y^*) - C(x, y^*, \theta) - W(x, x) = f(B, \theta)$$

where $y^* = y^*(x, B, \theta)$, $\frac{\partial f(B, \theta)}{\partial B} \geq 0$ and $\frac{\partial f(B, \theta)}{\partial \theta} \leq 0$

Now, a series of simplifying assumptions is made allowing us to transform the model into a testable linear regression model. First, assume RHS of equation (7) is an additive function of two linear components:

$$f(B, \theta) = \gamma B + \eta \theta$$  \hspace{1cm} 8

Second, the unobservable budget available for moving is assumed to be proportional to wage in the current location $x$:

$$B = \delta W(x, x)$$  \hspace{1cm} 9

Third, the costs function is not observed. It is modelled as the sum of a linear function of type $\theta$ and a second function which is proportional to distance between $x$ and $y$, $Dist(x, y)$:

$$C(x, y, \theta) = \lambda \theta + \beta Dist(x, y)$$  \hspace{1cm} 10
Fourth, the type $\theta$ is also not observed. The linear function $(\lambda + \eta)\theta$ can be modelled by a linear combination of the set of observed individual characteristics, $\sum_{j=1}^{n} \tau_{j}X_{j}$ or, in vector terms, $\tau X$:

$$(\lambda + \eta)\theta = \tau X$$ \hspace{1cm} (11)$$

Altogether, equation (7) is transformed with the help of assumptions (8)-(11) into the empirical model:

$$W(y) = \alpha W(x) + \beta \text{Dist}(x, y) + \tau X + \varepsilon$$

The equivalent panel data specification of this model is:

$$W_{it} = c + \alpha W_{i,t-1} + \beta \text{Dist}_{it} + \tau X_{it} + \lambda_{i} + \mu_{t} + \nu_{it} + \varepsilon_{it}$$

where $i$ stands for an individual, $t$ is a year, $\alpha$ is an autoregressive parameter (in the empirical model below three lags will be included instead of one), $\text{Dist}_{it}$ is a distance moved by the individual at year $t$. In addition the model includes common in panel econometrics unobserved individual characteristic $\lambda_{i}$, macroeconomic effect $\mu_{t}$, and regional dummy variable $\nu_{it}$.

This model has a similar view to the Mincer earning equation. However, traditional migration dummy is substituted by distance, the continuous variable. According to Proposition 3, parameter $\beta$, the derivative of wage with respect to distance, is larger for constrained individuals who are only able to finance their move to a local optimum at best. In order to capture this effect, in another model specification the interactive term of the lagged wage and distance, $W_{i,t-1}\text{Dist}_{it}$, is included in addition to other independent variables.
Distance should be treated as an endogenous variable in this model. The reason is that distance and wage are likely to be determined simultaneously since both are characteristics of the optimal destination. Distance is observed before move but total costs and benefits from migration are realised with some delay in time.

3.4 Empirical analysis

The econometric analysis is based on micro level data. A standard Mincer-type equation for individual earnings is estimated with a set of control variables which include personal characteristics common for the earnings equation including unobserved individual heterogeneity and unobserved regional characteristics. A dynamic panel data technique (the Arellano-Bond method) is used to find unbiased estimates of the model. In this section, the model and the methodology of the econometric estimation are described. The HILDA migration data is explained and descriptive statistics for the sample of individuals analysed is presented. Next, a system GMM method for the dynamic model on panel data is applied assuming the autoregressive term to be predetermined. Initially, migration distance, the variable of interest, is assumed to be exogenous. This assumption is relaxed to allow for weak endogeneity and the regression analysis is repeated assuming distance and its interactions with other variables to be predetermined.

3.4.1 Methodology

The model to be studied here is a modified Mincer-type earnings equation. In its classic form, wage is a function of individual characteristics such as age, work experience, and sometimes their squared terms, and education. In migration studies the model is augmented by present, past and future migration dummy variables (e.g. Yankow 2003). A short panel at hand does not allow using lags and leads other than an autoregressive term. The difference of the present model from earlier
earnings equations is that it has a dynamic form and instead of a traditional independent binary variable for migration the effect of migration distance is modelled in the following specification

\[ W_{it} = c + \alpha_1 W_{it-1} + \alpha_2 W_{it-2} + \alpha_3 W_{it-3} + \beta Dist_{it} + X_{it} + \lambda_i + \mu_i + \nu_{it} + \varepsilon_{it} \]  \hspace{1cm} (M1)

Where \( W_{it} \) is log of hourly wage at year \( t \) of person \( i \), \( \alpha \) is a constant term, \( X_{it} \) is a vector of individual characteristics at year \( t \) which includes education, \( Educ_{it} \), and labour force status, \( Dist_{it} \) is log of distance moved between years \( t-1 \) and \( t \), \( \lambda_i \) is the individual \( i \) fixed effect, \( \mu_i \) is a time dummy (unobserved macroeconomic effect at year \( t \)), \( \nu_{it} \) is a labour market/regional dummy (unobserved effect of a current region), and \( \varepsilon_{it} \) is an error of the model.

Several interesting estimation issues arise. First, OLS method cannot be used for the dynamic model estimation. This is due to two issues; serial autocorrelation of errors, and the autoregressive term appears to be predetermined; that is correlated with the past errors of the model. Indeed, since the error term is by definition correlated with the dependent variable then the autoregressive term is correlated with any lagged error, which in turn is positively correlated with the current realisation of the error. In this case the core assumption of consistency for OLS estimation is violated.

Second, the choice of three lags of the dependent variable is explained by post-estimation diagnostics. Only three lags led to non-rejection of the zero hypothesis of no overidentifying restrictions in this model estimated by GMM.

Distance variable is assumed to be exogenous in the first model (M1) and then predetermined (weakly endogenous) in the second model (M2). The weak endogeneity assumption, that is correlation between distance and the lagged error term in the earnings model, is realised when an individual is underpaid in the current location relative to somewhere else and, therefore, has a higher risk to move a large distance in the next period.
In the third model (M3) the effect of education on returns to migration is studied by means of interaction between distance and mean centred education, \( Dist_a (\text{Educ}_{a} - 12.6) \). In addition to this, the fourth specification (M4) models the differential effect of distance across wage levels, and includes predetermined interactive term between distance and the mean centred autoregressive wage, \( Dist_a (W_{a-1} - 3.0) \). In the fifth model (M5), the effect of income is studied in even more detail. The fourth model is reorganized such that instead of the lagged wage and distance variables use is made of interactive terms between distance and a set of four wage quartiles dummies

\[
Dist_a Q1_{a-1}, ..., Dist_a Q4_{a-1}.
\]

In the final model specification (M6) additional consideration is given to another interactive term for distance, wage groups and the group mean centred autoregressive term

\[
Dist_a Q1_{a-1}(W_{a-1} - W_{a-1} Q1_{a-1}) , ..., Dist_a Q4_{a-1}(W_{a-1} - W_{a-1} Q4_{a-1}) .
\]

Another model was also studied which includes all possible combinations of variables results in seven extra variables in addition to those in the model M6. These are three wage quartile dummies \( Q1_{a-1}, ..., Q3_{a-1} \) and interaction of the lagged wage and wage group dummies \( W_{a-1} Q1_{a-1}, ..., W_{a-1} Q4_{a-1} \). However, its results are not reported here for brevity as they are very close to those in M6. All interactive variables in models M1-M6 are assumed to be predetermined as well as the number of working hours.

The system GMM is applied for the dynamic model. This estimation technique is developed for dynamic panel data model with predetermined and endogenous independent variables (Arellano and Bond 1991; Arellano and Bower 1995; Blundell and Bond 1998). The system combines two equations, the shown equation in levels and its first differenced form. The system GMM provides
more efficient estimators as compared to the GMM estimator for the model in first differences or
the model in levels (Hayakawa 2007).

There are two moment conditions for the equations in levels and first differences. These can be
derived for any predetermined variable including the lagged dependent variable. The moment
conditions are based on the fact that correlation of the lag of predetermined variable and first
difference of the error is zero as well as correlation of first difference and the error. Instruments
used in the estimation are obtained from these moment conditions. As a result, the set of
instruments for the equation in levels includes first differences for predetermined variables while
instruments for the equation in first differences comprise the lag of predetermined variables. In
addition, all exogenous variables are used as instruments for themselves. As migration is measured
using a continuous variable, i.e. distance of move, this variable can be modelled as predetermined;
unlike the model with a migration dummy variable in which endogeneity is technically difficult to
control for. Although distance is a continuous variable, it is left censored to zero value in 96 percent
of observations; this does not appear to be a problem in the estimation process reported here.

Interestingly, this methodology implies that the coefficient for distance stems mostly from the
equation in levels.18 This is because for current migrants distance is instrumented by first difference
of distance, which is equal to distance for 85 percent of migrants, since 15 percent move two years
in a row. By contrast, in another equation in the system, for the abovementioned 85 percent of
migrants, first difference of distance, equal to distance, is instrumented by the lagged level; that is
by zero. Note also that in each of these two equations, zero is instrumented by zero in 97 percent of
the sample and that simple linear regression finds significant coefficients.

\[\text{Note: Mandatory citation}\]

18 Noteworthy, an estimation of the earnings equation in levels has dynamic panel bias because of correlation
between the autoregressive term and the fixed effects.
Following the recommendations of Roodman (2006) in his user guide for STATA routine xtabond2, time dummies were added to the model specification since with them an assumption of no correlation across individuals in the idiosyncratic disturbances is more likely to hold. A number of options with the routine xtabond2 were used including orthogonal deviations in order to maximise sample size in a panel with gaps. The two-step procedure was applied as it is asymptotically efficient and robust, although it may result in downward biased standard errors, and a small-sample correction to the covariance matrix estimate.

Post-estimation includes a test on overidentifying restrictions imposed by these assumptions on moment conditions. The Hansen test verifies joint validity of a list of instruments, separately for the equation in first differences and levels. The Arellano and Bond test is used to test autocorrelation in disturbance term. This test is used to verify validity of lags as instruments.

Selectivity of migrants is a potential problem recognized by researchers (Greenwood 1997). There are a number of reasons for sample selection in the present study. First, the analysis is based on wage earners and excludes from the sample three sub-categories of migrants who are in labour force. These exclusions are people unemployed in their pre-migration year but who found or were still looking for a job in new location, or were employed at their old location but unemployed post-migration. Second, sample selection can be caused by panel attrition which is in part due to migrants who were not tracked by the survey. The sample selection problem arising here is not studied in this chapter because the methodology for the panel data sample selection is not yet successfully realised in empirical studies. However, it is planned to explore and apply other methods and study the sample selection effect in future stages of this research. The present focus is

---

19 Technically, orthogonal transformation deducts future means from the contemporaneous level.
20 Although there are methods in Kyriazidou (1997, 2001), these are not yet widely used due to poor behaviour of the estimates. An alternative method which can be potentially used to address the problem here is found in Wooldridge (2002).
on reducing the effect by using a balanced panel of respondents participated in all seven waves of the HILDA survey and including a large set of explanatory variables.

3.4.2 Data

The data analysed here is longitudinal data from the HILDA Survey. The survey is a nationally representative sample of households. The panel is surveyed annually from August to March the following year. The first wave started in August 2001. The present analysis utilises data from the first seven waves, i.e. up to and including 2007, but the focus is essentially on the four latest waves. The use of the first three waves is very limited: only hourly wage is taken from waves 1 to 3 to construct independent variables and distance moved by migrant in wave 3 for instruments. The main advantage of the survey is that it tracks the initial sample of individuals even in case of internal migration. Initial inspection of the HILDA Survey found that about 17% of the population moved each year; mostly on short distances of less than 10 kilometres and due to personal or family reasons (Headey et al 2006).

However, there is a drawback in the HILDA survey. Attrition of the panel sample is to a large extent due to migrants not being successfully tracked. As a result the survey is potentially not representative for migrants, and this problem is aggravated in the course of time because HILDA does not attempt to replace these migrants by people and households with comparable risk of migration. The attrition problem is reduced in the analysis below which uses individual longitudinal weights for a balanced panel drawn from waves 1 to 7.
3.4.2.1 Dependent variable and sample

Respondents are asked about *usual* weekly wage in the main job and wage in all other jobs for the last week at time of interview.\(^{21}\) Since the model of interest is the dynamic earning equation with the first three lags as explanatory variables and the forth lag in the list of instruments, the sample is restricted to those who report a weekly wage at least for five successive waves. A real weekly wage is measured in 2001 Australian dollars and calculated as reported current weekly gross wages and salary from all jobs divided by the relevant regional state’s cumulative CPI. The measure of earnings used as a dependent variable is a real hourly wage defined as the real weekly wage divided by the total number of hours per week *usually* worked in all jobs (see also variables definitions in Table A2.1 of Appendix 2).

Seven waves of the HILDA survey contain weekly wage and hours of work for 49,487 respondents out of a total of 90,599 adult respondents. This study is able to make use 33,073 of hourly wage observations that are in a balanced panel for waves 1 to 7. In order to minimise the effect of outliers, the sample is further restricted to individuals with a real wage in the range 3 to 200 dollars per hour in 2001 prices. This is excludes less than 1 percent of observations; 0.7 percent of observations from the left tail of the hourly real wage distribution in the pooled sample and 0.07 percent of observations from the right tail. The sample is also restricted to include only individuals who are not full time students and are aged between 15 and 64 (for males) or 60 (for females). As the model is a dynamic model, over 60 percent of the sample is lost due to a missing value for a wage in any of four previous waves. These restrictions result in the total number of observations in all regressions being 11,846. In the results reported here, all calculations are based on a sample of

\(^{21}\) The usual weekly wage is used as it is probably the best available measure of wage in the HILDA. For example, last financial year income is not recorded for 6 percent of individuals who report weekly wage and obviously includes income before and after move in the case of migration.
these individual observations using responding person longitudinal weight balanced for waves 1 to 7.

3.4.2.2 Independent variables: distance

A common definition of internal migration is a change of usual place of residence or a move above some specified minimum distance (UN 1970).\textsuperscript{22} For practical purposes in this chapter migration is defined using the distance based definition. It is assumed that migration is a move of a distance over 30 kilometres as these movers are more likely to change their labour market. Such a small threshold value is chosen in order to keep more migrants in the model, although this may be at the expense of possibly reduced accuracy in the estimates. There are several different definitions of minimum distance in the empirical literature. Thus, in two papers migration was defined to have occurred if a respondent moved at least 50 miles or changed Metropolitan Statistical Area and moved at least 20 miles (Ham et al 2005, Baumann and Reagan 2002). Ham et al (2005) found that both prevailing in the USA definitions of migration, based on changing state or county, significantly underestimate and overestimate (minimum distance definition) migration respectively. Détang-Dessendre and Molho (1999) use the definition of long-distance migration as moves over 60 kilometres, which, they believe, almost definitely reduce a risk of commuting to previous district. In other three papers (Boyle (1995), Boyle \textit{et al.} (2001), Clark and Huang (2004)) a long-distance move is a move over 50 kilometres. For authors in the article Boyle \textit{et al.} (2001) it seems likely that most of migrants moving above 50 kilometres are motivated by employment factors.

In the sample used for the empirical model there were 384 migrations over 30 kilometres out of the 11,846 individual-year observations. These moves account for 3.2 percent points out of the 15 percent moving house within Australia annually. Not more than 30 migrations out of 384 were

\textsuperscript{22} It is also possible to apply another definition of migration which combines a change of both place of residence and place of work (Greenwood, 1997). This definition will be applied in future research on this project.
likely to be within a city boundary if the definition of a Major Statistical Region was adopted for the largest cities used in HILDA data. Nonetheless, these intra-city moves are of no particular importance, since exclusion of them from migration category does not change the results of the empirical analysis below as demonstrated by additional robustness checks (not reported).

The HILDA survey provides a distance measure for those individuals who have changed address and who were followed up beginning from the second wave in 2002. It is the great-circle distance, also known “as the crow flies” distance, measured in kilometres, between geocoded current and previous interviews residential addresses. These long distance moves represent roughly moves between labour markets which are often connected with a new job. In particular, 10 percent of the stayers in the sample have changed an employer since a previous interview as compared to 14 percent of the short distance movers and 49 percent of those who migrated on medium and long distance (author own calculations based on HILDA).

There is no detailed information about in-between moves if a respondent moved more than once between two subsequent interviews. It is known only when they moved to the current address. Also there is no way to distinguish between stayers and movers on minimum distance in the HILDA, which is approximated by 0 kilometres. Therefore, the data set zero value for distance in both categories.\(^\text{23}\)

There is additional evidence that distance may not be precisely measured since there are two types of errors found in a HILDA technical report of Watson and Wooden (2004). They observed that among 1915 respondents, who changed address between waves 1 and 2 according to the Household File used by interviewers, 119 indicated they did not move. Another 141 “movers” were identified as living at the same address.

\(^{23}\) Because of the rounding, this should mean that persons who moved less than 500 metres are assigned zero kilometres distance.
Distance varies between 1 and 3649 kilometres for movers. Interestingly but not surprisingly, the number of moves is a decreasing function of distance which can be described by the discrete probability (Poisson) distribution function. Thus, out of 1,642 moves the number of moves on distance 1 km is 214, on 2 km 155 moves, on 3 km 154 moves, on 4 km 113 and so on, on 31 km 10, on 32 km 4, on 33 km 5 and so on.

3.4.3 Descriptive analysis

Table A2.1 in Appendix 2 presents the definitions for all variables in the model. Summary statistics for these variables are in Table A2.2. A comparison of the characteristics of migrants on different distances and stayers is shown in Table 13. That Table reveals short distance (below 30 kilometres) movers have comparable average characteristics to medium distance (from 30 to 300 kilometres) and long distance (over 300 kilometres) migrants but are significantly different from stayers. Movers and migrants are better educated, younger, more often males, unmarried, and employed in small size businesses, have lower occupational tenure and fewer children. Table compares wages and working hours at current and previous years. Migration on medium and long distance increases real weekly wage by 6 to 7 percent respectively, whereas stayers and movers on short distances have 3 percent real wage growth. The decomposition of wage growth which can be made from Table demonstrates a structure of wage premium for migrants. Three percent growth of real weekly wage for both stayers and short distance movers stems entirely from similar growth of real hourly wage because their working hours does not change. Migrants on medium distance decrease working hours by 3 percent and have 9 percent real hourly wage growth but long distance migrants reduce working hours by 1 percent and still raise real wage per hour by 8 percent. Therefore, the average wage premium for migration is in the range of 3 to 4 percent for weekly earnings and between 5 and 6 percent for hourly wage.
Descriptive tables 14 to 20 below are constructed from individual observations weighted by the longitudinal sample weight, that is from the balanced panel. Table 14 reports the distribution of the sample by distance moved. Movers on distance less than 1 kilometre are treated as stayers. One can note that a sizable proportion in the range 11-13 percent of the sample moves between any two waves. Most of these moves, around 10 percent, can be classified as short distance moves within a radius of 30 kilometres. Such moves are usually localized within one labour market and not associated with a change of job.

Analysis of mobility among different income groups shows a non-linear pattern, see Table 15. Wage quartiles are defined for each wave according to weekly real wage distribution in the sample. The lowest and the highest income quartiles consist of more stayers than the two middle income groups. This is a U-shaped income effect for stayers. The most mobile group of population is in the second income group while the lowest quartile of population is the least mobile.

For short and medium distances, but not for long distances, an inverse U-shaped income effect is observed. This is in accordance with the income opportunity hypothesis. In contrast, the poorest quartile potentially could be the most mobile because of more income improving opportunities. But in reality one can observe that the first income group is much less mobile than the second and even the third groups. Moreover, it is the least mobile on long distances. This is apparently mostly due to financial constraints. The extra evidence of that: among all income groups the richest quartile is about 50 percent more likely to be involved in job related long distance migration (author’s own calculation).
Table 14. Descriptive statistics, movers vs. stayers*

<table>
<thead>
<tr>
<th></th>
<th>Stayers</th>
<th>≤30km</th>
<th>&gt;30km &amp; ≤300km</th>
<th>&gt;300km</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weekly wage previous year, log</td>
<td>6.63</td>
<td>6.61</td>
<td>6.63</td>
<td>6.71</td>
<td>6.63</td>
</tr>
<tr>
<td>Weekly wage current year, log</td>
<td>6.66</td>
<td>6.64</td>
<td>6.69</td>
<td>6.78</td>
<td>6.66</td>
</tr>
<tr>
<td>Hourly wage previous year, log</td>
<td>2.98</td>
<td>2.92</td>
<td>2.88</td>
<td>2.99</td>
<td>2.97</td>
</tr>
<tr>
<td>Hourly wage current year, log</td>
<td>3.01</td>
<td>2.96</td>
<td>2.97</td>
<td>3.07</td>
<td>3.00</td>
</tr>
<tr>
<td>Weekly working hours previous year, log</td>
<td>3.65</td>
<td>3.68</td>
<td>3.75</td>
<td>3.72</td>
<td>3.66</td>
</tr>
<tr>
<td>Weekly working hours current year, log</td>
<td>3.65</td>
<td>3.68</td>
<td>3.72</td>
<td>3.71</td>
<td>3.66</td>
</tr>
<tr>
<td>Distance, log</td>
<td>0</td>
<td>0</td>
<td>4.36</td>
<td>6.89</td>
<td>0.16</td>
</tr>
<tr>
<td>Education</td>
<td>12.6</td>
<td>12.7</td>
<td>12.7</td>
<td>13.0</td>
<td>12.6</td>
</tr>
<tr>
<td>Tenure</td>
<td>11.3</td>
<td>7.4</td>
<td>7.3</td>
<td>6.8</td>
<td>10.8</td>
</tr>
<tr>
<td>Age</td>
<td>42.4</td>
<td>35.5</td>
<td>37.6</td>
<td>35.9</td>
<td>41.5</td>
</tr>
<tr>
<td>Sex</td>
<td>0.58</td>
<td>0.60</td>
<td>0.61</td>
<td>0.63</td>
<td>0.58</td>
</tr>
<tr>
<td>Marital status</td>
<td>0.73</td>
<td>0.62</td>
<td>0.61</td>
<td>0.63</td>
<td>0.72</td>
</tr>
<tr>
<td>No. of children 0-14 years</td>
<td>0.64</td>
<td>0.50</td>
<td>0.57</td>
<td>0.60</td>
<td>0.63</td>
</tr>
<tr>
<td>Firm size unknown</td>
<td>0.005</td>
<td>0.003</td>
<td>0</td>
<td>0.009</td>
<td>0.004</td>
</tr>
<tr>
<td>Firm size 1-9</td>
<td>0.21</td>
<td>0.22</td>
<td>0.20</td>
<td>0.27</td>
<td>0.21</td>
</tr>
<tr>
<td>Firm size 10-49</td>
<td>0.30</td>
<td>0.34</td>
<td>0.41</td>
<td>0.32</td>
<td>0.30</td>
</tr>
<tr>
<td>Firm size 50-499</td>
<td>0.35</td>
<td>0.31</td>
<td>0.31</td>
<td>0.28</td>
<td>0.35</td>
</tr>
<tr>
<td>Firm size 500+</td>
<td>0.14</td>
<td>0.12</td>
<td>0.08</td>
<td>0.13</td>
<td>0.14</td>
</tr>
</tbody>
</table>

*weighted by individual longitudinal weights for balanced panel
Table 15. Distribution of the sample by distance of move, percentage

<table>
<thead>
<tr>
<th>Wave</th>
<th>Stayers</th>
<th>≤30km</th>
<th>&gt;30km &amp; ≤300km</th>
<th>&gt;300km</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>86.8</td>
<td>10.5</td>
<td>1.4</td>
<td>1.3</td>
<td>100.0</td>
</tr>
<tr>
<td>5</td>
<td>86.8</td>
<td>10.3</td>
<td>1.3</td>
<td>1.6</td>
<td>100.0</td>
</tr>
<tr>
<td>6</td>
<td>88.7</td>
<td>8.8</td>
<td>1.0</td>
<td>1.6</td>
<td>100.0</td>
</tr>
<tr>
<td>7</td>
<td>87.9</td>
<td>9.2</td>
<td>1.4</td>
<td>1.6</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>87.5</td>
<td>9.7</td>
<td>1.3</td>
<td>1.5</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 16. Distribution of the sample by wage quartiles and distance of move, percentage

<table>
<thead>
<tr>
<th>Wage quartile previous year</th>
<th>Stayers</th>
<th>≤30km</th>
<th>&gt;30km &amp; ≤300km</th>
<th>&gt;300km</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>89.3</td>
<td>8.6</td>
<td>1.1</td>
<td>1.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Q2</td>
<td>85.6</td>
<td>11.3</td>
<td>1.6</td>
<td>1.5</td>
<td>100.0</td>
</tr>
<tr>
<td>Q3</td>
<td>86.5</td>
<td>10.8</td>
<td>1.2</td>
<td>1.5</td>
<td>100.0</td>
</tr>
<tr>
<td>Q4</td>
<td>89.2</td>
<td>7.9</td>
<td>1.2</td>
<td>1.7</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>87.5</td>
<td>9.7</td>
<td>1.3</td>
<td>1.5</td>
<td>100.0</td>
</tr>
</tbody>
</table>

The next step is to proceed to data description which can support the hypothesis on the positive effect of distance on wage premium. For a moment the ceteris paribus assumption is relaxed and returns to distance are compared across income groups. Table 16 compares growth of mean wage
for stayers and movers by distance (columns) and initial income group (rows). Each cell of the table shows the proportion of the sample which has a positive real weekly wage change since a previous wave. The table shows 57 percent of the sample had real income growth independent of whether they stayed or moved any distance. Headey and Warren (2008) found a percentage of individuals whose equivalent income rose between the first and fifth waves of the HILDA survey to be equal to 58. Also, Table 16 demonstrates that the highest proportion of gainers, 86 percent, is observed among long distance migrants in the lowest income group. The lowest proportion, 44 percent, is for long distance movers from the highest income quartile. The poorest quartile gains the most in terms of wages after long distance migration, increasing proportion by 21 percent while the richest half of wage earners incur losses in terms of weekly wage at least in the short run. These observations are also confirmed from observations which can be made from Table 17.

Table 17. Proportion of real weekly wage gainers, by quartiles and distance moved, percentage

<table>
<thead>
<tr>
<th>Wage quartile previous year</th>
<th>Stayers</th>
<th>≤30km</th>
<th>&gt;30km &amp; ≤300km</th>
<th>&gt;300km</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>67</td>
<td>75</td>
<td>80</td>
<td>86</td>
<td>68</td>
</tr>
<tr>
<td>Q2</td>
<td>61</td>
<td>61</td>
<td>69</td>
<td>70</td>
<td>61</td>
</tr>
<tr>
<td>Q3</td>
<td>56</td>
<td>58</td>
<td>45</td>
<td>52</td>
<td>56</td>
</tr>
<tr>
<td>Q4</td>
<td>48</td>
<td>46</td>
<td>46</td>
<td>44</td>
<td>48</td>
</tr>
<tr>
<td>Total</td>
<td>57</td>
<td>58</td>
<td>57</td>
<td>57</td>
<td>57</td>
</tr>
</tbody>
</table>

One can look at income returns to distance from another angle, finding out a change of the mean real wage (in logs) in a particular income group across all waves, see Table 17. Calculations are
based on two tables (similar to Table 16) of average wages in a particular category by income group and distance. On average real wage growth between any two waves is 3 percent. By and large, conclusions are similar to those from the previous table.

**Table 18. Growth of mean real weekly wage, by quartiles and distance moved, percentage**

<table>
<thead>
<tr>
<th>Wage quartile previous year</th>
<th>Stayers</th>
<th>≤30km</th>
<th>&gt;30km &amp; ≤300km</th>
<th>&gt;300km</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>17</td>
<td>23</td>
<td>47</td>
<td>59</td>
<td>19</td>
</tr>
<tr>
<td>Q2</td>
<td>4</td>
<td>6</td>
<td>11</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Q3</td>
<td>2</td>
<td>1</td>
<td>-6</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Q4</td>
<td>-5</td>
<td>-6</td>
<td>-7</td>
<td>-7</td>
<td>-5</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

Short distance movers as anticipated do not differ markedly from stayers. Distance helps to improve wages for the bottom income group but discounts wages for the top income quartile. One can note only lower paid labour, with wages below median, can significantly improve wage by changing the labour market. The lowest income quartile benefits the most from medium and long distance migration earning from 47 to 59 percent more after migration as compared to 17-23 percent in the case they stay or move locally. Their wage premium for medium/long distance migration is estimated to be from 30 to 42 percent. For other income group wage premium is not so large. It is either small positive premium about 7 percent, for the second income group, or even negative for the above median income earners. This preliminary finding seems to support the hypothesis that the positive effect of distance on the wage premium is declining with wage.
Negative income growth among the higher income group is something of a surprise. However, a report of Headey and Warren (2008) presents another evidence of downward income mobility for the top income group. Using the HILDA data the authors constructed a transition matrix for

Table 19. Growth of mean real hourly wage, by quartiles and distance moved, percentage*

<table>
<thead>
<tr>
<th>Wage quartile previous year</th>
<th>Stayers</th>
<th>≤30km</th>
<th>&gt;30km &amp; ≤300km</th>
<th>&gt;300km</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>19</td>
<td>20</td>
<td>40</td>
<td>32</td>
<td>19</td>
</tr>
<tr>
<td>Q2</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Q3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Q4</td>
<td>-7</td>
<td>-9</td>
<td>-7</td>
<td>-5</td>
<td>-7</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>4</td>
<td>9</td>
<td>8</td>
<td>3</td>
</tr>
</tbody>
</table>

*Wage quartile is defined from the distribution of hourly wage in a particular year.

Table 20. Change of mean working hours, by quartiles and distance moved, percentage

<table>
<thead>
<tr>
<th>Wage quartile previous year</th>
<th>Stayers</th>
<th>≤30km</th>
<th>&gt;30km &amp; ≤300km</th>
<th>&gt;300km</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>1.6</td>
<td>1.8</td>
<td>3.1</td>
<td>8.9</td>
<td>1.7</td>
</tr>
<tr>
<td>Q2</td>
<td>0.0</td>
<td>0.4</td>
<td>-2.7</td>
<td>1.3</td>
<td>0.0</td>
</tr>
<tr>
<td>Q3</td>
<td>-0.5</td>
<td>-0.5</td>
<td>-2.0</td>
<td>-3.4</td>
<td>-0.6</td>
</tr>
<tr>
<td>Q4</td>
<td>-0.6</td>
<td>-0.8</td>
<td>-1.8</td>
<td>-3.6</td>
<td>-0.7</td>
</tr>
<tr>
<td>Total</td>
<td>-0.1</td>
<td>0.0</td>
<td>-1.5</td>
<td>-1.0</td>
<td>-0.1</td>
</tr>
</tbody>
</table>
Table 21. Growth of mean real annual wage, by quartiles and distance moved, percentage*

<table>
<thead>
<tr>
<th>Wage quartile previous year</th>
<th>Stayers</th>
<th>≤30km</th>
<th>&gt;30km &amp; ≤300km</th>
<th>&gt;300km</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>23</td>
<td>33</td>
<td>49</td>
<td>55</td>
<td>25</td>
</tr>
<tr>
<td>Q2</td>
<td>4</td>
<td>3</td>
<td>-1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Q3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Q4</td>
<td>-2</td>
<td>-3</td>
<td>4</td>
<td>-2</td>
<td>-2</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>11</td>
<td>5</td>
</tr>
</tbody>
</table>

*Wage quartile is defined from the distribution of annual wage in a particular year.

Equivalised income, according to which 42 percent of the (top) fifth income quintile moved down into the lower quintiles between 2001 and 2005.

In part, the positive income growth for poor and negative for rich shown in the previous table is due to changes in both the hourly wage and the working hours. Decomposition made in Tables 18 and 19 demonstrates this. Indeed, Table 18 reveals a positive change in hourly wage for the poor and negative for the rich which mostly explains weekly wage growth in both income groups. The poor increased hours of work substantially while the rich somewhat reduced them as evidenced by Table 19. The largest change in working hours, from 26.5 to 35.4 hours, is observed among the lowest income migrants on long distance (it is worthy of note that all together the lower income are still part-time employed, less than 29 hours per week on average), whereas highest income migrants experience the largest fall, from 49.8 to 46.2 hours per week. Table 20 reveals, however, that annual wage decrease is not as dramatic as the decrease in hourly wages in Table 18 for the top income
group. Moreover, Table 20 shows that annual wage premium is on average nil for movers, 1 percent for medium distance migrants and 6 percent for long distance migrants.

### 3.4.4 Econometric analysis

The detailed results for the system GMM analysis for the dynamic earnings model are available in Table A2.3 of Appendix 2. Dynamic components of current earnings, wages at three previous years, are all significant, with declining impact in the course of time. The model initially estimated with distance being exogenous, M1 in the Table. Distance of migration is found to have a positive, though insignificant, effect on hourly wages. Distance elasticity is equal to 0.005 and is insignificantly different from zero. The next column M2 shows the estimation results when distance is predetermined. The results of the model estimation under M2 suggest that distance elasticity falls to 0.004 and that the hypothesis of no effect of distance on wage cannot be rejected.

Models M3-M6 study the effect of education and pre-migration wage on returns to distance. The interaction of distance and mean centred education is significantly positive in every specification. According to model M3 for the average level of education, 12.6 years, distance is insignificant with elasticity 0.002. However, each additional year increases elasticity by 0.002. The statistically positive effect of education on distance elasticity is robust across all models M3-M6 which include the interaction between distance and education.

The model was then extended by including another predetermined variable, the interaction between distance and the mean centred autoregressive term, see M4. Both distance and the interactive term are found to be statistically insignificant. Results indicate a positive effect of distance conditional on a mean pre-migration wage, with elasticity 0.001. A negative interactive term between distance and pre-migration wage implies this positive effect reduces with wage.
In models 5 and 6 the differential effect of distance across wage quartiles was modelled. Model M5 says that the average distance elasticity for the lowest wage group is 0.019 and highly significant. Other three wage quartiles have statistically zero effect of distance. In model 6 in addition to this the non-linear effect of pre-migration wage for each income group is included. For the sake of simplicity of results interpretation, wages in models M6 are mean centred in their interaction with distance, i.e. the mean of the correspondent wage group is deducted.

It is found that the effect of distance conditional on a mean wage is significantly positive for the bottom quartile, with elasticity 0.023 on average. This effect is decreasing with higher wage, as shown by statistically negative interaction with wage. For other wage quartiles neither of these two coefficients is significant.

Diagnostics, as indicated by the Hansen test, performs well in all six models. The hypothesis of no overidentifying restrictions for the hourly wage equations is not rejected. The Arellano-Bond test indicates that there is first order serial autocorrelation of errors while there is no second order serial autocorrelation; both of which are in accordance with theoretical expectations.

The average effect of migration from the coefficients obtained in M5 suggests that the wage premium for lower income and well-educated migrants is large. For example, consider a migrant from the lowest income quartile who has 12 years of education and who moves a distance of 300 kilometres. This move, which is roughly the distance between Sydney and Canberra, results in a 10 percent hourly wage premium.24 A corresponding migrant with a bachelor degree would earn a 15 percent premium. If he/she moves instead 3,000 kilometres, equivalent to say from Sydney to Perth, the wage premium is higher at 21 percent. A move of only a small distance in this model, 30 kilometres, still is worthwhile with a premium of 9 percent.

24 Distance 300 kilometres corresponds to the average log of distance for migrants over 30 kilometres in the sample.
Finally, a number of checks to assess the robustness of the estimates were undertaken. First, migration was defined in two alternative ways (i) all movers even if the distance is less than 30 kilometres, (ii) all migrants over 50 kilometres. Estimated coefficients are in line with what was observed in the main models presented in Table A2.3.

In the second robustness check, models were estimated separately for a series of sub samples. These were: all males, all people aged below 36 years, females aged below than 36 years, all people aged over 45 years, people with wage below median wage, and people with incomes above the median wage. On other samples such as sample of all females and young males the system GMM estimates are not found. All results are close to the main results and not reported here.

In the third robustness check self-employed individuals were excluded from the sample. This did not alter the results significantly. The fourth check used the Windmeijer (2005) small-sample correction method for the two-step standard errors. Without such a correction standard errors tend to be severely downward biased. It was found, however, that new standard errors did not differ much from the original and as a result, significance of the estimates is not affected.

### 3.5 Conclusions

In this chapter several hypotheses concerning the effect of geographical mobility on the wage premium were tested. Previous literature and the spatial job search framework suggest that there are more opportunities with the greater distance, at least according to the distance opportunity hypothesis. The same prediction, of a higher wage premium for long distance moves because of various distant dependent costs associated with a move, arises from the costs hypothesis. Applying the idea of the optimal stopping rule for sequential job search it was hypothesized that the lower income groups may have larger returns to migration and distance than the higher income groups.
These hypotheses were tested using individual level data from the HILDA, a rich longitudinal survey of Australian households. Preliminary descriptive analysis demonstrated that moving on long distance may result in a large wage premium for relatively lower paid migrants and a negative premium for relatively higher paid migrants. Results obtained from the system GMM estimation of dynamic wage equation on panel data appear to confirm hypotheses. It was found that in contrast to the higher paid workers who do not have returns to distance the lower paid workers earn a large premium for migration and that this premium increases with the level of education and distance moved. Finally, some possible extensions of the empirical analysis are:

- to compare the earning equations for migrants and non-migrants,
- to explore an idea from compensating differentials theory and add individual characteristics which were observed in a sending region,
- to use the squared log of distance in addition to the linear term to better represent the distance opportunity hypothesis,
- to study the effect of purchasing power of income accounting for price levels rather than real income, and
- to address the sample selection problem.
## Appendix 2

### Table A2.1. Definition of variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition / Question</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Log) Real weekly wage, AUD 2001</td>
<td>Current weekly gross wages and salary from all jobs divided by the state CPI, CPI 2001=100</td>
<td>CPI is taken for the last quarter current year to the last quarter last year, assuming December 2001 quarter CPI=100. Data from ABS (2009b)</td>
</tr>
<tr>
<td>(Log) Real hourly wage, AUD 2001</td>
<td>Real weekly wage over the total number of hours per week usually worked in all jobs</td>
<td></td>
</tr>
<tr>
<td>(Log) Working hours</td>
<td>Hours per week usually worked in all jobs</td>
<td>Inaccuracy in hours of work: a week preceding an interview can be not representative, that is not usual for wages from all other jobs</td>
</tr>
<tr>
<td>(Log) Distance</td>
<td>The great-circle distance, also known &quot;as the crow flies&quot;, applied to latitude and longitude of geocoded addresses for the current and previous interviews (rounded to an integer number of kilometres). Log of distance for stayers and short distance moves up to 30 kilometres is zero.</td>
<td>There is no detailed information about in-between moves if a respondent had multiple moves between two consecutive interviews, only the most recent move, i.e. to the current address, is known.</td>
</tr>
<tr>
<td>Education</td>
<td>The number of years of formal education: 17 postgraduates - masters or doctorate, 16 graduate diploma or graduate certificate, 15 bachelors, 13 advanced diploma or diploma, 12 year 12 or certificate III or IV, 11 year 11 or certificate I or II, 10 year 10 or certificate not defined, 9 to 7 year 9 to 7, 4 'did not attend secondary school but finished primary school', and 2 'attended primary school but did not finish'</td>
<td></td>
</tr>
<tr>
<td>Tenure</td>
<td>The number of years in current occupation</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>Age of a person</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>0 for females and 1 for males</td>
<td></td>
</tr>
<tr>
<td>Marital status</td>
<td>0 for never married, separated, divorced or widowed, 1 legally or de facto married</td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Number of children</td>
<td>The number of household members minus the number of adults above 15 years</td>
<td></td>
</tr>
<tr>
<td>Firm size</td>
<td>The number of workers employed in current place of work, 5 categories</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Size unknown</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Size 1: 1-9 workers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Size 2: 10-49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Size 3: 50-499</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Size 4: 500+</td>
<td></td>
</tr>
<tr>
<td>Region</td>
<td>Statistical Divisions (SD), 55 categories</td>
<td></td>
</tr>
<tr>
<td>Occupation</td>
<td>Two-digit occupation, 37 categories</td>
<td></td>
</tr>
<tr>
<td>Industry</td>
<td>Two-digit industry, 53 categories</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------</td>
<td>-----------</td>
</tr>
<tr>
<td>(log) Hourly wage</td>
<td>3.002</td>
<td>0.421</td>
</tr>
<tr>
<td>(log) Hourly wage 1st lag</td>
<td>2.970</td>
<td>0.421</td>
</tr>
<tr>
<td>(log) Hourly wage 2nd lag</td>
<td>2.936</td>
<td>0.421</td>
</tr>
<tr>
<td>(log) Hourly wage 3rd lag</td>
<td>2.907</td>
<td>0.428</td>
</tr>
<tr>
<td>(log) Weekly working hours</td>
<td>3.620</td>
<td>0.429</td>
</tr>
<tr>
<td>(log) Distance</td>
<td>0.157</td>
<td>0.966</td>
</tr>
<tr>
<td>Education</td>
<td>12.64</td>
<td>2.131</td>
</tr>
<tr>
<td>Tenure</td>
<td>10.77</td>
<td>9.541</td>
</tr>
<tr>
<td>Age</td>
<td>41.55</td>
<td>10.39</td>
</tr>
<tr>
<td>Sex</td>
<td>0.579</td>
<td>0.494</td>
</tr>
<tr>
<td>Marital status</td>
<td>0.717</td>
<td>0.451</td>
</tr>
<tr>
<td>No. of children</td>
<td>0.627</td>
<td>0.944</td>
</tr>
<tr>
<td>Firm size unknown</td>
<td>0.004</td>
<td>0.066</td>
</tr>
<tr>
<td>Firm size 1-9</td>
<td>0.212</td>
<td>0.409</td>
</tr>
<tr>
<td>Firm size 10-49</td>
<td>0.303</td>
<td>0.459</td>
</tr>
</tbody>
</table>
Firm size 50-499  |  0.345  |  0.475  |  0  |  1
Firm size 500+ |  0.135  |  0.342  |  0  |  1

* Number of observations is 11846. All statistics are weighted by individual longitudinal weights for the balanced panel.

Table A2.3. Wage equation, system GMM results

<table>
<thead>
<tr>
<th></th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wage, 1st lag</td>
<td>0.309</td>
<td>0.307</td>
<td>0.309</td>
<td>0.307</td>
<td>0.307</td>
<td>0.314</td>
</tr>
<tr>
<td>(LWage)</td>
<td>[0.026]<em><strong>[0.025]</strong></em>[0.025]<em><strong>[0.024]</strong></em>[0.023]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage, 2nd lag</td>
<td>0.113</td>
<td>0.111</td>
<td>0.113</td>
<td>0.110</td>
<td>0.109</td>
<td>0.108</td>
</tr>
<tr>
<td></td>
<td>[0.021]<em><strong>[0.020]</strong></em>[0.020]<em><strong>[0.019]</strong></em>[0.018]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage, 3rd lag</td>
<td>0.086</td>
<td>0.085</td>
<td>0.087</td>
<td>0.081</td>
<td>0.084</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>[0.018]<em><strong>[0.018]</strong></em>[0.017]<em><strong>[0.017]</strong></em>[0.016]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance</td>
<td>0.005</td>
<td>0.004</td>
<td>0.002</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.008] [0.003] [0.003] [0.003]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>0.023</td>
<td>0.024</td>
<td>0.024</td>
<td>0.024</td>
<td>0.024</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>[0.008]<em><strong>[0.003]</strong></em>[0.003]<em><strong>[0.003]</strong></em>[0.003]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance*mcEducation</td>
<td>0.002</td>
<td>0.003</td>
<td>0.003</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.001]* [0.002]* [0.002]* [0.001]**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance*mcLWage</td>
<td>-0.008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.009]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance*Q1</td>
<td>0.019</td>
<td>0.023</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.007]<em><strong>[0.007]</strong></em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance*Q2</td>
<td>-0.008</td>
<td>-0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.006] [0.005]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Coefficient</td>
<td>Standard Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------------</td>
<td>----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance*Q3</td>
<td>-0.003</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.008</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance*Q4</td>
<td>-0.003</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.008</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance<em>Q1</em>mcLWage</td>
<td>-0.082</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.022]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance<em>Q2</em>mcLWage</td>
<td>0.064</td>
<td>[0.064]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance<em>Q3</em>mcLWage</td>
<td>-0.010</td>
<td>[0.013]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance<em>Q4</em>mcLWage</td>
<td>-0.005</td>
<td>[0.013]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hours of work</td>
<td>-0.620</td>
<td>[0.044]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.610</td>
<td>[0.042]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.610</td>
<td>[0.042]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.609</td>
<td>[0.041]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.612</td>
<td>[0.041]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.595</td>
<td>[0.040]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tenure</td>
<td>0.004</td>
<td>[0.001]**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.004</td>
<td>[0.001]**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.004</td>
<td>[0.001]**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.004</td>
<td>[0.001]**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.003</td>
<td>[0.001]**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Tenure^2)/100</td>
<td>-0.004</td>
<td>[0.006]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.005</td>
<td>[0.003]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.005</td>
<td>[0.003]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.005</td>
<td>[0.003]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.005</td>
<td>[0.003]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.009</td>
<td>[0.009]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.007</td>
<td>[0.007]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.007</td>
<td>[0.007]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.008</td>
<td>[0.008]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.008</td>
<td>[0.008]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Age^2)/100</td>
<td>-0.013</td>
<td>[0.011]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.011</td>
<td>[0.011]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.011</td>
<td>[0.011]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.011</td>
<td>[0.011]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>0.158</td>
<td>[0.036]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.164</td>
<td>[0.015]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.164</td>
<td>[0.015]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.165</td>
<td>[0.014]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.164</td>
<td>[0.014]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.159</td>
<td>[0.014]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital status</td>
<td>0.018</td>
<td>[0.026]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.013</td>
<td>[0.011]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.013</td>
<td>[0.010]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.012</td>
<td>[0.010]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.012</td>
<td>[0.010]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.012</td>
<td>[0.010]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.012</td>
<td>[0.010]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M1</td>
<td>M2</td>
<td>M3</td>
<td>M4</td>
<td>M5</td>
<td>M6</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>No. of children</td>
<td>-0.017</td>
<td>-0.014</td>
<td>-0.013</td>
<td>-0.013</td>
<td>-0.013</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>[0.015]</td>
<td>[0.007]**</td>
<td>[0.007]**</td>
<td>[0.006]**</td>
<td>[0.006]**</td>
<td>[0.006]*</td>
</tr>
<tr>
<td>Household size</td>
<td>-0.0002</td>
<td>-0.002</td>
<td>-0.003</td>
<td>-0.002</td>
<td>-0.002</td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td>[0.010]</td>
<td>[0.005]</td>
<td>[0.005]</td>
<td>[0.005]</td>
<td>[0.005]</td>
<td>[0.004]</td>
</tr>
<tr>
<td>Firm size unknown</td>
<td>-0.096</td>
<td>-0.085</td>
<td>-0.086</td>
<td>-0.086</td>
<td>-0.080</td>
<td>-0.087</td>
</tr>
<tr>
<td></td>
<td>[0.059]</td>
<td>[0.058]</td>
<td>[0.057]</td>
<td>[0.057]</td>
<td>[0.056]</td>
<td>[0.053]</td>
</tr>
<tr>
<td>Firm size 1-9</td>
<td>-0.158</td>
<td>-0.149</td>
<td>-0.147</td>
<td>-0.148</td>
<td>-0.151</td>
<td>-0.146</td>
</tr>
<tr>
<td></td>
<td>[0.042]<em><strong>[0.016]</strong></em>[0.016]<em><strong>[0.015]</strong></em>[0.015]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm size 10-49</td>
<td>-0.074</td>
<td>-0.071</td>
<td>-0.070</td>
<td>-0.070</td>
<td>-0.071</td>
<td>-0.068</td>
</tr>
<tr>
<td></td>
<td>[0.016]<em><strong>[0.012]</strong></em>[0.012]<em><strong>[0.012]</strong></em>[0.012]***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm size 50-499</td>
<td>-0.025</td>
<td>-0.023</td>
<td>-0.022</td>
<td>-0.022</td>
<td>-0.023</td>
<td>-0.021</td>
</tr>
<tr>
<td></td>
<td>[0.012]<strong>[0.011]</strong>[0.010]<strong>[0.010]</strong>[0.010]**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm size 500+</td>
<td>base</td>
<td>base</td>
<td>base</td>
<td>base</td>
<td>base</td>
<td>base</td>
</tr>
<tr>
<td>Constant</td>
<td>3.529</td>
<td>3.511</td>
<td>3.503</td>
<td>3.537</td>
<td>3.569</td>
<td>3.495</td>
</tr>
<tr>
<td></td>
<td>[0.219]<em><strong>[0.216]</strong></em>[0.216]<em><strong>[0.214]</strong></em>[0.210]<em><strong>[0.206]</strong></em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-Bond A(1) test, p-value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>A-Bond A(2) test, p-value</td>
<td>0.168</td>
<td>0.160</td>
<td>0.158</td>
<td>0.175</td>
<td>0.143</td>
<td>0.123</td>
</tr>
<tr>
<td>Hansen test, p-value</td>
<td>0.206</td>
<td>0.379</td>
<td>0.524</td>
<td>0.335</td>
<td>0.668</td>
<td>0.629</td>
</tr>
<tr>
<td>Hansen test excluding group, p-value</td>
<td>0.234</td>
<td>0.221</td>
<td>0.336</td>
<td>0.253</td>
<td>0.316</td>
<td>0.474</td>
</tr>
<tr>
<td>Difference (null H = exogenous), p-value</td>
<td>0.274</td>
<td>0.545</td>
<td>0.619</td>
<td>0.452</td>
<td>0.805</td>
<td>0.651</td>
</tr>
</tbody>
</table>

(a) Number of observations is 11846, number of persons is 3793.
(b) Standard errors in brackets. * significant at 10%; ** at 5%; *** at 1%.
(c) P-values are reported for diagnostics tests: A-B is Arellano-Bond test for AR(1) or AR(2) in first differences; the last two are Difference-in-Hansen tests of exogeneity of GMM instruments for levels.
(d) Coefficients received for dummies for year, region, occupation, and industry are not reported.
(e) mc in the first column indicates the variable is mean centred.
Proof of Proposition 1.

Intuitively, a rich agent has a larger set of affordable destinations and therefore, at least the same destination as a poor or at best another destination not reachable by the poor, which provides the rich with higher returns. Proof can be seen in the special case of differentiable wage and costs functions, $W(x, y)$ and $C(x, y)$, both are defined on an everywhere dense set $X$. The Karush–Kuhn–Tucker conditions for the optimal solution $y^*$ of the maximization problem are:

$$\frac{\partial \pi}{\partial y} = \mu \cdot \frac{\partial C(x, y)}{\partial y}$$

stationarity condition

$C(x, y) \leq B$

primal feasibility

$\mu \geq 0$

dual feasibility

$\mu \cdot (B - C(x, y)) = 0$

complementary slackness

These conditions are necessary for a local maximum. They are also sufficient if profit and costs are concave functions. We assume these conditions hold at least in the neighbourhood of the maximum. First, define the Lagrange function $\Lambda(x, y, B, \mu) = \pi(x, y) + \mu \cdot (B - C(x, y))$

Then, applying the envelope theorem with respect to parameter $B$ and taking into account the dual feasibility condition we derive the positive sign of the derivative of the optimal profit

$$\frac{\partial \pi^*(x, B)}{\partial B} = \frac{\partial \Lambda(x, y, B, \mu)}{\partial B} \bigg|_{y = y^*(x, B)} = \mu, \text{ where } \mu \geq 0.$$ 

In the general case without differentiability assumption one can apply a proof by contradiction. Take two agents out of which the first has higher budget $B_1 > B_2$. Assume on the contrary $\pi^*(x, B_1) < \pi^*(x, B_2)$. Consider the two optimal points $y^*(x, B_1)$ and $y^*(x, B_2)$. They are different, since otherwise the profits are equal. Moreover, inequality $\pi^*(x, B_1) < \pi^*(x, B_2)$ means the point $y^*(x, B_2)$, for which $C(x, y^*(x, B_2)) \leq B_2$, is not feasible for the first agent, that is, $C(x, y^*(x, B_2)) > B_1$. These two inequalities contradict the assumption the first has higher budget than another agent, $B_1 > B_2$.

Q.E.D.
Proof of Proposition 2.

As in the proof of Proposition 1, first, this property can be shown for differentiable wage and costs functions. The envelope theorem together with the dual feasibility and the monotonicity of the costs function with respect to $\theta$ lead to:

$$\frac{\partial \pi^*(x, B, \theta)}{\partial \theta} = -\mu \cdot \frac{\partial C(x, y^*, \theta)}{\partial \theta} \leq 0$$

Also, one can provide here a proof by contradiction in the general case. Assume $\theta_1 < \theta_2$, what implies $C(x, y, \theta_1) < C(x, y, \theta_2)$ for any unequal $x$ and $y$, and the contrary to what we are eager to show

$$\pi^*(x, B, \theta_1) < \pi^*(x, B, \theta_2)$$

Then the two optimal points are $y^*(x, B, \theta_1)$ and $y^*(x, B, \theta_2)$. It follows from the assumption that point $y^*(x, B, \theta_2)$ is not feasible for the first agent, $C(x, y^*(x, B, \theta_2), \theta_1) > B$, what contradicts other two conditions, $C(x, y^*(x, B, \theta_2), \theta_2) \leq B$ and $C(x, y^*(x, B, \theta_2), \theta_1) < C(x, y^*(x, B, \theta_2), \theta_2)$. 

Q.E.D.
Proof of Proposition 3.

Assume differentiability of the optimal profit with respect to distance and internal global maximum. The global or a local maximum is an internal solution of the profit maximization problem when budget constraint is non-binding at the optimal point, \( C(x, y^*, \theta) < B \). On the contrary, a profit maximizing point which is neither the global no a local maximum, is a boundary solution of the optimization with binding budget constraint, \( C(x, y, \theta) = B \), and with a positive Lagrange multiplier \( \mu \) in the Lagrange function \( \Lambda(x, y, B, \theta, \mu) = \pi(x, y) + \mu \cdot (B - C(x, y, \theta)) \).

Differentiating the Lagrange function with respect to distance, one finds in optimum \( y^* \):

\[
\frac{\partial \Lambda(x, y, B, \theta, \mu)}{\partial \text{DIST}} = \frac{\partial \pi(x, y)}{\partial \text{DIST}} - \mu \cdot \frac{\partial C(x, y, \theta)}{\partial \text{DIST}} = 0
\]

Then for the global and the local maximum complementary slackness condition, \( \mu \cdot (B - C(x, y^*)) = 0 \), implies \( \mu = 0 \). Therefore, from stationarity condition

\[
\frac{\partial \pi(x, y^*)}{\partial \text{DIST}} = \mu \cdot \frac{\partial C(x, y^*)}{\partial \text{DIST}} = 0
\]

in the global / local maximum. Contrariwise, in other optimums stationarity condition is:

\[
\frac{\partial \pi(x, y^*)}{\partial \text{DIST}} = \mu \cdot \frac{\partial C(x, y^*)}{\partial \text{DIST}} = \mu \cdot MC > 0.
\]

Q.E.D.
4 Theoretical models: destination choice and search

4.1 Introduction

In this chapter, migration is modelled as an outcome of two principally different approaches, namely the fixed sample destination search and the sequential destination search. The first is known as the fixed-sample-size (FSS) approach as proposed by Stigler (1962). Time is not an important factor in this model. An agent can patiently wait until the end of the period when their decision is made. During this period the agent collects all necessary information about \( N \) destinations and assigns a number, which is the post-migration utility level, to each of them. The number of destinations, \( N \), is modelled to be an exogenous number in the beginning and an endogenous number at the end of this chapter. It is assumed the searcher must make a choice of a single destination. The selected destination can coincide with the origin in which case there will be no migration. For exogenously given \( N \) an agent assesses utility in all \( N \) destinations and selects the optimal destination with net utility above the reservation utility \( u_0 \). Utility can be understood as the net present value of a stream of utilities associated with a given location. Utility function is assumed to be an additive function of three variables, utility at current location, migration costs, and information costs. Net utility in destination \( i \) is defined as utility in this destination \( u_i \) less the costs of migration \( c_i \) to the destination \( i \) and costs of information collection and analysis \( c(N) \). A block diagram below demonstrates a sequence of the agent’s decisions and payoffs.
Diagram 5. Block diagram of decisions and payoffs, N destinations

- Migration
  - Migration to 1
  - ...
  - Migration to N

- No migration
The second approach is based on the sequential search. A searcher is restricted to receiving only one observation each period paying for it with the costs \(c\). Before the next period starts they decide whether to continue the search or stop it. This decision depends on whether the expected gain of continuing to search is above the costs of searching or, equivalently, whether the current offer exceeds their reservation value (Lippman and McCall 1976). The calculation of the expected value of searching depends on a reservation value, below which an offer is rejected. If costs of the search in the first period, \(c(1)\), are greater than the expected gain from this search, then the optimal decision is not to start a search. The search is finished because the expected value of searching in the next period is not greater than for the current period due to non-increasing reservation value in the course of time. Increasing reservation value is not optimal because the present value of the expected gain will be higher if the current reservation value is raised to the level of the next period reservation wage. If it is worthwhile to search in the first period then the search can finish after random periods of time when the agent is lucky enough to get an offer above the reservation value in that period. For example, an offer in the second period above the reservation value makes a search in the third period unattractive due to higher opportunity costs. This approach was used in the consumer search literature, e.g. searching for the lowest price in Rothschild (1974) and Kohn and Shavell (1974). Also it was adopted by economists for the job search problem (Lippman and McCall 1976). This method can be represented by a different block diagram, where the expected gain will be defined by a sequence in section 4.3.
Diagram 6. Block diagram of decisions, sequential search
4.2 Model of destination choice

4.2.1 Exogenous \( N \): no uncertainty

Theoretical model presentation starts from the simplest model set-up with no uncertainty. Costs of information search are an increasing function of the number of destinations with no fixed costs:

\[
c(N), \ c'(N) > 0, \ c(0) = 0
\]

No assumption about concavity/convexity of the information costs function is made. The only assumption here is that the information costs are affordable for any agent; that is they are below or equal to the available budget. By backward induction an agent is able to define, first, post-migration utility, which is at maximum across destinations:

\[
UM \equiv \max (u_1 - c_1, \ldots, u_N - c_N) - c(N)
\]

Second, because there is an option not to search, the utility after searching is a maximum of two numbers:

\[
US \equiv \max (u_0, UM) \equiv \max (u_0, u_1 - c_1 - c(N), \ldots, u_N - c_N - c(N))
\]

Since all information is available and there is no uncertainty about utilities and costs across destinations in this simple set-up, the solution of a utility maximization problem is deterministic. This is the simplest maximization problem in which the agent finds the optimal destination identifying the maximum number out of \( N + 1 \) options.

4.2.2 Exogenous \( N \): uncertainty

The number of destinations \( N \) is assumed to be given, in other words \( N \) is exogenous. A destination utility is randomly drawn from a continuum set of real numbers \(-\infty \leq x \leq \infty\). Uncertainty implies
that with some probability after search and even post-migration utility will be below an initial level. In the worst case without migration utility is $u_0 - c(N)$. However, in the case of good luck, utility exceeds the initial level, the most desirable outcome of search for the migrant. Assume that at each destination utility $U$ is random with c.d.f. $F_U(x)$, migration costs associated with this destination are random variable $C$ with c.d.f. $F_C(x)$, and that both random variables are independently distributed.

Now, define a new random variable for maximum net utility on the sample of the fixed size $N$:

$$\Psi = \max_{i=1,...,N} (u_i - c_i)$$

where $u_i$ and $c_i$ are realizations of the random variables $U$ and $C$ at destination $i$. Then the c.d.f. of the function $\Psi$ is $F_N(\cdot)$, where $F(\cdot)$ is the c.d.f. of the random variable $U - C$ which is a gross benefit to migration. Therefore after the fixed-size sampling migration decision is done with probability:

$$PM(u_0, N) = P\{\Psi > u_0 - c(N)\} = 1 - P\{\Psi < u_0 - c(N)\}$$

$$= 1 - P\{u_1 - c_1 < u_0 - c(N), ..., u_N - c_N < u_0 - c(N)\}$$

$$= 1 - \prod_{i=1}^{N} P\{u_i - c_i < u_0 - c(N)\} = 1 - F_N(u_0 - c(N))$$

**Proposition 4.** The probability of migration is increasing with costs of searching $c(N)$ and number of destinations $N$ and decreasing with initial utility $u_0$.

Proof. It is sufficient to find a sign of the respective derivative:
\[
\frac{\partial \text{PM}(u_0, N, c(N))}{\partial c(N)} = N\Phi^{-1}(u_0 - c(N)) \rho(u_0 - c(N)) > 0. \text{ Also,}
\]
\[
\frac{\partial \text{PM}(u_0, N)}{\partial N} = -F(N)(u_0 - c(N)) \left( \ln \left( F(u_0 - c(N)) \right) - \frac{Nc'(N) \rho(u_0 - c(N))}{F(u_0 - c(N))} \right) > 0
\]

since both terms in the expression in brackets are negative. Finally,
\[
\frac{\partial \text{PM}(u_0, N)}{\partial u_0} = -N\Phi^{-1}(u_0 - c(N)) \rho(u_0 - c(N)) < 0
\]

Q.E.D.

In other words, this proposition tells that the more that is spent on search and the more destinations are sampled the higher the probability of migration is. Though it is practically impossible to find the derivative with respect to the mean utility and mean costs of migration in the general case, one can do it in some special cases such as the normal distribution.

For the following proposition it is assumed that utility \( U \) and cost of migration \( C \) are independent normally distributed random variables on the infinite interval \((-\infty, \infty)\), \( U \sim N(\bar{u}, \sigma_u^2) \) and \( C \sim N(\bar{c}, \sigma_c^2) \). Then their difference is a random variable with normal distribution \( U - C \sim N(\bar{u} - \bar{c}, \sigma_u^2 + \sigma_c^2) \) and the following c.d.f. \( F(x) \):
\[
F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} \exp \left( -\frac{(y - \bar{u} + \bar{c})^2}{2\sigma^2} \right) dy
\]

where \( \sigma^2 = \sigma_u^2 + \sigma_c^2 \).
Proposition 5. For a normally distributed benefit the probability of migration is increasing with the mean utility $\bar{u}$, decreasing with the mean costs $\bar{c}$, and non-monotonic with both the variances of utility and costs of migration, $\sigma_U$ and $\sigma_C$.

Proof. For normal distribution it is possible to calculate respective two derivatives and to determine their signs:

$$\frac{\partial PM(u_0, N)}{\partial u} = -NF^{-1}(u_0 - c(N)) \frac{\partial F(u_0 - c(N))}{\partial u} = NF^{-1}(u_0 - c(N)) \rho(u_0 - c(N)) > 0$$

$$\frac{\partial PM(u_0, N)}{\partial c} = -NF^{-1}(u_0 - c(N)) \frac{\partial F(u_0 - c(N))}{\partial c} = -NF^{-1}(u_0 - c(N)) \rho(u_0 - c(N)) < 0$$

The two derivatives of the c.d.f. $F(u)$ were separately calculated here:

$$\frac{\partial F(u)}{\partial u} = -\rho(u)$$
$$\frac{\partial F(u)}{\partial c} = \rho(u)$$

For example, the first of these two derivatives:

$$\frac{\partial F(u)}{\partial u} = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y - \bar{u} + \bar{c})^2}{2\sigma^2}\right) dy = -\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y - \bar{u} + \bar{c})^2}{2\sigma^2}\right)_{-\infty}^{u} = -\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(u - \bar{u} + \bar{c})^2}{2\sigma^2}\right) = -\rho(u) < 0$$
Another derivative has the same absolute value but the opposite sign. This can be noted if one
substitutes $\bar{c}$ for $\bar{u}$ in the latest derivation formulas.

Non-monotonicity of the probability of migration with respect to a variance follows from the
changing sign of two derivatives, $\frac{\partial F(u)}{\partial \bar{\sigma}_U}$ and $\frac{\partial F(u)}{\partial \bar{\sigma}_c}$, shown in Proposition 11 below.

Q.E.D.

Because in general the agent has to decide whether to invest in a migration destination search, the
expected utility gain needs to be calculated. Before that one can extend formula (13) and derive a
formula for the expected utility:

$$EU = E \max \left( u_0 - c(N), \max (u_1 - c_1, \ldots, u_N - c_N) \right)$$

$$= u_0 - c(N) + E \max (0, \Psi - u_0 + c(N))$$

$$= u_0 - c(N) + E \left( \Psi - u_0 + c(N) \right) I \{ \Psi > u_0 - c(N) \} + 0 I \{ \Psi < u_0 - c(N) \}$$

$$= u_0 - c(N) + \int_{u_0 - c(N)}^{\infty} (u - u_0 + c(N)) dF^N(u)$$

The expected utility gain of searching is:

$$EUG(u_0, N) \equiv EU - u_0 = -c(N) + \int_{u_0 - c(N)}^{\infty} (u - u_0 + c(N)) dF^N(u)$$  \hspace{1cm} (15)

The integral in this formula has a positive value. Alternatively, applying the integration by parts to
(15) one finds the equivalent formula:

$$EUG(u_0, N) = -c(N) + \int_{u_0 - c(N)}^{\infty} (1 - F^N(u)) du$$  \hspace{1cm} (16)
The second item in this formula, the integrated probability of migration, is the expected benefit/increase in gross utility from migration. A condition for an agent to invest in a search is that the expected utility gain is non-negative or, equivalently, the expected benefit must be at least the costs of search:

\[
\int_{u_0-c(N)}^{\infty} \left(1 - F^N(u)\right) du \geq c(N)
\]

As can be checked from (16), there is no gain with no search; that is with zero destinations, \( N = 0 \):

\[
EUG(u_0, 0) = -c(0) = 0.
\]

When the number of destinations is one, \( N = 1 \), (15) implies that:

\[
EUG(u_0, 1) = -c(1) + \int_{u_0-c(1)}^{\infty} \left(u - u_0 + c(1)\right) \rho(u) du
\]

The integral in the expression can be explicitly found in a special case with normal distributions defined in (14). Since for any \( x \)

\[
\int_{x}^{\infty} \left(u - u_0 + c(1)\right) \rho(u) du = \int_{x}^{\infty} \left(u - m - m - u_0 + c(1)\right) \rho(u) du = \sigma^2 \rho(x) + (m - u_0 + c(1))(1 - F(x))
\]

the expected utility with a single destination in this special case is a function of the initial utility, average utility, dispersion of utility, and costs of one destination search:

\[
EUG(u_0, 1) = -c(1) + \sigma^2 \rho(u_0 - c(1)) + (m - u_0 + c(1))(1 - F(u_0 - c(1)))
\]
This function is positive for a sufficiently small initial utility and negative for a sufficiently large one:

$$\lim_{u_0 \to -\infty} EUG(u_0, 1) = \infty$$

and

$$\lim_{u_0 \to \infty} EUG(u_0, 1) \leq -c(1).$$

Moreover, it is a monotonic function in this special case since its derivative is negative:

$$\frac{dEUG(u_0, 1)}{du_0} = \sigma^2 \rho'(u_0 - c(1)) - (m - u_0 + c(1)) \rho(u_0 - c(1)) - (1 - F(u_0 - c(1)))$$

$$= -\left(1 - F(u_0 - c(1))\right) < 0$$

Here the derivative of the density of normal distribution is:

$$\rho'(x) = -\frac{x - m}{\sigma^2} \rho(x).$$

It is possible to generalise these observations in the following two propositions.

**Proposition 6.** The expected gain is a decreasing function of the initial utility and costs of search.

Proof. It is straightforward to show that derivatives with respect to initial utility and costs of search are negative by differentiation of formula (16):

$$\frac{\partial EUG(u_0, N)}{\partial u_0} = -\left(1 - F^N(u_0 - c(N))\right) < 0$$
and after noting that the expected utility gain depends negatively on how much an agent pays for the search of $N$ destinations:

$$\frac{\partial EUG\left(u_0, N, c(N)\right)}{\partial c(N)} = -F^N\left(u_0 - c(N)\right) < 0$$

Q.E.D.

**Proposition 7.** The expected utility gain is a non-monotonic function of the number of destinations $N$.

Proof. Indeed, the derivative of the expected utility gain with respect to $N$ consists of two summands:

$$\frac{\partial EUG\left(u_0, N\right)}{\partial N} = -c'(N)F^N\left(u_0 - c(N)\right) - \int_{u_0-c(N)}^{\infty} F^N(u)\ln\left(F(u)\right)du$$

The sign of this expression is indeterminate as the first item is negative and the second one is positive. Therefore the sum can be either negative or positive depending on the initial utility, the number of destinations and that other two functions in the expression, the costs function and c.d.f. .

Q.E.D.

This link will be investigated further in section 4.2.4 with endogenous $N$.

**Proposition 8.** For a normally distributed benefit to migration the expected utility gain is an increasing function of the mean utility and both variances, $\sigma_U$ and $\sigma_C$, and a decreasing function of the mean cost of migration.
Proof. It is enough to find two derivatives of the expression (16):

\[
\frac{\partial EUG(u_0, N)}{\partial u} = - \int_{u_0 - c(N)}^{\infty} \frac{\partial F^N(u)}{\partial u} \, du = - \int_{u_0 - c(N)}^{\infty} NF^{N-1}(u) \frac{\partial F(u)}{\partial u} \, du = \int_{u_0 - c(N)}^{\infty} NF^{N-1}(u) \rho(u) \, du > 0
\]

and

\[
\frac{\partial EUG(u_0, N)}{\partial c} = - \int_{u_0 - c(N)}^{\infty} \frac{\partial F^N(u)}{\partial c} \, du = - \int_{u_0 - c(N)}^{\infty} NF^{N-1}(u) \frac{\partial F(u)}{\partial c} \, du = - \int_{u_0 - c(N)}^{\infty} NF^{N-1}(u) \rho(u) \, du < 0
\]

The two derivatives of the c.d.f. were already calculated in Proposition 5.

The negative sign of the derivatives with respect to variance,

\[
\text{e.g.} \quad \frac{\partial EUG(u_0, N)}{\partial \sigma_U} = \int_{u_0 - c(N)}^{\infty} NF^{N-1}(u) \frac{\partial F(u)}{\partial \sigma_U} \, du < 0
\]

follows from Proposition 11.

Q.E.D.

4.2.3 Comparative statics

In this section the comparative statics, that is the effect of parameter change on migration behaviour, is studied.
Definition 1. An agent is said to be indifferent between search and no search if the expected gain is zero:

\[
\int_{u_0 - c(N)}^{\infty} (1 - F^N(u)) du = c(N)
\]

Proposition 9. In order to guarantee the existence of the indifferent agent for a given \(N\) it is sufficient to assume that the integral in (17) is converging; that is for any \(a > 0\):

\[
\int_{a}^{\infty} (1 - F^N(u)) du < \infty
\]

Proof. According to the definition of integral convergence, for any number \(\varepsilon > 0\) there is a number \(b(\varepsilon)\) such that

\[
\int_{b}^{\infty} (1 - F^N(u)) du < \varepsilon \text{ for any } b \geq b(\varepsilon)
\]

It follows from this assumption that for given \(N\) and \(c(N)\) there is an initial utility such that the search is not beneficial. Indeed, let \(\varepsilon = c(N)\) then there is \(b_1\) such that

\[
\int_{b_1}^{\infty} (1 - F^N(u)) du < \varepsilon.
\]

Therefore, this initial utility is \(u_0 = b_1 + c(N)\).

One can find an agent who gains from the investment in search. The following integral

\[
\int_{-\infty}^{\infty} (1 - F^N(u)) du = \infty
\]

diverges due to the monotonicity of the c.d.f. \(F^N(\cdot)\) and due to a property that the integrated function tends to 1 at negative infinity, \(\lim_{u \to -\infty} (1 - F^N(u)) = 1\). Therefore, for any
\( \varepsilon > 0 \), in particular, \( \varepsilon = c(N) \), there is a number \( b_2 \) such that \( \int_{b_2}^{\infty} (1 - F^N(u)) du > \varepsilon \). Then the initial utility of an agent who gains from the search is \( u_0 = b_2 + c(N) \).

As a result, the existence of the indifferent agent follows from the monotonicity of the expected utility gain, which is the LHS minus the RHS in (17).

Q.E.D.

Condition (18) imposes restrictions from above on the growth of the c.d.f. \( F(\cdot) \). The c.d.f. of a normal distribution is an example of such a function. An example of a function which does not satisfy this condition is:

\[
F(x) = 1 - \frac{1}{1+x}, \quad 0 \leq x \leq \infty, \quad \text{with density} \quad \rho(x) = \frac{1}{(1+x)^2}.
\]

All agents with an initial utility below that of the indifferent agent are migrants. A reaction of the indifferent agent, which can also be called the marginal agent, to a change in any parameter of the model gives the direction of the relationship between the total migration rate and the parameter. For the two propositions below there will be an assumption that benefits to migration are normally distributed, that is they have the c.d.f. defined in (14).

**Proposition 10.** The marginal agent’s utility \( u_0 \) increases with the average utility \( \bar{u} \) and decreases with the average costs of migration \( \bar{c} \).

Proof. The implicit function theorem can be used to calculate two derivatives:
\[
\frac{du_0}{du} = -\frac{\int_{u_0}^{\infty} \frac{\partial F^N(u)}{\partial u} du}{1 - F^N(u_0)} > 0
\]

and

\[
\frac{du_0}{dc} = -\frac{\int_{u_0}^{\infty} \frac{\partial F^N(u)}{\partial c} du}{1 - F^N(u_0)} < 0
\]

where the signs of the nominators of these two fractions were calculated in the proof of Proposition 8.

Q.E.D.

As a check of formulas derived in the proof of the last proposition a special case \( N = 1 \) can be examined. Using the two above formulas for the derivatives one finds

\[
\frac{du_0}{du} = \frac{\int_{u_0}^{\infty} \frac{\partial F(u)}{\partial u} du}{1 - F(u_0)} = \frac{\int_{u_0}^{\infty} \rho(u) du}{1 - F(u_0)} = 1
\]

and

\[
\frac{du_0}{dc} = -\frac{\int_{u_0}^{\infty} \frac{\partial F(u)}{\partial c} du}{1 - F(u_0)} = -\frac{\int_{u_0}^{\infty} \rho(u) du}{1 - F(u_0)} = -1
\]

The only solution to the system of these two differential equations is the indifferent agent’s utility \( u_0 = \bar{u} - \bar{c} \).
Proposition 10 also implies the unsurprising result that there are more migrants if the average utility is higher and if the average cost of migration is lower. Finally, it will be proved that there are more migrants if there is more heterogeneity in the distributions of utility and costs; that is if either of these two variances is higher.

**Proposition 11.** The marginal agent’s utility \( u_0 \) increases with both the variances of utility and costs of migration, respectively \( \sigma_u \) and \( \sigma_c \).

Proof. As in the proofs of propositions above one needs to find the derivatives of the implicit function and show they have the positive sign:

\[
\frac{du_0}{d\sigma_u} = -\frac{\int_{u_0}^{\infty} \frac{\partial F^N(u)}{\partial \sigma_u} du}{1 - F^N(u_0)} = -\frac{\int_{u_0}^{\infty} NF^{N-1}(u) \frac{\partial F(u)}{\partial \sigma_u} du}{1 - F^N(u_0)} > 0
\]

In order to show the positive sign in the expression one needs to calculate, first, the derivative under the integral:

\[
\frac{\partial F(u)}{\partial \sigma_u} = u \frac{\partial \rho(y)}{\partial \sigma_u} dy = \frac{\sigma_u}{\sqrt{2\pi}} \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} \left( \frac{y - \bar{u} + \bar{c}}{\sigma_u^2 + \sigma_c^2} \right)^2 \exp\left( -\frac{(y - \bar{u} + \bar{c})^2}{2(\sigma_u^2 + \sigma_c^2)} \right) dy
\]

It is not difficult to see that this function decreases with respect to \( u \), changing sign from the positive to the negative in \( u = \bar{u} - \bar{c} \), and is antisymmetric with respect to this point; that is:

\[
\frac{\partial F(\bar{u} - \bar{c})}{\partial \sigma_u} = 0
\]

and
\[
\frac{\partial F(\bar{u} - \bar{c} - u)}{\partial \sigma_U} = -\frac{\partial F(\bar{u} - \bar{c} + u)}{\partial \sigma_U}
\]

This is due to symmetry of the integrated function (since both the pdf and second moment are symmetric), with a mirror point \( u = \bar{u} - \bar{c} \). In particular, the first formula:

\[
\frac{\partial F(\bar{u} - \bar{c})}{\partial \sigma_U} = \frac{\sigma_U}{\sqrt{2\pi(\sigma_U^2 + \sigma_C^2)^3}} \int_{-\infty}^{\bar{u} - \bar{c}} \exp\left(-\frac{(y - \bar{u} + \bar{c})^2}{2(\sigma_U^2 + \sigma_C^2)}\right) dy = 0
\]

because

\[
\int_{-\infty}^{\bar{u} - \bar{c}} \frac{(y - \bar{u} + \bar{c})^2}{\sigma_U^2 + \sigma_C^2} \exp\left(-\frac{(y - \bar{u} + \bar{c})^2}{2(\sigma_U^2 + \sigma_C^2)}\right) dy = \int_{-\infty}^{\bar{u} - \bar{c}} \exp\left(-\frac{(y - \bar{u} + \bar{c})^2}{2(\sigma_U^2 + \sigma_C^2)}\right) dy
\]

Finally, note that the symmetry implies that

\[
\int_{u_{\text{c}(N)}}^{\infty} \frac{\partial F(u)}{\partial \sigma_U} du < 0
\]

since the left tail of the distribution is cut off. Furthermore, after multiplication by the increasing function \( F^{N-1}(u) \) the sign is still negative:

\[
\int_{u_{\text{c}(N)}}^{\infty} F^{N-1}(u) \frac{\partial F(u)}{\partial \sigma_U} du < 0.
\]

A similar calculation shows that the derivative with respect to another variance is also positive:

\[
\frac{du_0}{d\sigma_C} = -\frac{\int_{u_{\text{c}(N)}}^{\infty} \frac{\partial F^N(u)}{\partial \sigma_C} du}{1 - F^N(u_0)} = -\frac{\int_{u_{\text{c}(N)}}^{\infty} \frac{\partial F^N(u)}{\partial \sigma_C} du}{1 - F^N(u_0)} > 0
\]
due to the negative derivative:

\[ \frac{\partial F(u)}{\partial \sigma_c} < 0 \]

Q.E.D.

### 4.2.4 Endogenous information search

In section 4.2.5 below, the endogeneity of the information search will be shown to lead to the acquisition of information about the infinite number, or in general the maximum affordable number, of destinations by each agent no matter what their initial utility is. Only in that case individual expected gain is maximised. This is true if information is costless in terms of time and budget which is unlikely to be true. In this section not only are the costs of searching introduced but information is collected optimally. The exogeneity assumption is relaxed and the expected utility gain is maximised by \( N \). Since \( N \) is not given and there is uncertainty in this model then an agent needs to select the optimal number of destinations based on available information about distributions of utility and costs. In other words, the individual decides how much information to buy before the migration decision is made. As information arrival is modelled by a random process this decision is based on maximization of the expected gain.

Algebraically, each agent with reservation utility \( u_0 \) finds the optimal number of destinations to search, \( 0 \leq N \leq \bar{N} \), where \( \bar{N} \) is the maximum affordable number of destinations about which the agent can collect information. The agent’s budget constraint is not considered in this problem for analytic tractability.

The agent maximises their expected utility, which is determined by the initial utility, the number of destinations, cost function, and the two c.d.f. functions
\[
\max_{0 \leq N \leq \infty} EUG(u_0, N, C(\phi), F_U(\phi), F_C(\phi))
\]

The equivalent problem is maximization of the expected utility gain, given in (16):

\[
\max_{0 \leq N \leq \infty} EUG(u_0, N) = \max_{0 \leq N \leq \infty} \left( -c(N) + \int_{u_0 - c(N)}^{\infty} (1 - F^N(u)) du \right)
\]

Denote its solution by \( N^* \):

\[
N^* = \arg \max_{0 \leq N \leq \infty} EUG(u_0, N)
\]

Note the expected gain is zero if there is no search, \( N^* = 0 \). Therefore, the optimal expected gain of a migrant is at least zero. As was demonstrated in Proposition 7 it is not generally true that the expected gain is a monotonic function of the number of destinations, even though both the function under the integral and the cost function increase with \( N \). Such a deduction can give either a monotonic (increasing or decreasing function of \( N \)) or a non-monotonic function. The former case has a unique corner solution of the maximization problem, that is either \( N^* = 0 \) or \( N^* = N \), or both in some rare cases.

Below, a solution of the optimization problem (19) in integer numbers is found. The main result of this section, that only a minimum or maximum number of destinations is possible, will be derived. First of all, it should be noted that the discussion on possible internal solutions below uses rational numbers \( N^* \). This is done for simplicity but does not restrict the generality as one can see below in Proposition 12.

For an internal solution in real numbers, \( 0 < N^* < \bar{N} \), the following first and second order conditions are sufficient:
\[
\frac{\partial EUG(u_0, N)}{\partial N} = -c'(N) F^N(u_0 - c(N)) - \int_{u_0-c(N)}^{\infty} F^N(u) \ln(F(u)) \, du = 0
\] 20

SOC:

\[
\frac{\partial^2 EUG(u_0, N)}{\partial N^2} = -c''(N) F^N(u_0 - c(N))
\]

\[
-c'(N) F^{N-1}(u_0 - c(N))(F(u_0 - c(N)) \ln(F(u_0 - c(N))) - Nc'(N) \rho(u_0 - c(N)))
\]

\[
- \int_{u_0-c(N)}^{\infty} F^N(u) \ln(F(u)) \, du - \int_{u_0-c(N)}^{\infty} F^N(u) \ln^2(F(u)) \, du < 0
\] 21

**Proposition 12.** For a concave costs function, \( c''(x) < 0 \) for any \( x \), there is only a corner solution of the expected utility gain maximization problem, that is either \( N^* = 0 \) or \( N^* = N \) or both.

Proof. It is sufficient to prove that there is no internal solution satisfying conditions (20) and (21).

In order to show this one needs to prove that all internal extremums, that is solutions \( N^* \) of equation (20), are minimums. Note if there are no internal solutions \( N^* \) then there must be a border solution of the expected utility maximization problem.

Consider internal solutions of the FOC (20), \( 0 < N^* < N \). In these points the expected utility gain could be either minimum, or maximum, or neither. A sufficient condition for a minimum is that the expected utility gain is a convex function at these extremums, \( \frac{\partial^2 EUG(u_0, N^*)}{\partial N^2} > 0 \).

The last item in formula (21) can be bounded from below:
The integral on RHS of this inequality as well as in the penultimate item of formula (21) can be derived from the FOC. Hence,

\[- \int_{u_0-c(N)}^{\infty} F^N(u) \ln^2(F(u))du > -\ln(F(u_0-c(N))) \int_{u_0-c(N)}^{\infty} F^N(u) \ln(F(u))du\]

Therefore the second derivative at extremums can be bounded from below:

\[\frac{\partial^2 EUG(u_0, N^*)}{\partial N^2} > F^{N^*}(u_0 - c(N^*)) \left[ c'(N^*) - c''(N^*) + N^* c'(N^*) \right]^2 \rho(u_0 - c(N^*)) \] 22

The expression in the square brackets is positive for any \( u_0 \) when \( c'(N) \geq c''(N) \). This inequality always holds for a concave costs function since \( c''(N) < 0 \).

Note it is possible that the expected utility function is not convex in points other than extremums. Though this possibility is hard to prove, it is not important for the proof. What is important is that there are no internal solutions in real numbers. Therefore, the solution is one of two corners or both.

Q.E.D.

**Remark 1.** Concavity is a strict condition for a border solution because for a positive value in the square brackets in (22) it is sufficient to require that the costs function is not “very convex”; that is \( c'(N) \geq c''(N) \). For example, for the exponential costs function \( c(N) = \exp(N) - 1 \) the first and
second derivatives are equal. Hence, a border solution is ensured. However, for a slightly more convex function such as \( c(N) = \exp(\alpha N) - 1 \), where \( \alpha > 1 \), the second derivative exceeds the first one for any argument \( N \). It seems that for such a relatively convex function there will be an internal solution of problem (19). An example of this situation is however beyond the scope of this thesis.

**Corollary 1.** If the derivative of the expected utility function at zero is positive,

\[
\frac{\partial EUG(u_0, N)}{\partial N} \bigg|_{N=0} > 0,
\]

then the solution is the right corner, \( N^* = \overline{N} \).

Proof. The possibility of an internal solution is ruled out in Proposition 12. Therefore, for every real number \( N \), \( 0 \leq N \leq \overline{N} \), the derivative is everywhere positive. The maximum of the increasing function is in the right corner.

Q.E.D.

**Remark 2.** It follows from this corollary that there will be migrants in this model. This is because the expected utility gain at zero is positive for sufficiently small initial utilities:

\[
\frac{\partial EUG(u_0, N)}{\partial N} \bigg|_{N=0} = -c'(0) - \int_{u_0}^{\infty} \ln\left(F(u)\right) du > 0
\]

for all \( u_0 < \overline{u}_0 \), where \( \overline{u}_0 \) is a solution of the equation:

\[
\int_{u_0}^{\infty} \ln\left(F(u)\right) du = -c'(0)
\]
In one special case when the derivative of the costs function is equal to zero at zero, \( c'(0) = 0 \), then the derivative of the expected utility gain is positive at zero for any initial utility:

\[
\frac{\partial EUG(u_0, N)}{\partial N} \bigg|_{N=0} > 0 \text{ for } \forall u_0.
\]

Therefore, every agent in the model will invest in a search of destinations if their maximum number of destinations is at least one, \( \bar{N} \geq 1 \).

**Corollary 2.** For an agent with an initial utility below some level the optimal solution is the left corner, \( N^* = 0 \), for others it is the right corner, \( N^* = \bar{N} \).

Proof. It follows from (16), (17) and Proposition 9 that there is a unique sequence of indifferent agents with initial utilities \( \{\mu_0(n)\}_{n=1}^{\infty} \) such that the expected utility gain is zero, \( EUG(\mu_0(n), n) = 0 \). It is not clear whether this sequence is monotonic or not. However, whether the maximum of the expected utility gain is in the left corner or in the right corner depends on the relative location of the indifferent agent. In particular,

\[
N^* = \begin{cases} 
0 & \text{if } u_0 \geq \bar{u}_0(\bar{N}) \\
\bar{N} & \text{if } u_0 \leq \bar{u}_0(\bar{N})
\end{cases}
\]

This is because the expected utility gain is a decreasing function of the initial utility. Note, two optimal solutions are possible only for an indifferent agent.

Q.E.D.
Remark 3. It should be noted here that the initial utility of the indifferent agent in corollary 2 is greater than another initial utility, the solution of equation (23), $\bar{u}_0(N) > \bar{u}_0$. This is because the positive derivative of the expected utility gain at zero is not a necessary condition for maximum. In essence, the expected utility is non-monotonic when initial utility $u_0$ is between these two utilities, $u_0 < u_0 < \bar{u}_0(N)$.

Up to now Chapter 4 has presented two models with an exogenous as well as an endogenous number of destinations $N$. As an alternative approach, a search model will be considered in section 4.3. The following section presents a summary of examples based on a special case of the FSS migration model without search costs.

### 4.2.5 Examples: model of migration on circle

In this section a model of migration without information costs is considered. The goal is to develop a model which shows the relationship between information, migration, distance, and pre-migration and post-migration utility. The model presented here is close to the classical FSS problem by Stigler (1962). A modification of this model is done by means of randomly distributed income and costs of migration.

It is assumed that continuum agents are evenly located on a circle with a unit circumference. The circle is taken instead of the commonly used interval $[0,1]$ in order to allow similar migration opportunities for all agents; that is to exclude agents who live close to a border and therefore, have half the opportunities as compared to those residing in the centre. Utility is a random variable uniformly distributed on $[0,1]$. This contrasts with the model considered in the preceding section where normal distribution of utilities and costs was assumed. Unfortunately, for normal distribution

---

25 One may notice the resemblance to the distribution of Australian cities located along the coastal line.
it appears to be impossible to derive properties received in this section for simpler and less realistic uniform distribution.

The model considers an agent with utility $z \in [0,1]$, who can get information about a fixed number of random destinations, and then estimates the probability of migration. Using uniform distributions in the Stigler type FSS model of migration on the circle one can find the total migration rate, as well as the average distance moved by migrants and average pre-migration and post-migration utility. All correspondent formulas for two special cases, with one and two destinations, and then for two general cases, with partial and with full information about destinations, are derived in Appendix 3 at the end of this chapter.

In Table 22 below all results are summarized for different numbers of destinations, $N$, from one to infinity, including a general case of an arbitrary finite number. For the general case all variables other than the migration rate are expressed in terms of migration rate, $M_N$. Note that in order not to overload cells in the table some formulas contain the average distance, $\bar{\alpha}_N$, which is a function of $M_N$.

Note also that the case of “full” information does not necessarily require information about all, that is continuum, destinations (hence, the word full is placed in inverted commas). For properties to hold, it suffices to collect information for a countable subset, an infinite number of random destinations. The case of full information is not looking as unrealistic in the modern IT era as one might think. There is actually a need for a government agency or a private company which could provide access possibly to “projected” data, tailored to an individual’s needs.
### Table 22. Summary of theoretical results

<table>
<thead>
<tr>
<th>Number of destinations</th>
<th>Partial information</th>
<th>&quot;Full&quot; information</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>N</td>
</tr>
</tbody>
</table>

- **Migration rate, $M_N$**
  - $\frac{1}{6}$
  - $\frac{17}{60}$
  - $\sum_{k=1}^{N} \frac{(-1)^{k+1}C_k^k}{2^k(2k+1)}$ for $N \in \mathbb{N}$
  - 1

- **Average distance moved by migrants, $\bar{\alpha}_N$**
  - $\frac{1}{4}$
  - $\frac{4}{17}$
  - $\frac{1}{2} - \frac{1}{N+1} \left(1 - \frac{1}{2^{N+1}}\right)$ for $M_N$
  - 0

- **Average utility of stayers, $\bar{z}_s$**
  - $\frac{11}{20}$
  - $\frac{51}{86}$
  - $\frac{1}{2} + \bar{\alpha}_N \frac{M_N}{1-M_N}$ for $N \in \mathbb{N}$
  - 1

- **Average pre-migration utility of migrants, $\bar{z}_m$**
  - $\frac{1}{4}$
  - $\frac{9}{34}$
  - $\frac{1}{2} \bar{\alpha}_N$ for $N \in \mathbb{N}$
  - $\frac{1}{2}$

- **Average post-migration utility of migrants, $\bar{x}$**
  - $\frac{3}{4}$
  - $\frac{13}{17}$
  - $1 - \bar{\alpha}_N$ for $N \in \mathbb{N}$
  - 1

- **Average returns, $\bar{r} = \bar{x} - \bar{z}_m - \bar{\alpha}$**
  - $\frac{1}{4}$
  - $\frac{9}{34}$
  - $\frac{1}{2} \bar{\alpha}_N$ for $N \in \mathbb{N}$
  - $\frac{1}{2}$

One can conclude from this table that the migration rate increases with the amount of information agents are able to collect prior to migration. Also this is proved in Proposition A3.3 in Appendix 3. At the same time the more information, the lower the costs of migration for an average migrant and for migrants altogether since both the average distance and total distance moved are decreasing. The average post-migration utility of migrants is higher by the constant increment, $\frac{1}{2}$, than their average pre-migration utility, but a part of this utility gain has to be paid for relocation.

An increasing volume of information available to each agent results in a higher average utility of stayers and movers. This leads to the somewhat surprising conclusion that information induces
more migration but the effect is mostly seen not on the lower end of the utility distribution but among the relatively wealthier. This is due to a marginal agent who is generally not in either tail of the distribution.

4.3 Migration as search

Stigler (1961, 1962) is referred to as a pioneer in the economics of information and search. He is famous for his seminal paper on the FSS model of search. However, Hey (1981) mentions an article by Simon (1955), missed by many scholars, in which the sequential search model was first introduced. Fixed-sample-size studied in the preceding part is not a dominant approach in job search theory. The usual critique addresses the inefficiency of having a definite sample size (Morgan and Manning 1985). The opponents assert it is unlikely job offers come as a bundle. This is because of infrequent job offer arrival combined with the low duration of an offer. Under the sequential search approach, job offers are arriving stochastically, usually only one or zero per period, and generally subject to Poisson distribution (Mortensen 1986). It is optimal not to sample in bundles but to critically evaluate each arriving job offer and either to accept or reject that offer and continue to search. The sample size is, therefore a random variable determined by an optimal “stopping” rule for the migration destination (or job) search: stop searching when a destination (job offer) with utility (wage) above the reservation utility (wage) is found.

The optimality of this stopping rule is easy to understand. The optimal search should last until the cost of one extra search is equal to the expected gain from that search. Since the search is costly with a generally increasing marginal cost of the extra search and the expected marginal benefit of the extended search is generally declining and also because of time and financial constraints, the search is restricted in time.
The existence of the reservation wage is proved for the simplest search models. However, in more complex search models such as those with Bayesian leaning the reservation wage property may not hold (Rothschild 1974, Kohn and Shavell 1974, p.102). In models considered thereafter the reservation wages are either described or shown in the form of an implicit function where possible.

Before the model of search presentation is started, there is one remark regarding the link between the spatial job search and the migration destination search. There are similarities if the information regarding different destinations arrives randomly. For example, an individual may have some information about a utility distribution but for the exact value he/she needs to ask trustworthy people, say relatives or friends, currently living or who previously lived in that location. The set of destinations is restricted by the circle of his/her own and family acquaintance. If he/she is able to collect information all at once then the model is FSS which was considered in the preceding sections. However, if the information necessary for the decision making is arriving by parts in some random order then the model is similar to the search model. Because “offers” do not expire then the model is equivalent to the search model with recall. Below, models of search with and without recall and a number of their variations are analysed in the context of migration.

### 4.3.1 Model of search without recall

First, following the model presented in Lippman and McCall (1976) and defining $EG_T^t(x)$ to be the maximum expected gain from a sequential search with the optimal strategy and horizon $T$, where $x$ is a current utility draw and $t$ is a number of remaining periods, $1 \leq t \leq T \leq \infty$:

$$
EG_T^t(x) = \max \left( x, -c + \int_{-\infty}^{\infty} EG_{T-1}^t(u) dF(u) \right)
$$
The second term in the maximised expression is the expected gain from the optimal continuation of the search, denoted by $R_t^T$:

$$R_t^T = -c + \int_{-\infty}^{\infty} EG_{t-1}^T(u)dF(u)$$

where $F(\cdot)$ is the c.d.f. of the random variable for benefit to migration, in line with the assumption in section 4.2.2. Therefore the expected gain is determined sequentially:

$$R_t^T = -c + \int_{-\infty}^{\infty} \max(u, R_{t-1}^T)dF(u) \text{ for } t \geq 1, \text{ and } R_0^T = -\infty$$

The initial condition implies that the reservation utility is negative infinity, and if so, any offer received at the last period is accepted:

$$EG_0(x) = x$$

Therefore the reservation utility at period 1 is the expected utility, which is the mean utility minus the mean costs of migration $\bar{u} - \bar{c}$, minus the costs of searching:

$$R_1^T = -c + \int_{-\infty}^{\infty} udF(u) = \bar{u} - \bar{c} - c$$

The expected gain of searching at period 1 after offer $x$ is:

$$EG_1^T(x) = \max(x, R_1^T)$$

and so on for other periods:

$$EG_t^T(x) = \max(x, R_t^T), 1 \leq t \leq T$$
The sequence $R^T_t$ is a series of so-called reservation utilities which determines the optimal stopping rule:

**Stopping rule.** Observe $x$ and
\[
\begin{cases}
\text{stop search if } x \geq R^T_t \\
\text{continue if } x < R^T_t
\end{cases}.
\]

There are two possible cases:

(i) a finite number of periods $T$ before search must be terminated, $T < \infty$;

(ii) an infinite number of periods remaining, $T = \infty$.

These two cases are to be studied now

4.3.1.1 **Case of finite $T$, $T = \infty$**

**Proposition 13.** Reservation utility increases with remaining time, that is $R^T_T > R^T_{t-1}$.

Proof. It is also shown in a different, perhaps more accurate way in Lippman and McCall (1976). Not a big task to see this property from the sequential rule for the reservation utility (25) since in each additional period remaining until the termination of search, the function under the integral is increasing. The final period reservation utility is minimal, $R^T_0 = -\infty$, then some finite number $R^T_t = \bar{u} - \bar{c} - c$, etc. From formula (27) below one may see that the reservation utility is not infinite, otherwise the LHS in this formula is zero but RHS is equal to negative cost, $-c$.

Q.E.D.
This result is intuitively clear, the fewer periods remaining for the search the less gain of search is expected, other things being equal. Every additional period available for search improves the outcomes of searching and makes a searcher more patient and less restrictive in the beginning of the search.

4.3.1.2 Case of infinite $T$, $T = \infty$

Before proceeding to the property of the reservation wage, the expression (25) is rewritten in the form:

\[ R_T^T - R_{T-1}^T = -c + \int_{R_{T-1}^T}^{\infty} \left(u - R_{T-1}^T\right) dF(u) \]

Proposition 14. For an infinitely long time horizon, the reservation utility is constant, $R_T^\infty \equiv R^\infty$.

Proof. Because any two periods of time are equivalent a searcher always has an infinite horizon. From equation (27) one can find the limit when the number of remaining periods tends to infinity. Since the limit of LHS is zero when $T \to \infty$, the implicit function for the reservation utility $R_T^\infty$ is the following:

\[ 0 = -c + \int_{R_T^\infty}^{\infty} \left(u - R_T^\infty\right) dF(u) \]

After integration by parts this equation is transformed to a simpler expression:

\[ \int_{R_T^\infty}^{\infty} (1 - F(u)) du = c \]
From it the reservation utility $R_t^* \equiv R_t$ is constant.

Q.E.D.

Recall, this condition is equivalent to the definition (17) of the marginal agent who is indifferent between moving after sampling one draw and staying. This corresponds to a special case with one destination, $N = 1$, and costs of information search, $c \equiv c(1)$. Therefore the stopping rule is always determined by the initial utility $\bar{u}_0$ of the marginal agent in the FSS(1) model, namely $R_t^* \equiv \bar{u}_0 - c(1)$ for any $1 \leq t \leq T < \infty$.

Thus, comparative statics for the reservation utility is similar to the comparative statics for the indifferent agent in the FSS model. So there is no need for proof of the following:

Proposition 15. For each cost of searching an extra period, $c$, there is a unique reservation utility, solution of (28). It decreases with the cost $c$ and average costs of migration $\bar{c}$ and increases with the average utility $\bar{u}$ and with both variances of utility and cost, $\sigma_u$ and $\sigma_c$.

4.3.2 Search without recall in general case: optimal sample size

Generally, a sequential search leads to a higher expected utility as compared to a fixed-sample-size search. This conclusion is made from the fact that the FSS is more costly because the optimal destination can only be chosen with the entire sample. It can be shown that the FSS is more optimal than the sequential search under some conditions (Morgan and Manning 1985). However, neither of these two approaches is optimal in the general case (Gal et al. 1981, Morgan 1983 and 1985). What
is optimal in the general case is a compound strategy of search combining features of both the FSS and sequential search models. This type of search is the subject of our consideration in the following sections.

### 4.3.2.1 Case of finite $T$, $T < \infty$

The model of sequential search was extended in the search literature in order to capture the features of fixed sample search and to be more realistic. In such a model the agent draws a sample of independent and identically distributed observations from a population with known distribution. The size of this sample is determined for each period. Thus, the optimal stopping rule is characterized by the optimal sample size $N^*_T$ in addition to the sequence of reservation utilities $R^*_T$. Since the search is stopped after the first offer above the reservation utility is obtained then the optimal sample size is zero after that, $N^*_T \geq 0$. Gal et al. (1981) proved that the reservation utility increases with the number of remaining periods, $R^*_T < R^*_{T+1}$, and the sample size, which they call intensity of searching, is non-increasing, $N^*_T \geq N^*_{T+1}$.

The intuition behind reservation utility decreasing in time is similar to the case of a single observation per period. In addition to the proposed above explanation, note, the fewer the periods that remain until the forced termination of searching, the less likely is the searcher to get a draw (or maximum draw in the current period sample) above any fixed number in spite of the possibility of choosing any number of draws each period. The increasing intensity of searching in time reflects the property of decreasing marginal returns to investment in the intensity of searching within any period. This property is shown by Gal et al. (1981).

Although the search process is often modelled without an option to recall the rejected offers, in the framework of the migration model developed in this chapter it will be more realistic to assume
recall and endogenous sample size in each period. To the author’s best knowledge, this type of search model has not been developed in the literature yet.

4.3.2.2 Case of infinite \( T, \ T = \infty \)

There is an important distinction between the two cases of finite and infinite horizons of searching. In the latter the number of periods remaining for each period is the same and equal to infinity. Such being the case, there is no difference between any two periods. As a result the reservation wage and optimal intensity of searching each period are stable, \( R^* = R^\infty \) and \( N^* = N^\infty \) for any \( 0 \leq t \leq \infty \). A constant optimal size of the sample for each period is shown by Morgan (1983). Moreover, he found that this optimal intensity of searching is not greater than any period of optimal intensity of searching for restricted horizon, \( N^* \leq N^T \) for any \( t \) and \( T \) such that \( 0 \leq t \leq T < \infty \).

4.3.3 Migration as search with recall, \( N = 1 \)

A search process with option to recall implies all past destinations are not lost but can be restored. This model has appeared in the consumer search literature. In it a consumer is shopping across a number of firms with known price distribution searching for the lowest price.

The analysis starts from the simplest case with one offer per period. In addition to what is included in the expected gain of extra search in formula (24) one must add the expected gain from the recalled observations:

\[
\bar{EG}_T^T(x) = \max \left( x, -c + \int_{-\infty}^{\infty} \bar{EG}_{T-1}^T(u)dF(u) + \bar{EG}_{T-1}^T(x)F(x) \right)
\]

The bar is added in notations in order to distinguish the case of recall from no recall. It was proved by Lippman and McCall (1976) there is a reservation utility for this problem, \( \bar{R}^T \); that is:
The possibility to restore previous offers increases the expected gain to search and reservation utility. It therefore makes the duration of searching longer. The reservation utility is shown to be not lower than in the case of no recall, \( \bar{R}_i^T \geq R_i^T \) where \( R_i^T \) is from (25) (Lippman and McCall 1976). An important result they obtained is that the reservation utility of searching with recall is constant and coincides with the reservation utility for infinite horizon \( T, \bar{R}_i^T = \bar{R}^T = \bar{R}^\infty \) for any \( 1 \leq T \leq \infty, 0 \leq t \leq T \). Moreover, Lippman and McCall proved that the reservation utilities in the infinite horizon search with and without recall are equal, \( R^\infty = \bar{R}^\infty \).

As one may see, the possibility of recall in the search process is “equivalent” in the strategy, but not in the outcome, to the possibility of searching without time restriction. The outcome of the finite time search with recall may be very different from that in the infinite time search. In the former there is the possibility that all offers would be rejected and the best of them, below the reservation wage, recalled in the end. In the latter the accepted offer is at least the reservation wage. This also implies that the expected duration of searching is longer in the infinite time horizon.

### 4.3.4 Search with recall in general case: optimal sample size

In the general case it is possible to choose how many potential destinations to inspect in one period before deciding whether to choose the “best” destination and stop or proceed with further search.

#### 4.3.4.1 Case of finite \( T, T < \infty \)

As was said in the beginning of section 4.3, the case of finite horizon search with recall and optimal sample size is probably the most realistic model describing migration among all models presented in this chapter. The first result obtained in the literature was a stable reservation utility, \( \bar{R}_i^T \equiv \bar{R}^T \) for
any $t \geq 0$ (Gal et al. 1981). Morgan (1983) extended this result showing that the optimal sample size can be non-monotonic. This property is caused by two opposite forces. On the one hand, the optimal sample size is (not strictly) reducing in time after each “lucky” period when the best observation (i.e. maximum is higher than before) is collected. On the other hand, with time approaching the search horizon it is optimal to (not strictly) increase the sample size to speed-up the process of searching.

A conjecture proposed by Morgan (1983) says that a searcher with full recall never has an optimal sample larger than that in the case of no recall; that is $N_i^T \leq N_i^T$. It was not proved in the general case but was only shown for the penultimate period, $t = T - 1$.

### 4.3.4.2 Case of infinite T, $T = \infty$

The result on constant reservation utility remains for the infinite horizon, $\bar{R}_T^\infty = \bar{R}^\infty$ (Lippman and McCall 1976, Landsberger and Peleg 1977). Non-monotonicity disappears from the property of the optimal sample size because the second force, mentioned in the previous paragraph, is irrelevant. The optimal sample size is not strictly monotonically decreasing, $\bar{N}_t^\infty \leq \bar{N}_{t+1}^\infty$, due to the fact that only the first force mentioned in the previous section is active in this case (Morgan 1983). A necessary but not sufficient condition for strict inequality is that the best observation is obtained in period $t + 1$.

Although the compound strategy is optimal in the general case, Morgan and Manning (1985) claimed that in a wide class of search problems, sequential search may be suboptimal and they studied conditions under which FSS and sequential search are optimal. They demonstrated that sequential search; that is $\bar{N}_t^\infty = 1$, is optimal for a relatively small subset of full recall models with infinite horizon which includes some class of search problems close to that considered in this
section. However, sequential search ceases being optimal when the future is discounted. They also found that an FSS strategy is optimal if the marginal costs of searching or discounting are sufficiently large.

**4.3.5 Extensions of search models**

**4.3.5.1 Search with costly recall**

Basically results of a costly recall search model are unknown as yet. The only paper available in this area is a recent work by Janssen and Parakhonyak (2008). One may not be persuaded by the method and conclusions in this working paper. The authors argue that the reservation wage property is not static and depends on the preceding wage offers. This is unlikely to be the case for both infinite and finite $T$. The basic argument is that recall is never realised, at least before all offers are rejected and time expired, and the authors come to this conclusion usually in the job search framework. What may be affected by cost of recall is the reservation wage but only for the finite horizon. This is easy to illustrate using a passage to the limit. If cost of recall is zero then the reservation price is a positive constant according to the main result above. If it is infinite or sufficiently large then the reservation wage must be zero when one period remains. Before that the reservation wage is also zero only because it excludes the possibility of extremely costly recall. And so on. For the intermediate case of a relatively small non-zero cost of recall the hypothesis is that never realised recall makes the reservation wage be time invariant. Variable reservation wage does not submit to a passage to the limit. The reservation wage only declines with the cost of recall, varying from zero to the maximum value corresponding to not costly recall. These results were explained intuitively and strict proof still needs to be done, though it is beyond the scope of the thesis.
4.3.5.2 Search with uncertain recall and with unknown distribution

The intermediate case between search with and without recall is when a part of the observations is recalled with uncertainty. Karni and Schwartz (1977) considered a case with a solicitation of a past observation which has a Bernoulli distribution. The authors demonstrated that the search in this case is more intensive; that is, the time interval between two consecutive searches is smaller than in the cases of search with and without recall, and increases with past observations of offers especially those with higher probability of recall. The reservation utility is shown to be bound from below and above by the reservation utilities of searching without and with recall respectively.

In the most theoretical models of search perfect knowledge about distribution of offers is assumed. This supposition seems to be implausible for some scholars (e.g. Maier 1985). Rothschild (1974) was the first who considered unknown parameters in the distribution of offers. He assumed that in the search for a lower price problem prices with unknown distribution belong to a finite set and used a searcher’s prior Dirichlet distribution. He concluded that results do not change generally, that is, optimal search has the same properties as in the case of known distribution.

A multiregional version of the model with imperfect information about wage offers is highly intractable. This is because the Bellman equation for the expected returns to search becomes a system of Bellman equations in the general case. With perfect information, job search after migration is the only valid strategy as Maier (1985) noted. With imperfect information the model of job search before migration was initially considered by Maier (1983). In another paper Maier (1985) modelled the situation when search is done after migration. He considered two strategies for the individual, either to make the migration decision or to buy additional information. This model has a limitation since it restricts the migration decision to pair wise comparison of one region with the other regions.
Molho (2001) presented another model with different wage distributions at different locations. He distinguishes between the local labour market job search and a global (the origin and destination labour markets) search. The model showed that under some conditions search can be combined with a speculative migration, that is an unemployed person migrates first and then searches for a job locally. Under other conditions search precedes contracted migration. In this case the unemployed person accepts an offer above the reservation wage and only then migrates.

4.3.5.3 Migration as spatial search

Surprisingly, a real life situation when search items are allocated in space has not been widely studied and barely will be a popular direction of research in the near future due to its complexity. As a rule, in addition to an optimal stopping rule along a route a searcher must develop an optimal route of searching in space, one among overwhelming possible routes. Not many results are available in the area of spatial search but research is progressing and much is still awaiting to be investigated. Surveys of basic results with respect to migration are available in Molho (1986, 2001) and Maier (1990, 1991, and 1993), and for other areas in Maier (1995). Spatial search became an area of intensive research not only in Economics and Mathematics but also in other disciplines, most prominently in Regional Science, in particular in areas of agglomeration, market areas and location, and spatial interaction models.

A classical routing problem considers a travelling sales person who is searching for the optimal path in space. In the general case it is known as a NP (nondeterministic polynomial time) problem, one of the most difficult classes of problems in mathematics. Such problems do not have a computationally feasible solution. This is because CPU time necessary for an almost exhaustive search is proportional to the power function with degree equal to the number of searching nodes (Applegate et al. 2006).
For an illustration of this problem one can consider an Australian who is planning to migrate from Sydney to another State or Territory capital and needs to travel to all of them before making a decision. An optimal route minimises the total, say flight, distance. Then the optimal route is not difficult to find since all cities except Hobart and Adelaide are located roughly on a circle and travel to these two cities apparently has to be linked to the closest pair of other cities on the circle. It seems like the optimal route would be Sydney-Canberra-Hobart-Melbourne-Adelaide-Perth-Darwin-Brisbane-Sydney. Amazingly, a circular flight in the reverse order will probably be somewhat longer due to known phenomenon of a longer flight in the direction of Earth rotation. An optimal route, which minimizes total airfare, could be nontrivial.

4.4 Directions of future research

One possible direction of future theoretical research could be a study of the role of psychic costs in migration. They could be both the cause and consequence of migration. The economic literature so far has considered mostly the latter role. An idea that they could be a cause of migration adds some interesting insights into migration research.

For example, under some circumstances migration is possible even if income in the destination is less than in the origin. Suppose, there are large psychic and other costs such as costs of living in the current location which can be far from where a person is keen to live or where they regularly travel such as the home of fathers. In this case migration can effectively eliminate entire psychic costs or some part of them. This idea may help to explain the return migration phenomenon in the framework of investment in human capital in the case when expectations in the destination, e.g. employment or income level, were not realised on the desired level.

Moreover, it could be a good idea to explain the driving force of repeat migration not, or not only, by a higher average income/utility but by higher psychic and other costs which make the current
location unaffordable or less attractive than another location, the optimal from individual perspectives. An example of a life event which creates such psychic costs is a change in the household composition, e.g. marriage of two young persons who have been living with parents in different cities. It causes migration which creates new psychic costs, say, a need to travel regularly to parents, but this migration can minimise total psychic costs. Now if marriage and migration are followed by the couple’s separation then both of them may want to move back to their parents to eliminate all remaining psychic costs. In another case, if a child was born and therefore, a care from one of the grandmothers is badly needed, then the family can move back, say to the mother’s parents. In reality the full picture is even more complicated because there could be individual, family, and extended family psychic costs.

Remarkably, even if population tends to be perfectly homogeneous in terms of abilities and incomes across cities, migration is likely to be observed at significant levels due to previous migrations which created some distribution of psychic costs. Though, this distribution will be endogenous to many factors exogenous in individual migration decisions.

From the perspective of the theoretical model presented in this chapter, the idea is that psychic costs of living far from friends and relatives, generally not observed by researchers, change both pre-move and post-move utility distribution in the model. Since it is assumed that pre-move utility is normally distributed as well as random draws of utility for those considering migration, after deduction of normally distributed unobserved psychic costs, a random variable for net utility in destinations has a lower mean and higher variation. Thus, the psychic costs are introduced into the model by means of another random variable deducted from both the initial utility and utility across the destinations. This changes the model results but only insignificantly. Although the probability of being a migrant is different now for everyone, because psychic costs now change both the initial utility and the post-move utility, comparative statics studied in this chapter is valid here.
Another possible extension of the model developed in this chapter is to consider the migration experience. This variable, described in Chapter 2, was found to significantly affect the probability of migration. The history of previous migrations is important therefore. The exact change in the model results depends on the type of move, either initial (say, moving out of the parents’ house, maybe not for the first time) or repeat (moving away from other than the parents’ house). For initial migration both costs are correlated, thereby magnifying the total effect. In contrast, for repeat migration they are not correlated because home is not the origin any more unless the repeat migrant starts a new migration from the home which is not unlikely in general. As a result initial migration is less likely to be done than repeat migration.

Formally this follows from the inequality:

\[ P\{x-2\alpha > z\} < P\{x-\alpha - \beta_2 > z\} \]

where \( z \) is the initial utility and \( x \) is utility in the destination, both are i.i.d. random variables and hence, for migrants \( z \) is less than the mean utility in most cases. For an initial migration \( \alpha \), costs of migration with a positive mean, is correlated with the psychic costs \( \beta_1 \) since it is assumed that in order to compensate for these costs one needs to move \( \alpha \) in the opposite direction. Both costs in general are i.i.d. random variables. They are equal on the LHS of the inequality. This is why the sum of \( \alpha + \beta \) is replaced by \( 2\alpha \). On the RHS \( \alpha \) is costs of migration and \( \beta_2 \) is psychic costs of the repeat move. Both random variables are i.i.d. now since a secondary move is random and therefore, not correlated with the random initial move. Now there are two random variables \( \Psi_1 = x-2\alpha \) and \( \Psi_2 = x-\alpha - \beta_2 \) which have the same mean, since

\[ E(\Psi_1) = E(x) - 2E(\alpha) = E(\Psi_2) \]

but different dispersions
\[ D(\Psi_1) = D(x) + 4D(\alpha) < D(\Psi_2) = D(x) + 2D(\alpha) \]

This inequality guarantees the desired inequality \( F_1(z) \equiv P\{\Psi_1 > z\} < F_2(z) \equiv P\{\Psi_2 > z\} \) for agents with a low initial utility \( z < E(x) - 2E(\alpha) \), that is for those who are more likely to be migrants in the model. Inversely, higher initial utility agents are more likely to be initial migrants than repeat migrants but the likelihood of even the first migration is relatively lower than for lower utility agents. As a result of this discussion one can conclude that migrants are selected not only because they are more likely to have a lower initial utility but also higher psychic costs (for repeat moves). Both are expected to regress to the mean after migration in our random migration framework.

An interesting extension of the theoretical model on the circle arises in the case of repeat migration. One can envisage that there is more than one period and in each period information about \( N \) destinations is obtained and the migration decision is made. Then, it is possible to show that the two period model is equivalent in its aggregate effect to the \( 2N \) destinations single period model for those who migrated in period one and to the \( N \) destinations model for period one stayers. This extension adds dynamics into the static model. As a result, there will be a dynamic distribution of utility which changes after some agents migrate each period.

In the model of migration on a circle in Appendix 3 the endogeneity of the information search was considered as a special case. It would be interesting to study the effect of budget constraints in this model as well as in the model with an exogenous number of destinations. Regarding the model with endogenous information search it is possible to endogenise the distribution of initial utilities and budgets available for search and migration but this can be done only in the dynamic contest such as the one proposed above. More mobile people are expected in general to have a higher initial utility.
and budget for search. As a result under some assumptions there could be either convergence or even divergence of utilities for mobile and immobile populations in the course of time.

Another possible research direction is to model the migration destination search in a general equilibrium framework. This approach can be similar to equilibrium in job search models. Since the current model of migration does not consider any markets one needs to explicitly model either the labour or the housing market. There are models which consider these two markets in the equilibrium (Combes et al. 2008).

The equilibrium in the model with migration can also arise as a result of interjurisdictional competition where people “vote with their feet” towards jurisdictions providing better public goods in addition to exogenously given amenities (Tiebout hypothesis). Public goods provision may depend on congestion in order not to make all populations migrate to one point.

Search in space is another intriguing topic, though a comprehensive and tractable model is very difficult to develop. Relative to other labour markets the geographical location of an individual starting a search matters, because information is limited and its collection depends on distance. Initially people living far from the largest cities could be in an inferior equilibrium, say, with lower income, lower price of housing, lower quality of life, but higher price level due to restricted competition. There could be two forces which affect their mobility in different directions. On the one hand, they have less information because of remoteness which is an impediment to migration. On the other hand, they have more opportunities for income and quality of life improvement if they are able to overcome these difficulties linked to the lack of information and distance barrier.

**4.5 Conclusions**

Theoretical models of migration in economics traditionally view an agent as one seeking potential gains in other destinations. The search literature adds uncertainty to the process of job search in
distant labour markets in the same manner it does for a local job search. In this chapter two
approaches to the migration destination search with uncertainty were adopted. The first is a fixed-
sample-size information search which showed that a volume of information on alternative
destinations is important. The most interesting result, perhaps, is the new result obtained for the
model with endogenous sample size. The optimal volume of information is shown to depend on
individual initial conditions. Consequently, richer agents will prefer not to start searching
information and not to undertake migration. However, poorer will search the optimal volume of
information which entirely depends on the budget available for search. In the end, of those who are
poorer, those who can afford to buy relatively more information than others should be expected to
be better off. Thus, the chapter found evidence of the liquidity constraints hypothesis in migration
quite commonly met in microeconomics. For the special case of migration on a circle, the migration
rate is shown to be increasing with amount information available to each agent, with a maximum in
the case of full information, but with decreasing costs associated with migration.

The second model of sequential information search treats search as done more optimally. Namely,
only one destination or, in the general case, an optimal number of destinations, is assessed each
period. It led to conclusions about the level of the reservation utility, any utility above which the
process of search is terminated. This level of the reservation utility, as shown in the traditional
search literature, depends mostly on the number of periods remaining and is different for any finite
and infinite time horizon.

If one were to choose between different models with a sequential search, then the model with recall
is preferred. Migration destinations evaluated in the process of search can be restored. This
contrasts the model of migration destination search with the traditional model of job search because
job offers tend to expire.
A few remarks about whether the assumptions made in the theoretical models are realistic should be
given. The first approach is probably less realistic since it seems doubtful that everyone collects all
necessary information about the fixed number destinations in a given period. In addition, it is likely
that information about some destinations will be incomplete. It is more realistic that every agent
decides independently whether to do a search and how many destinations to explore, that is the
number of destinations is endogenous.

It is also doubtful that the decision is made exactly at the end of a period. It is more likely that
information or important parts of it turn up in some random ways. Thus, the second approach seems
to be more reliable. It is in essence the process of sequential search similar to a search for a job or
minimum price. The process studied here, however, is more general because the job search process
generates offers which are inseparable from other characteristics of the location and therefore, from
utility an individual assigns to the destination.

Both theoretical models considered in this chapter seem to be equally important because they
basically cover the main ways individuals collect information about destinations under uncertainty
about costs of search and move and benefits of migration.
Appendix 3

A3 Model of migration on circle

This section presents a simple model of migration on a circle with a continuum population. Migration is represented as an outcome of a free lottery. An example of such a lottery is Green the Card lottery in the US. However, in the model everyone receives an invitation to a particular random destination with information about the utility there, which can be either accepted or rejected.

A3.1 One destination

An agent knows utility $x$ at some point on a unit circle at distance $\alpha \in [0,1]$. Cost of a move is a linear function and equal to the distance for simplicity. Then the agent with initial utility $z$ migrates to that point only if net utility after move is greater than the current utility:

$$x - \alpha > z$$  \hspace{1cm} (A1)

The relationship between parameters of the model and probability of move are very simple in this case. The agent has a higher chance of migration if utility at destination is higher or distance is lower or agent’s initial utility is higher.

Now assume that each agent out of the continuum set simultaneously draws a lottery ticket and gets information about utility $x$ at some random for each agent point located on distance $\alpha$, $\alpha \in [0,1]$, and decides whether to move to that point or stay. Three random variables are i.i.d. with uniform distribution on $[0,1]$ and c.d.f. $F(x) = x$.

One can find a proportion of agents who migrate, average distance of migration, average utility of stayers, and average pre-migration and post-migration utility of migrants.
Migration rate or proportion of movers is the probability of inequality (A1). It is equal to a triple integral:

\[ M = \int \int \int dz \, dx \, d\alpha = \frac{1}{6}. \]

Therefore, the proportion of stayers is equal to:

\[ S = \int \int \int dz \, dx \, d\alpha = 1 - M = \frac{5}{6}. \]

Average distance and other average characteristics are calculated only for movers disregarding stayers who move zero distance. Thus, the average distance is a ratio of the integrated distance and the proportion of movers:

\[ \bar{\alpha} = \frac{\int \int \int \alpha \, dz \, dx \, d\alpha}{M} = \frac{1}{4}. \]

Now migrants’ utility gain can be identified. The average pre-migration utility of migrants is:

\[ \bar{z} = \frac{\int \int \int z \, dz \, dx \, d\alpha}{M} = \frac{1}{4}, \]

and their average post-migration utility is:

\[ \bar{x} = \frac{\int \int \int x \, dz \, dx \, d\alpha}{M} = \frac{3}{4}. \]

Therefore, the average net returns to migration are:

\[ \bar{r} = \bar{x} - \bar{\alpha} - \bar{z} = \frac{1}{4}. \]
Also, the average utility of stayers is determined by the integral:

\[
\overline{z}_s = \frac{1}{1} \int_0^1 \int_0^{\max\{x-\alpha,0\}} z \, dz \, dx \, d\alpha / S = \frac{11/24}{5/6} = \frac{11}{20}.
\]

As one can notice, on average a stayer has a utility slightly above the median utility, \( \frac{1}{2} \). A migrant triples his/her utility on average, having a half of the median utility before move and getting the median utility as returns.

**A3.2 Two destinations**

Suppose an agent can get two lottery tickets, that is can randomly select two destinations on distance \( \alpha \) and \( \beta \), \( \alpha, \beta \in [0,1] \), \( \alpha < \beta \) without restricted generality, and can determine utility at these locations, respectively \( x \) and \( y \), without any costs. Then the agent decides whether to move or stay at the initial point. Therefore, a destination is determined by maximum utility net less travel costs. It is a solution of the problem:

\[
\max \{z, y-\beta, x-\alpha\}
\]

The agent stays if:

\[
z > \max \{y-\beta, x-\alpha\}
\]

He/she goes to the lower costs destination if:

\[
x-\alpha > \max \{y-\beta, z\}
\]

or to the higher cost destination if:

\[
y-\beta > \max \{x-\alpha, z\}
\]
Thus, the probability of being a migrant is an increasing function of both the destinations’ utility and the decreasing of both distances and current utility.

As in section A1, it is assumed that all five parameters of the model \( \alpha, \beta, x, y, \) and \( z \) are i.i.d. random variables with uniform distribution on the segment \([0,1]\). Then the proportion of the immobile population is determined by the probability of inequality (A2) being held, which is a quintuple integral in the case of two destinations:

\[
S = 2 \int_0^1 \int_0^1 \int_0^1 \int_0^1 \int_0^1 dx\,dy\,dz\,d\alpha\,d\beta
\]

Note the multiplier 2 appears in this formula since the assumption made in the very beginning, that \( \alpha < \beta \), restricts the number of possible cases by half. It implies that the symmetric case, when \( \alpha > \beta \), with a similar formula is omitted and, therefore, the multiplier 2 is used in the formula here and thereafter.

It will be useful for further computation to find an auxiliary double integral for an agent with utility \( z \) and two destinations on some fixed distance \( \alpha \) and \( \beta \) with random utility \( x \) and \( y \) respectively.

Then, the probability of stay for the agent is integral:

\[
S(z, \alpha, \beta) = \int_0^{\min\{z+\beta,1\}} \int_0^{\min\{z+\alpha,1\}} dx\,dy = \begin{cases} 
(z+\alpha)(z+\beta) & \text{if } z < 1-\beta \\
z+\alpha & \text{if } 1-\beta < z < 1-\alpha \\
1 & \text{if } 1-\alpha < z
\end{cases}
\]

One can observe that for two given distances agents with a high initial utility do not migrate. In contrast, those with a lower initial utility have a lower probability of stay, which increases with the initial utility, first, quadratically and then linearly.

It is straightforward to find the total number of stayers:
\[
S = 2 \int_{0}^{1} \int_{0}^{1} S(z, \alpha, \beta) dz \, d\alpha \, d\beta = \frac{43}{60}
\]

Therefore the migration rate is the proportion of the population which is mobile:

\[
M = 1 - S = \frac{17}{60}
\]

Now one can find the probability of a move to nearby and distant destinations, that is the probability of inequalities (A3) and (A4). Similar to the way it was done for stayers, the first step is to find the probabilities of migration to both destinations, keeping the three parameters \( z, \alpha, \) and \( \beta \) fixed:

\[
M_\alpha(z, \alpha, \beta) = P\{x - \alpha > \max\{y - \beta, z\}\} = \int_{\min\{z+\beta,1\}}^{1} \int_{0}^{\min\{z+\beta,1\}} dy \, dx + \int_{\min\{z+\beta,1\}}^{1} \int_{y-\beta+\alpha}^{1} dx \, dy
\]

\[
\begin{cases}
\frac{1}{2} + \beta - \alpha - \frac{1}{2}(z + \beta)^2 & \text{if } z < 1 - \beta \\
1 - z - \alpha & \text{if } 1 - \beta < z < 1 - \alpha \\
0 & \text{if } 1 - \alpha < z
\end{cases}
\]

and

\[
M_\beta(z, \alpha, \beta) = P\{y - \beta > \max\{x - \alpha, z\}\} = \int_{\min\{z+\alpha,1\}}^{1} \int_{0}^{\min\{z+\alpha,1\}} dx \, dy + \int_{\min\{z+\alpha,1\}}^{1} \int_{x-\alpha+\beta}^{1} dy \, dx
\]

\[
\begin{cases}
\frac{(1+\alpha-\beta)^2}{2} - \frac{1}{2}(z+\alpha)^2 & \text{if } z < 1 - \beta \\
0 & \text{if } 1 - \beta < z
\end{cases}
\]

For small enough \( z \) and any \( \alpha \) and \( \beta \) the probability of migration to the smallest distance is greater than that to the largest distance. Only for agents with a higher initial utility, \( z \geq 1 - \alpha \), both probabilities are zero.
Second step, the proportions of migrants on both distances are:

\[ M_\alpha = 2 \int_{0}^{\beta} \int_{0}^{\alpha} M_\alpha(z, \alpha, \beta) dz d\alpha d\beta = \frac{7}{30} \quad \text{A6} \]

and

\[ M_\beta = 2 \int_{0}^{\beta} \int_{0}^{\alpha} M_\beta(z, \alpha, \beta) dz d\alpha d\beta = \frac{1}{20} \quad \text{A7} \]

Following the logic, the sum of all three probabilities for stayers and movers to the smaller and the larger distances found in (A5)-(A7) is equal to 1:

\[ S + M_\alpha + M_\beta = 1 \]

In proportions, 72 percent of population does not move, 23 percent moves a shorter distance and 5 percent moves a longer distance. In addition to that average distances for the lower and the higher distance movers are equal to respectively:

\[ \bar{\alpha} = 2 \int_{0}^{\beta} \int_{0}^{\alpha} \alpha M_\alpha(z, \alpha, \beta) dz d\alpha d\beta \Bigg/ M_\alpha = \frac{17}{84} \]

and

\[ \bar{\beta} = 2 \int_{0}^{\beta} \int_{0}^{\alpha} \beta M_\beta(z, \alpha, \beta) dz d\alpha d\beta \Bigg/ M_\beta = \frac{7}{18} \]

On average a short distance migrant goes 0.20, which is approximately a half of the average long distance move equal to 0.39. The average distance moved by all migrants:
\[ d = \frac{\bar{\alpha} M_\alpha + \bar{\beta} M_\beta}{M} = \frac{4}{17} \]

The next step is to find the average utility of stayers:

\[ \bar{z}_s = 2\int_0^1 \int_0^1 \int_0^1 zS(z, \alpha, \beta) dz \, d\alpha \, d\beta \bigg/ S = \frac{51}{86} \]

The average pre-migration utility of the smaller and the larger distance migrants is respectively:

\[ \bar{z}_\alpha = 2\int_0^1 \int_0^1 \int_0^1 zM_\alpha(z, \alpha, \beta) dz \, d\alpha \, d\beta \bigg/ M_\alpha = \frac{23}{84} \]

and

\[ \bar{z}_\beta = 2\int_0^1 \int_0^1 \int_0^1 zM_\beta(z, \alpha, \beta) dz \, d\alpha \, d\beta \bigg/ M_\beta = \frac{2}{9} \]

The movers to the larger distance are about 23 percent better off initially than the movers a shorter distance, indicating affordability of such distant moves.

Moreover, the average pre-migration utility of migrants is:

\[ \bar{z} = \frac{\bar{z}_\alpha M_\alpha + \bar{z}_\beta M_\beta}{M} = \frac{9}{34} \]

which is approximately 45 percent of that of stayers.

Finally, the average post-migration utility for both migrants' categories is the mathematical expectation

\[ \bar{x}_\alpha = E \left[ x \mid x - \alpha > \max \{ y - \beta, z \} \right] \]
and

\[ \bar{x}_\beta = E[y | y - \beta > \max \{x - \alpha, z\}] \]

As usual, one can start with the calculation of integral for an agent with utility \( z \) and the fixed distances \( \alpha \) and \( \beta \):

\[
\bar{x}_\alpha(z, \alpha, \beta) = \int_{\min[z+\beta, 1]}^{1} \int_{\min[z+\alpha, 1]}^{\min[z+\beta, 1]} x \, dy \, dx + \int_{\min[z+\beta, 1]}^{1} \int_{\min[z+\beta, 1]}^{\min[z+\beta, 1]} x \, dx \, dy
\]

\[
= \left\{ \begin{array}{ll}
\frac{1}{3} + \frac{1}{2} (z+b)^2 & \text{if } z < 1 - \beta \\
\frac{1}{6} (z+b)^3 + \frac{1}{2} (b-a) \left(1-(z+b)^2\right) - \frac{1}{2} (z+b) & \\
\frac{1}{2} - \frac{1}{2} (z+a)^2 & \text{if } 1 - \beta < z < 1 - \alpha \\
0 & \text{if } z > 1 - \alpha
\end{array} \right.
\]

and

\[
\bar{x}_\beta(z, \alpha, \beta) = \int_{\min[z+\alpha, 1]}^{1} \int_{0}^{\min[z+\beta, 1]} y \, dx \, dy + \int_{\min[z+\alpha, 1]+\beta}^{1} \int_{\min[z+\alpha, 1]+\beta}^{\min[z+\alpha, 1]+\beta} y \, dy \, dx
\]

\[
= \left\{ \begin{array}{ll}
\frac{1}{2} (z+a) \left(1-(z+b)^2\right) - \frac{1}{6} (1+a-b)^3 + \frac{1}{6} (z+a)^3 + \frac{1}{2} (a-b) \left(1+(a-b)^3 -(z+a)^2\right) & \\
\frac{1}{2} - \frac{1}{2} (z+b) + \frac{1}{2} (b-a) \left(1-(z-b)^2\right) & \text{if } z < 1 - \beta \\
0 & \text{if } z > 1 - \beta
\end{array} \right.
\]

Then after a move the three average utilities are equal to

\[
\bar{x}_\alpha = 2 \int_{0}^{\beta} \int_{0}^{1} \bar{x}_\alpha(z, \alpha, \beta) \, dz \, d\alpha \, d\beta / M_\alpha = \frac{3}{4}.
\]
\[ \bar{x}_\beta = 2 \int_0^1 \int_0^1 \int_0^1 \bar{x}_\beta (z, \alpha, \beta) \, dz \, d\alpha \, d\beta / M_\beta = \frac{5}{6}, \]

and

\[
\bar{x} = \frac{\bar{x}_\alpha M_\alpha + \bar{x}_\beta M_\beta}{M} = \frac{13}{17}
\]

The average net return for lower distance migrants is:

\[
\bar{r}_\alpha = \bar{x}_\alpha - \bar{\alpha}_\alpha - \bar{z}_\alpha = \frac{23}{84}
\]

and for higher distance:

\[
\bar{r}_\beta = \bar{x}_\beta - \bar{\alpha}_\beta - \bar{z}_\beta = \frac{2}{9}
\]

Finally, the total average net returns:

\[
\bar{r} = \bar{x} - \bar{\alpha} - \bar{z} = \frac{9}{34}
\]

**A3.3 Full information**

Up to now cases with partial information about costs and utility have been considered. A possible extension is an assumption about full information. Under full information any agent knows utility at any point on the circle.

Assume for a moment there is a finite number of random destinations, points on the circle, \( \alpha_1, \alpha_2, \ldots, \alpha_n \in [0,1] \). Distances to them and utilities in these destinations are given for an agent with
utility $z$. Rearrange $\alpha$-s in ascending order $\alpha_1 < \alpha_2 < \ldots < \alpha_n$. Then the probability of migration is represented by a formula:

$$P\left\{ \max (x_1 - \alpha_1, x_2 - \alpha_2, \ldots, x_n - \alpha_n) > z \right\} = 1 - \prod_{k=1}^{n} P\{x_k - \alpha_k < z\} = 1 - \prod_{k=1}^{N} (z + \alpha_k) \quad \text{A8}$$

where $N$ is determined by $\alpha_1, \alpha_2, \ldots, \alpha_n$:

$$\begin{cases} 
\alpha_k < 1 - z & \text{if } k \leq N \\
\alpha_k \geq 1 - z & \text{if } k > N 
\end{cases} \quad \text{A9}$$

Since $\alpha$-s are random then it is not difficult to estimate this probability when $n$ tends to infinity. For any single destination out of $n$ the probability of being better than the initial point is:

$$P\{x - \alpha > z\} = \int_{0}^{1-z} \int_{z+\alpha}^{1} dx d\alpha = \frac{(1 - z)^2}{2} \quad \text{A10}$$

This formula shows that for sufficiently large $n$ the number of destinations in which the agent with utility $z$ is better off is close to a real number:

$$n \frac{(1 - z)^2}{2} \quad \text{A11}$$

Note also that for large $n$ the number of factors in the formula for probability (A8) is approaching $N = n(1 - z)$ which is greater than the number of better off points in (A11).

Intuitively, the best destination has a utility equal to 1. Moreover in any $\varepsilon$-circle there are points with utility infinitely close to 1. Keeping this in mind it is possible to assert the following proposition.
**Proposition A3.1.** Under full information any agent with a non-unit utility is a migrant with unit probability.

Proof. It is enough to demonstrate that for any \( z \in [0,1] \) and random numbers \( \alpha_1, \alpha_2, \ldots, \alpha_n \in [0,1] \) the following limit is equal to zero:

\[
\lim_{N \to \infty} \prod_{k=1}^{N} (z + \alpha_k) = 0
\]

where \( N \) is determined by condition (A9).

This follows from the fact that for any \( \varepsilon < \frac{1-z}{2} \) there exists a sufficiently large \( N = N(\varepsilon, z) \) such that:

\[
\prod_{k=1}^{N} (z + \alpha_k) < \varepsilon
\]

Indeed according to formula (A10) for large enough \( N \) the proportion of better off points within distance \( \varepsilon \) (that is all points with \( \alpha < \varepsilon \)) is close to \( \varepsilon n \left( \frac{1-z}{2} \right)^2 \). Then one can replace \( \varepsilon \) for these \( \alpha \) and \( 1-z \) for other \( \alpha \) in the formula with the product of \( N \) multipliers. Therefore for any \( n \),

\[
n > N = \frac{\ln \varepsilon}{\varepsilon} \frac{2}{(1-z)^2 \ln \frac{1+z}{2}}
\]

one can get the desired bound from above:

\[
\prod_{k=1}^{N} (z + \alpha_k) < (z + \varepsilon)^{\varepsilon n(1-z)^2/2} < \left( \frac{1+z}{2} \right)^{\varepsilon n(1-z)^2/2} < \varepsilon
\]
Q.E.D.

One remark needs to be made here. For this proposition to be true it is enough to assume a weaker condition that utility is known at a numerable set of random destinations, that is with the mathematical measure of the set equal to zero, but still consisting of an infinite number of random points.

Thus, it is found that all continuum agents on the circle except for the richest agent with unit utility are migrants. As a consequence the migration rate is equal to one, $M_\infty = 1$, and the proportion of stayers is 0, $S_\infty = 0$. The average pre-migration utility of movers is a half, $\bar{z}_\infty = \frac{1}{2}$, and the average utility of stayers is 1, $\bar{z}_\infty = 1$. The average post-migration utility is 1, $\bar{z}_\infty = 1$. The average distance moved by each migrant is equal to zero, $\bar{\alpha}_\infty = 0$ since any utility arbitrarily close to 1 is found in an arbitrary small neighbourhood. Therefore, the average returns to move are one half, $\bar{r}_\infty = \bar{x}_\infty - \bar{z}_\infty - \bar{\alpha}_\infty = \frac{1}{2}$. These results were summarized among others in Table 21 in Chapter 4. Next is a case with an arbitrary number of destinations.

A3.4 General case: N destinations

First, a formula for the proportion of stayers $S_N$ is derived. One can fix utility $z$ and find the probability of stay using formula (A8):

$$S_N(z) = P\left\{ \max(x_1 - \alpha_1, x_2 - \alpha_2, ..., x_N - \alpha_N) < z \right\} = \prod_{k=1}^{N} P\{x_k - \alpha_k < z\}$$

$$= \left( P\{x - \alpha < z\} \right)^N = \left( 1 - \frac{(1-z)^2}{2} \right)^N$$
Note that Proposition A3.1 follows from this formula immediately.

**Proposition A3.2.** Probability of stay is increasing with $z$ and decreasing with $N$.

Proof. It is straightforward after two derivatives are calculated:

$$\frac{\partial S_N(z)}{\partial z} = N(1-z) \left(1 - \frac{(1-z)^2}{2}\right)^{N-1} > 0$$

and

$$\frac{\partial S_N(z)}{\partial N} = \left(1 - \frac{(1-z)^2}{2}\right)^N \ln \left(1 - \frac{(1-z)^2}{2}\right) = S_N(z) \ln(S_1(z)) < 0$$

since it is the product of positive and negative terms.

Q.E.D.

It should be noted here that a case of endogenous $N$ does not add new insights into the model and results. If the model allows agents to choose the number of offers $N$, they would prefer to have an infinite number of random destinations, $N = \infty$. This is possible only if a search for destinations is costless.

Define a random variable $\Psi = \max_{i=1,...,\xi} (\xi_i - \eta_i)$ with a c.d.f. $F_\psi(z) = P\{\Psi < z\} = S_N(z)$, which is probability of stay at a current point with utility $z$.

The expected proportion of stayers in this model can be represented either in the integral form or in series.
\[ S_N = P\{\Psi < \xi\} = \int_0^1 F_\Psi (z) \, dz = \int_0^1 S_N (z) \, dz = \int_0^1 \left(1 - \left(\frac{1-z}{2}\right)^2\right)^N \, dz = \sum_{k=0}^N \frac{(-1)^k \, C_N^k}{2^k (2k+1)} \quad A12 \]

Therefore the expected proportion of migrants or migration rate, evaluated by the integral or series, has a view:

\[ M_N = 1 - S_N = 1 - \int_0^1 \left(1 - \left(\frac{1-z}{2}\right)^2\right)^N \, dz = \sum_{k=1}^N \frac{(-1)^{k+1} \, C_N^k}{2^k (2k+1)} \quad A13 \]

Note the only difference in the last formula from the previous one is in the omitted first term under the summation.

It is not difficult to check that the result obtained in the cases of one and two destinations follows from the general case since:

\[ M_1 = \frac{(-1)^3 \, 1}{2^1 (2 \times 1 + 1)} = \frac{1}{6} \]

and

\[ M_2 = \frac{(-1)^3 \, 2}{2^2 (2 \times 1 + 1)} + \frac{(-1)^2 \, 1}{2^2 (2 \times 2 + 1)} = \frac{2}{6} - \frac{1}{20} = \frac{17}{60} \]

**Proposition A3.3.** The migration rate increases with the number of destinations.

Proof. It can be noted from that the integral in formula (A12) the proportion of stayers \( S_N \) declines with \( N \).

Q.E.D.
Utility of stayers after application of the definition for conditional mathematical expectation

\[ z_{r,N} = E \{ \xi E < \xi \} = E \{ E \{ \xi | x < \xi \} |_{x=\Psi} \} = E \{ z dz \}_{x=\Psi} = \frac{1}{2} - \frac{1}{2} E \{ \Psi^2 \} \]

where in its turn \( E \{ \Psi^2 \} = \int_0^1 z^2 dF_\Psi(z) \).

Using integration by parts, the mathematical expectation can be rewritten in a form:

\[ E \{ \Psi^2 \} = z^2 F_\Psi(z)|_0^1 - 2 \int_0^1 z F_\Psi(z) dz = 1 - 2 \int_0^1 F_\Psi(z) dz + 2 \int_0^1 (1-z) F_\Psi(z) dz \]

Though one can further simplify this expression, to:

\[ 1 - 2 \int_0^1 z F_\Psi(z) dz , \]

it is possible to do another trick here. It is easy to calculate the third term in the expression:

\[ 2 \int_0^1 (1-z) F_\Psi(z) dz = 2 \int_0^1 \left( 1 - \frac{(1-z)^2}{2} \right)^N d \left( 1 - \frac{(1-z)^2}{2} \right) \]

\[ = 2 \left. \left( 1 - \frac{(1-z)^2}{2} \right)^N \right|_0^1 = 2 \left. \left( 1 - \frac{1}{2^{N+1}} \right) \right|_0^1 \]

and finally get:
\[ E\{\Psi^2\} = 1 - 2S_N + \frac{2}{N+1} \left(1 - \frac{1}{2^{N+1}}\right) \]

From which the average utility of stayers is found:

\[ \bar{z}_{s,N} = \frac{1}{2} - \frac{1}{2} E\{\Psi^2\} S_N = \frac{S_N}{N+1} \left(1 - \frac{1}{2^{N+1}}\right) - \frac{1}{N+1} \left(1 - \frac{1}{2^{N+1}}\right) \]

Pre-migration utility of migrants:

\[ z_{m,N} = E\{z^{s_N} > \xi^{s_N}\} = E\left\{E\{z^{s_N} | x^{s_N} > \xi^{s_N}\} \right|_{x^{s_N} = \Psi^{s_N}} \right\} = E\int_{x^{s_N} = \Psi^{s_N}} z \, dz \bigg|_{x^{s_N} = \Psi^{s_N}} = \frac{1}{2} E\{\Psi^2\} \]

Proposition A3.4. The sum of total utility for stayers and migrants before they move is equal to a half, \( z_{s,N} + z_{m,N} = \frac{1}{2} \).

Proof. It is intuitively clear that total utility does not change and is still equal to the expected value of utility, \( \frac{1}{2} \). Also this immediately follows from formulas (A13) and (A15).

Q.E.D.

The average pre-migration utility of migrants is:
\[
\overline{z}_{m,N} = \frac{1}{M_N} \cdot \frac{1}{2} E\{\Psi^2\} = 1 - \frac{1}{2} \cdot \frac{1}{N+1} \left(1 - \frac{1}{2^{N+1}}\right)
\]

\[
= 1 - \frac{1}{2} \cdot \frac{1}{N+1} \left(1 - \frac{1}{2^{N+1}}\right) = 1 - \frac{1}{2} \cdot \frac{1}{N+1} \left(1 - \frac{1}{2^{N+1}}\right)
\]

This result can also be derived from Proposition A3.4:

\[
\overline{z}_{m,N} + \overline{z}_{s,N} = \frac{1}{2}
\]

First, one can find the second summand

\[
\overline{z}_{s,N} = \int_0^1 S_N(z) dz = \int_0^1 \left(1 - \frac{(1-z)^2}{2}\right)^N dz = \int_0^1 \left(1 - \frac{(1-z)^2}{2}\right)^N dz + \int_0^1 \left(1 - \frac{(1-z)^2}{2}\right)^N dz
\]

\[
= -\int_0^1 \left(1 - \frac{(1-z)^2}{2}\right)^N dz = S_N = \frac{1}{N+1} \left(1 - \frac{1}{2^{N+1}}\right)
\]

Finally, the average pre-migration utility of migrants:

\[
\overline{z}_{m,N} = \frac{1}{M_N} \left(\frac{1}{2} - S_N + \frac{1}{N+1} \left(1 - \frac{1}{2^{N+1}}\right)\right) = 1 - \frac{1}{2} \cdot \frac{1}{N+1} \left(1 - \frac{1}{2^{N+1}}\right)
\]

the same formula as (A14).

Total post-migration utility of migrants can be expressed in the form of the mathematical expectation:

\[
x_{m,N} = E\{\Psi|\Psi > \xi\} = E\left(E\{\Psi|\Psi > y\}\right)_{y=\xi} = \int_0^1 \int_y^1 dF_{\Psi}(z) dy
\]

It is calculated using a different approach below.
The total distance migrants move is the expected value of the random variable \( \eta \) conditional on migration

\[
\alpha_{m,N} = E\{\eta|\Psi > \xi\}
\]

where \( \eta \) is one of \( \eta_1, \ldots, \eta_N \) for which maximization is achieved, \( \eta = \arg \max_{i=1,\ldots,N} (\xi_i - \eta_i) \). Let \( x, x_i \), and \( \alpha_i \) be realizations of the random variables \( z, \xi_i \) and \( \eta_i \) respectively. In order to represent this formula in a computable form assume that \( \max_{i=1,\ldots,N} (x_i - \alpha_i) = x_i - \alpha_i \) without any loss of generality.

Therefore the total distance moved by the population of agents can be written as:

\[
\alpha_{m,N} = E\{\alpha_i \mid \max_{i=1,\ldots,N} (x_i - \alpha_i) = x_i - \alpha_i, x_i - \alpha_i > x\} = E\{\alpha_i \mid \alpha_i < x_i - x, \alpha_i < x_i - x_2 + \alpha_2, \ldots, \alpha_i < x_i - x_N + \alpha_N\}
\]

This mathematical expectation can also be calculated below. Calculation of the average distance and utility after move require some preliminary computations. First, find the probability of a move to a particular destination preferred over both the initial point and other \( N-1 \) alternative destinations. Fix the destination utility \( x \) and distance \( \alpha \). Imagine that every agent has the best destination (among \( N \)) which is characterized by utility \( x \) and distance \( \alpha \) from the initial point.

Then the probability of migration to this utility point and on this distance is

\[
P_N(x, \alpha) = P\{x - \alpha > \max\{z, x_1 - \alpha_1, \ldots, x_{N-1} - \alpha_{N-1}\}\} = P\{x - \alpha > z\} P\{x - \alpha > x_1 - \alpha_1\} \ldots P\{x - \alpha > x_{N-1} - \alpha_{N-1}\}
\]

\[
= \left\{ \begin{array}{ll}
(x - \alpha) \prod_{k=1}^{N-1} \int_0^{\min(1,x - \alpha + \alpha_k)} dx_k d\alpha_k & \text{if } x > \alpha \\
0 & \text{if } x \leq \alpha
\end{array} \right.
\]
where \( \int_0^{\min(1, x+\alpha_k)} dx_x d\alpha_k = 1 - \frac{(1-x+\alpha)^2}{2} \) for any \( k = 1, \ldots, N-1 \)

and hence, the product of \( N-1 \) multipliers is equal to:

\[
\left( 1 - \frac{(1-x+\alpha)^2}{2} \right)^{N-1}
\]

Therefore, the final view of the probability is:

\[
P_N(x, \alpha) = \begin{cases} 
(x-\alpha) \left( 1 - \frac{(1-x+\alpha)^2}{2} \right)^{N-1} & \text{if } x > \alpha \\
0 & \text{if } x \leq \alpha 
\end{cases}
\]

Now one can find the expected migration on distance \( \alpha \) integrating this function by the destination utility \( x \):

\[
P_{1,N}(\alpha) = \int_0^1 P_N(x, \alpha) dx = \int_0^1 (x-\alpha) \left( 1 - \frac{(1-x+\alpha)^2}{2} \right)^{N-1} dx \]

\[
= \int_\alpha^{1} \left( 1 - \frac{(1-x+\alpha)^2}{2} \right)^{N-1} dx - \int_\alpha^{1} (1-x+\alpha) \left( 1 - \frac{(1-x+\alpha)^2}{2} \right)^{N-1} dx \]

\[
= \sum_{k=0}^{N-1} (-1)^k \frac{C_{N-1}^k}{2^k} \int_\alpha^{1} (1-x+\alpha)^{2k} dx - \int_\alpha^{1} \left( 1 - \frac{(1-x+\alpha)^2}{2} \right)^{N} dx \]

\[
= \sum_{k=0}^{N-1} (-1)^k \frac{C_{N-1}^k (1-\alpha^2)^{2k+1}}{2^k (2k+1)} - \frac{1}{N} \left( \frac{1-\alpha^2}{2} \right)^N \left( \frac{1}{2} \right)^N
\]

Similarly after integration of \( P_N(x, \alpha) \) by distance \( \alpha \) the expected migration to destination with utility \( x \) is:
\[ P_{2,N}(x) = \int_0^1 P_N(x, \alpha) d\alpha = \int_0^1 (x - \alpha) \left[ 1 - \frac{(1-x+\alpha)^2}{2} \right]^{N-1} d\alpha \]

\[ = \sum_{k=0}^{N-1} (-1)^k \frac{C_{N-1}^k (1-(1-x)^{2k+1})}{2^k (2k+1)} - \frac{1}{N} \left( \left( \frac{1}{2} \right)^{N} - \left( \frac{1}{2} \right)^{N} \right) \]

Actually, there was no need to calculate this integral because one can do a linear replacement in order to derive \( P_{2,N}(x) \) from \( P_{1,N}(\alpha) \), which is \( x' = 1 - \alpha, \alpha' = 1 - x \). Also, note if \( x \) is replaced by \( 1 - \alpha \), formula (A19) is equivalent to (A18).

\[ P_{2,N}(1-\alpha) = P_{1,N}(\alpha) \]

Resulted probability of migration to the best destination among \( N \) is shown in the following proposition.

**Proposition A3.5.** \( P_N = \int_0^1 P_{1,N}(\alpha) d\alpha = \int_0^1 P_{2,N}(x) dx = \frac{1}{N} \sum_{k=1}^{N} \frac{(-1)^{k+1} C_N^k}{2^k (2k+1)} \)

Proof. In order to see this formula, first, rewrite the probability with the help of (A18)

\[ P_N = \int_0^1 P_{1,N}(\alpha) d\alpha = \int_0^1 \left( \sum_{k=0}^{N-1} (-1)^k \frac{C_{N-1}^k (1-\alpha^{2k+1})}{2^k (2k+1)} - \frac{1}{N} \left( \left( \frac{1}{2} \right)^{N} - \left( \frac{1}{2} \right)^{N} \right) \right) d\alpha \]

\[ = \sum_{k=0}^{N-1} (-1)^k \frac{C_{N-1}^k \left( 1 - \frac{1}{2k+2} \right)}{2^k (2k+1)} - \frac{1}{N} \int_0^1 \left( 1 - \frac{\alpha^2}{2} \right)^N d\alpha - \left( \frac{1}{2} \right)^N \]

Spectacularly, the first term is not a series but polynomial since it can be represented by a computable integral:
\[
\frac{1}{2} \sum_{k=0}^{N-1} (-1)^k \frac{C^k_{N-1}}{2^k (k+1)} = \frac{1}{2} \int_0^1 \left(1 - \frac{x}{2}\right)^{N-1} \, dx = \frac{1}{2} \int_0^1 \left(1 - \frac{x}{2}\right)^{N-1} \, d\left(1 - \frac{x}{2}\right) = \frac{1}{N} \left(1 - \left(\frac{1}{2}\right)^{N}\right)
\]

It was used here as:

\[
\int_0^1 \left(1 - \frac{x}{2}\right)^{N-1} \, dx = \int_0^{N-1} \frac{(-1)^k C^k_{N-1}}{2^k} \, x^k \, dx = \sum_{k=0}^{N-1} \frac{(-1)^k C^k_{N-1}}{2^k (k+1)}
\]

Finally:

\[
\frac{1}{N} \left(1 - \left(\frac{1}{2}\right)^N\right) - \frac{1}{N} \int_0^N \left(1 - \frac{\alpha^2}{2}\right)^N \, d\alpha = \frac{1}{N} \left(1 - \int_0^N \left(1 - \frac{\alpha^2}{2}\right)^N \, d\alpha\right) = \frac{1}{N} \sum_{k=1}^N (-1)^{k+1} \frac{C^k_N}{2^k (2k+1)}
\]

Q.E.D.

Now it is easy to note that the probability of migration is just one \(N\)-th of the migration rate:

\[
P_N = \frac{M_N}{N}
\]

This is not an unanticipated result given all destinations among \(N\) are equally possible due to their random nature.

The average distance moved by migrants and the average post-migration utility is the mathematical expectation divided by the probability of migration, i.e.:

\[
\bar{\alpha}_N = \frac{\int_0^1 \alpha P_N(\alpha) \, d\alpha}{P_N}
\]
\[
\bar{x}_N = \frac{\int_0^1 xP_{2,N}(x)dx}{P_N}
\]

\[
\int_0^1 \alpha P_{1,N}(\alpha)d\alpha = \frac{1}{\alpha} \left( \sum_{k=0}^{N-1} (-1)^k C_{N-1}^k \left( 1 - \alpha^{2k+1} \right) \right) + \frac{1}{N2^N} \left( \frac{1 - \alpha^2}{2} \right)^N d\alpha
\]

\[
= \sum_{k=0}^{N-1} \frac{(-1)^k C_{N-1}^k}{2^{k+1}(2k+3)} + \frac{N+2-2^{N+1}}{N(N+1)2^{N+1}}
\]

Inserting this result back to formula (A22) and using formula (A21), the average distance moved by migrants is calculated in the following clumpy view:

\[
\bar{\alpha}_N = \frac{N \sum_{k=0}^{N-1} (-1)^k C_{N-1}^k}{2^{k+1}(2k+3) + (N+1)2^{N+1}} \frac{N+2-2^{N+1}}{N(N+1)2^{N+1}}
\]

but can be simplified a bit to a more compact form:

\[
\bar{\alpha}_N = \frac{1}{2} - \frac{1}{N+1} \left( \frac{1 - 1}{2^{N+1}} \right) - \frac{1}{2} \sum_{k=1}^{N} \frac{(-1)^k C_{N}^k}{2^k (2k+1)}
\]

Proposition A3.6. The average utility of stayers and movers can be expressed as a function of the average distance for migrants and the migration rate

\[
\bar{z}_{\alpha,N} = \frac{1}{2} + \bar{\alpha} \circ \frac{M_N}{1-M_N}
\]

and
\[ \bar{z}_{m,N} = \frac{1}{2} - \bar{\alpha}_N \]

Proof. Rewriting (A14) with the help of (A24):

\[ \bar{z}_{x,N} = 1 - \frac{1}{N+1} \left( \frac{1 - 1}{N^{N+1}} \right) = 1 + \left\{ \frac{1}{2} - \frac{1}{N+1} \left( \frac{1 - 1}{N^{N+1}} \right) \right\} \frac{M_N}{1-M_N} \]

\[ = \frac{1}{2} + \bar{\alpha}_N \frac{M_N}{1-M_N} \]

This formula can be expressed as a function of \( N \) and \( M_N \):

\[ \bar{z}_{x,N} = 1 - \frac{1}{N+1} \left( \frac{1 - 1}{N^{N+1}} \right) \]

In a similar manner it follows from (A17) that:

\[ \bar{z}_{\bar{m},N} = 1 - \frac{1}{N+1} \left( \frac{1 - 1}{N^{N+1}} \right) = 1 - \left\{ \frac{1}{2} + \frac{1}{N+1} \left( \frac{1 - 1}{N^{N+1}} \right) \right\} \frac{1}{2} \]

\[ = \frac{1}{2} - \bar{\alpha}_N \]

Q.E.D.

**Proposition A3.7.** The sum of the average distance moved and the average post-migration utility of migrants is unity, \( \bar{\alpha}_N + \bar{z}_N = 1 \).

Proof. Using the relationship between \( P_{\bar{z},N}(x) \) and \( P_{\bar{z},N}(\alpha) \), derived in equation (A20), the integral in (A23) can be calculated from (A22) after replacing \( 1-\alpha \) for \( x \)
\[
\int_0^1 x P_{2,N} (x) \, dx = \int_0^1 (1-\alpha) P_{2,N} (1-\alpha) \, d\alpha = \int_0^1 (1-\alpha) P_{1,N} (\alpha) \, d\alpha = \int_0^1 P_{1,N} (\alpha) \, d\alpha - \int_0^1 \alpha P_{1,N} (\alpha) \, d\alpha \\
= P_N - \int_0^1 \alpha P_{1,N} (\alpha) \, d\alpha
\]

Therefore for the average post-migration utility:

\[
\bar{x}_N = \frac{\int_0^1 x P_{2,N} (x) \, dx}{P_N} = \frac{P_N - \int_0^1 \alpha P_{1,N} (\alpha) \, d\alpha}{P_N} = 1 - \bar{\alpha}_N
\]

Q.E.D.

**Proposition A3.8.** The average utility of migrants increases by a half, \( \bar{x}_N = \bar{z}_{m,N} + \frac{1}{2} \).

Proof is immediately derived from Propositions A3.6 and A3.7:

\[
\bar{x}_N = 1 - \bar{\alpha}_N = 1 - \frac{1}{2} + \bar{z}_{m,N} = \frac{1}{2} + \bar{z}_{m,N}
\]

Q.E.D.

In the end of this section the expected utility gain for an agent with initial utility \( z \) can be defined as:

\[
EUG (z) = EU (z) - z = E \{\Psi | \Psi > z\} P \{\Psi > z\} + zP \{\Psi < z\} - z \\
= E \{\Psi - z | \Psi > z\} P \{\Psi > z\} + 0P \{\Psi < z\} = \int_\zeta^1 (y - z) \, dF_\Psi (y) = \int_\zeta^1 (1 - F_\Psi (y)) \, dy
\]
where \( F_\psi(y) = \left(1 - \frac{(1-y)^2}{2}\right)^N \).

It is a decreasing function of \( z \) bounded by zero from below:

\[
EUG'(z) = -1 + F_\psi(z) < 0 \text{ for any } z, \quad 0 \leq z < 1; \quad EUG(1) = EUG'(1) = 0.
\]

Also it is an increasing function of \( N \) due to the positive sign of the derivative:

\[
\frac{\partial EUG}{\partial N} = \frac{\partial}{\partial N} \left(1 - zF_\psi(z) - \int_z^1 F_\psi(y) \, dy\right) = -\left(z \frac{\partial F_\psi(y)}{\partial N} + \int_z^1 \frac{\partial F_\psi(y)}{\partial N} \, dy\right) > 0
\]

because the derivative under the integral is negative:

\[
\frac{\partial F_\psi(y)}{\partial N} = \left(1 - \frac{(1-y)^2}{2}\right)^N \ln \left(1 - \frac{(1-y)^2}{2}\right) < 0
\]
5 Conclusions

Migration theories proposed by scholars in various disciplines tend to focus only on a few aspects specific to their own discipline. There is a need for a unifying theoretical approach which could combine theories and, more importantly, hypotheses across disciplines. Economics provides the rationalist approach, one of the most widely used ideas in research on migration, according to which an individual or household changes the place of living pursuing their own well-being interests. The idea of self-interest conforms well with theories employed by geography, demography, law, and political science.

In theoretical models, developed in the thesis, two economic ideas, self-interest of an individual and a notion of utility, were incorporated into the investment in human capital approach. The model of migration as investment in human capital, which goes back to Becker and Sjaastad and which will celebrate its fifty year anniversary soon, is applied throughout the thesis. The basic assumption in this type of model is about initial disequilibrium in a heterogeneous economy, that is there are potential gains in terms of utility for those considering the migration decision. Agents are keen to change their place of living if the expected post-migration utility exceeds the expected pre-migration utility. The expected utility in the destination takes into account all risks associated with the destination and costs of migration and information search. The migration decision can be represented as a solution to the expected utility maximization problem. Uncertainty about utility in both the origin and the destination is important in this decision.

However, there are still some unclear elements in this model which need to be clarified. The two most important ones are that researchers do not know how potential migrants select the list of potential destinations and how they collect information about destinations. The theoretical model presented in Chapter 4 addresses this problem in a manner common in economics. Potential
destinations are assumed to arrive randomly either in the sample of fixed size (the FSS search) or sequentially (the sequential search). The former approach is met in situations when information comes in clusters, for example, from statistical books. In the latter approach information arrives discretely. This situation is common in job search which often leads to either zero or few vacancies per period followed by zero or maximum one job offer.

In the FSS model, the migration decision is made at the end of a single period of search. The agent proceeds to migration only in the case of good luck, that is, when at least one destination from the drawn random sample gives a better utility than the initial location. Under an assumption of normally distributed costs of migration and utility both the probability of migration and the expected utility gain are proved to increase with costs of searching, the mean gain in utility, and both variances and to decrease with initial utility. However, in the case of generic (not normal) distributions the probability of migration increases with sample size but the expected utility gain is generally not monotonic. This leads to the idea that the exogenously given sample size could be less than optimal for many agents. Indeed, it was demonstrated that in the case of the endogenous sample size and concave, or not very convex, information costs function, the expected utility gain is maximised in one of the two corners. Those who are relatively wealthier do not invest in the information search but the remainders are willing to collect maximum information, that is, to spend all their budgets available for search.

In another model, with sequential information search of migration destination, duration of search is not restricted by one period but is a random variable. This model was shown to be not different from the model of job search or the minimum price search. The only difference between the models is that in the case of migration destination search there is a possibility to restore destinations met during the previous periods of search. A general conclusion for the model is that migration will be to the first destination, met during search, which has utility, adjusted by costs of search and
migration, above the reservation utility. The reservation utility depends on the distribution of utilities, costs, and the type of search and can be non-constant.

Future theoretical research could be to study the migration destination search in a general equilibrium framework. This idea is similar to equilibrium in job search models. Another interesting extension of the current model is to place it into a dynamic context, that is, to consider initial migration for some agents and repeat migration for others. Two groups of potential primary and repeat migrants may have in general different characteristics due to self-selection of migrants. Moreover, the repeat migrants are in addition characterized by psychic costs originated from the previous migration. If one models psychic costs as a function of distance from the initial location, such as the place of birth, then the primary migrants and the repeat migrants solve different optimization problems.

In the empirical models, estimated in Chapters 2 and 3 on the HILDA, the sole longitudinal household survey in Australia, causes and consequences of internal migration were studied. Chapter 2 constructed an empirical model of migration decisions on an individual level. It estimated a probit model for the likelihood of the migration decision on a panel sample of all adults from the HILDA survey excluding only full-time students. Migration was defined as a move of a distance over 30 kilometres, which is long enough to change the labour market in most cases.

Several economic variables were shown to be of high importance in migration decisions. In addition to the unemployed, who were always found more mobile than the employed, those who are not in the labour force are no less mobile. Results obtained with respect to wage indicate that even in such a higher income country as Australia there are financial constraints. Wage was found to have a positive impact on the likelihood of migration but those without wages are much more likely to migrate than those with low wages. Another result, obtained in the model for migration decisions, is
about the effect of remoteness on migration. Those who are living in remote areas are much more likely to migrate than others. In part this could be due to the elimination of psychic costs of living far from relatives and friends, the idea proposed in the theoretical section.

Many directions for further research were proposed in Chapter 2, two of which are the most intriguing. One is to model destination choice, for which the nested logit model is available but it may have limitations in case IIA property does not hold. Another is to introduce a spatial structure into the model.

In Chapter 3 economic consequences of geographical mobility were studied. The main focus was on the wage premium for migration. Earlier studies on the effect of income on migration usually did not consider potential endogeneity of the migration decision. Hypotheses developed from various approaches in Chapter 3 suggested that there is a higher wage premium for long distance moves. A dynamic wage equation was constructed on a sample of earning individuals from the HILDA survey. Descriptive analysis found a large wage premium for lower paid migrants and a negative premium for relatively higher paid migrants. The dynamic wage equation, estimated by system GMM which allows controlling endogeneity of the distance, demonstrated that in contrast to the higher paid workers who do not have returns to distance the lower paid workers earn a large premium for migration and that this premium increases with the level of education and distance moved.


- 208 -


