The role of distance in returns to geographical mobility: Evidence from HILDA survey[1]

Yury Andrienko[2]

Abstract

While income-distance trade-off has been well studied at the macro level, there has been little attention to individual returns to migration matched to distance of move. This paper studies wage returns to distance for internal migrants in Australia. Traditional human capital theory suggests the cost hypothesis, according to which returns / wage premium (defined as wage in destination less wage in origin) are a positive increasing function of migration distance because of higher costs associated with a move. The income opportunity hypothesis states that poor may move to a local but not to the global optimum because of liquidity constraints, therefore marginal returns to distance are not only greater than marginal costs but also diminishing with wage. Using individual level data from HILDA, I show the wage premium varies not only across distance moved but also with wage before a move and reasons of move. Applying a system GMM for dynamic panel earning model, I find positive returns to distance for individuals migrating for economic reasons with and negative returns for moving due to family reasons. Positive returns for economic migrants decrease with age and disappear for wages above median. Average returns to economic migration in Australia are estimated to be 4 percent in the short run increasing with distance and for poor.

Keywords: geographical labour mobility, migration, earning equation, dynamic panel data

JEL classification: J31, J61

---

[1] This paper 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 paper, however, are those of the author and should not be attributed to either FaHCSIA or the MIAESR.

[2] Economics, The University of Sydney. Address: Room 370 Merewether Building, University of Sydney, NSW 2006 Australia; e-mail: y.andrienko@econ.usyd.edu.au

This paper is an empirical part of my doctoral research. I sincerely appreciate the help of Russell Ross whose attention and support to my research activity is invaluable. All the remaining errors are my own.

Introduction

The mobility of the labour force is a phenomenon which has been paid a lot attention by economists. There are many types of mobility which can improve well-being, including human capital development (better education, occupation, and job) and other spheres of life (better house, suburb, and city). The returns to new jobs and education are positive (Krieg 1997, Trostel et al 2002). For example, the returns to an additional year of higher education in Australia are estimated to vary between 6 and 12 percent in terms of the wage depending on data and methodology (Preston 1997, Miller et al 1995 and 2006, Leigh and Ryan 2005). Economic literature has accumulated various research results about returns to (geographical) mobility.

The subject of this study is the impact of geographical mobility or migration on income.[3] The results from previous research have found the returns to migration are ambiguous. For example, earlier studies have identified a long-term negative effect of migration on income; see survey in Greenwood (1997). However, more recent studies based on individual level data have underlined the importance of a time effect in the human capital investment process and identified positive returns to migration within a few years following a move (Yankow 1999 and 2003, Boheim and Taylor 2007). For young workers migration is often well motivated by better career perspectives. However, not all migration is motivated by income gains. For example, retirees often move to their family roots or to a sea coast in the case of Australia. Even for working people if migration occurs because of family reasons such as marriage, divorce or rejoining the family one may expect even negative returns. The loss in income is compensated by non-monetary but valuable factors such as close relations with family and friends or a better environment and climate (amenities). Thus, there are many personal and economic reasons for a move and hence, there are likely to be a large variation in the returns to migration.

The distance people move during migration could explain some part of this variation. Distance is known to be a deterrent factor which reduces number of migrants between two locations. However, conditional on migration, wage should include a reward for every additional kilometre moved farther away. Income gains are expected to be increasing with distance people move. There are several reasons for this. First, this is because the costs associated with a move may be increasing with distance. People are willing to move if their expected wage at destination exceeds their reservation wage (i.e. the current or potential wage at origin) at least by the costs associated with moving. Costs include not only transport expenditures and the opportunity costs of time spent on a move but other costs such as search

[3] The words geographic mobility and migration are used interchangeably in this paper.

for information about potential destinations, psychic costs of missing family and friends, costs of regular travel back to the origin, communication costs such as telephone calls, and the opportunity costs of time. Some of these costs are non-permanent costs, e.g. expenditures on the information search and a house move but others are permanent. The further an individual moves the greater the costs s/he incurs and therefore demands higher income. Alternatively, if she moves to a place which substantially reduces costs (e.g. young person moving back to their parents' house where there is no need to pay rent) then this move can lead to lower income but still be economically rational. In this case wage gains could be negative. This supply side story is clearly demonstrated in reality by the existence of the supply of internal migrants for any local labour market.

Second, the supply of migrants is biased towards unsatisfied population which is often relatively poor individuals who are trying to find a better paid job or simply change their life. Economists think about labour as rational agents choosing locations by maximizing their utility subject to many constraints. As a result of this optimization poor are expected to be farther from globally optimal destination in contrast to rich because poor are more constrained to do search and moving (liquidity constraints hypothesis, see, for example, Andrienko and Guriev (2004)). Therefore, poor have higher marginal returns to distance than rich, *ceteris paribus* (income opportunity hypothesis). Keeping it simple, marginal benefits for a moving an additional kilometre is equal to marginal costs for rich but for poor marginal benefits exceed marginal costs. This hypothesis results in higher elasticity of wage with respect to distance for poor. This is also shown in a theoretical model developed in this paper. The income opportunity hypothesis means also that the positive effect of distance on wage is decreasing and may disappear after some threshold level of income is reached. These are arguments based on the supply side of the labour market.

Third, there is the demand side explanation. There are more opportunities to improve a wage with the greater distance (distance opportunity hypothesis, see for example Sjaastad (1962)). Distance elasticity of wage is not only positive but can even be increasing due to the spatial distribution of income. This can be shown under some simple assumptions. The further a migrant extends her / his search in a heterogeneous labour market the higher the wage and returns. This is in contrast to zero returns in a homogeneous labour market. For example, if a large number of similarly sized labour markets (cities) are located on a circle then the number of cities located within the circle of a given radius is proportional to the squared radius of the circle. Then, for the uniform distribution of income across cities, the expected returns are an increasing concave function of a search radius.

The decision to move depends on whether costs are less than an increase in income. The costs can be either low or high depending on the assumptions and type of costs. One can model costs of the move and psychic costs as linear functions of distance.[4] But total costs could be a quadratic function of distance if information search costs are assumed to be proportional to the number of destinations and information about all cities in the cirle of a given radius is collected and analysed.

Summarizing the hypotheses to be tested empirically, returns to migration are a positive function of migration distance with a decreasing rate for higher income.

This paper is organized into the following sections. First, I review the economic literature on migration with an emphasis on the returns to migration. Second, a theoretical model of optimal destination search is presented. Third, in the empirical part, I estimate a dynamic earnings equation as a function of distance and other explanatory variables using longitudinal data from 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 paper some concluding remarks are offered.

1. Literature

Traditionally economists and often other social scientists model the migration as an investment decision of a rational agent. Migration of an individual occurs if the expected benefit from a move outweighs the costs of the move. This is a traditional human capital approach which dates back to the seminal papers of Becker (1962) and Sjaastad (1962). The model of migration as investment to human capital has become the most popular and influential thereafter.

The relationship between distance and migration arrived in the literature in part due to gravity models. The Newtonian law of gravitation was found to be useful when describing not only migration but also other flows in international trade and transportation (see, for example, the article about gravity models in Wikipedia). According to this Law, the number of people migrating between two areas is proportional to both body masses equal to population size in each locality and the inverse distance between them. In modified gravity models, other push and pull factors for the source and host regions are added; including average income, unemployment rates, public goods, etc (see, for example, 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. Firstly, information is diminishing with

[4] Following an idea of Schwartz (1973) psychic costs are proportional to the transportation costs times to the number of visits to an old location within a year. Since psychic costs represent multiple visits, they could be much larger than costs of a move.

distance (Schwartz, 1973). Secondly, as Greenwood (1997) noted, costs of information search are an increasing function of distance. Therefore, it is costly to have the same level of certainty about more distant labour market conditions as compared to a close market. Thirdly, according to Schwartz (1973), psychic costs of moving are a positive function of distance as far as they could be offset by more frequent trips.

Generally, small distance moves which do not cross administrative unit borders are not considered as migration.[5] As a result labour migrants often change not only house but also job and commuting to previous address or job is generally unlikely. Unobserved skills and abilities introduce unobserved heterogeneity into empirical analysis since labour with different skills might have different returns to migration. One of the common results is that better educated people are more mobile and seem to move greater distances (Greenwood, 1997). Also, the income distance trade-off (elasticity of distance with respect to income) is higher for more educated migrants (Courchene, 1970). The reason mentioned in the literature is that "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).

Empirical evidence about returns to migration is still 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 income on mobility 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 HILDA supports this finding for international migrants (Belkar, 2005). There are methodological difficulties in the approach used in this type of studies 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 USA, Yankow 1999 and 2003; in Britain, Boheim and Taylor 2007). Returning (from overseas) males but not female migrants are shown to have a wage premium in Ireland (Barrett and O'Connell 2000).

[5] Migration is defined as a move on distances greater than some minimum or across administrative borders (UN 1970). I discuss a problem of distance based migration definition in more details in the empirical section below.

Even the negative effect of migration on income does not necessary contradict the rational choice theory because migrants may get some compensatory utility from consumption of other goods such as local amenities at the destination location. There are other reasons for a negative effect of migration on income which are generally ignored in analysis. Return migration may be followed by a loss in income. Also, analysis does not generally consider different price levels across locations.

There was some further development of theoretical approach in the migration literature. Some researchers treat the migration decision as undertaken not by individuals but rather by families. As a consequence of the theory, a household head and a 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. In some papers a sample for an econometric model is restricted to young males in order to get rid off problems with tied movers or stayers and restrict the effect of non-economics reasons of the move (Yankow 1999 and 2003, Boheim and Taylor 2007). In this paper I use a sub sample of youth to verify robustness of results.

It has been shown that returns to migration vary significantly across groups of a population (Greenwood, 1975). Bartel (1979) argued that job mobility should be taken into account in studying the consequences of migration. His results indicate that for young males' only transfers but not job quits and layoffs lead to higher wage in case of migration.

Another important side of the migration decision is the link with different personal characteristics and life cycle. One of the best established relationships is declining mobility starting in the mid-twenties age group (Greenwood 1997). Greenwood points out another universal relationship that better educated people are more mobile. Family factors were recognized as influential on the migration decision much earlier in non-economics disciplines (Lucas 1997).

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 (Courchene 1970 as cited in Greenwood 1997). 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) has estimated 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.

Selectivity of migrants is a potential problem recognized by scholars (Greenwood 1997). Greenwood in his survey mentioned 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 characteristics, e.g. demographic factors such as age, marital status, family size, migration history, etc. which characterize a population at risk of migration. Some of these variables may successfully identify migrants to have a significant effect in an "identification" equation but without that effect in a "consumption" equation. Avoiding selectivity problem may cause a selection bias in the empirical estimation. For example, Détang-Dessendre et al (2004) have studied the impact of migration on wages using two surveys for young French males. They found positive selection for highly educated males and no selection effect for lower educated. Unfortunately for my paper, a dynamic panel approach applied here can not address the sample selection problems outlined above. However, I reduce the effect of at least two sources of selection, (1) and (2), using balanced panel of respondents participated in all waves of the HILDA survey.

2. Theoretical model

A model constructed below is developed to describe the relationship between wage and distance moved by a migrant. It identifies the trade-off between a wage in a destination and the costs of a moving, the latter are assumed to be a positive function of distance. The trade-off means the higher costs of a moving, which means fro the empirical purpose the longer distance moved, are associated with a higher wage in a destination. The model shows the trade-off originates from spatial maximization of benefits, which is a wage less the costs of moving from an origin to a destination. Financial constraints and higher costs of move per given 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 and benefit.

2.1. Model of migration

Let's consider population of agents living in a point x which belongs to a set X. There is no any assumption on X, which could be either discrete or continuous set in some space. For some reasons population is heterogeneous only in costs of moving and budget available for a moving between points. Population residing in origin x is homogeneous in human capital meaning that everyone from origin x is equally productive and has the same wage in any point of X. However, wage varies across different points of X. For example, similarly educated people may have children and therefore, have less tangible, and live near parents, that is have higher psychic costs of moving. Wage and costs are exogenous in the model. Wage of an agent in the destination y, $y \in X$, is defined by a function $\mathrm{W(y)}$. For the sake of brevity other variables are omitted. Assume each agent from the origin x is free to choose a city y, migration destination, entertaining his/her profit due on migration, $\pi(x,y)$, and a positive budget (financial constraint) which is a part of wage earning in origin x available to finance a moving from x to y, B, $B>0$. Profit from the investment in a move is the wage in a destination y less the wage in the origin x and the costs of migration from the origin x to the destination y, $C(x,y)$

$$\pi(x,y) = \mathrm{W(y)\text{-}C(x,y)\text{-}W(x)} \qquad (1)$$

Profit must be non-negative in order to migrate.

The set of feasible cities is a set $Z = \{y \in X : C(x,y) \le B\}$, a subset of X, $Z \subset X$, which depends on the point x, the agent's costs function and budget, $Z = Z(x, C(x,\cdot), B)$

Verbally, an agent moves to a city with maximum profit from migration choosing among cities feasible by his budget constraint. Analytically, the profit maximization problem for any agent residing in a point $x \in X$

$$\pi(x,y) \to \max \text{ s.t. } y \in Z \qquad (2)$$

This is equivalent to the following problem of net wage maximization

$$W(y) - C(x,y) \to \max \text{ s.t. } y \in Z \qquad (3)$$

What if there is no such an economically attractive point y ? Then the agent remains in the point x. Therefore, the minimum possible set Z consists of the only point x, $Z = \{\mathrm{x}\}$. For simplicity, we may suppose decision not to move implies zero costs, $\mathrm{C(x,x)} = 0$, that is there are no fixed costs of moving. In the case when $Z = \{\mathrm{x}\}$ the optimal profit is equal to zero. However, it is possible to move to another point y if the optimal profit is equal to zero. In this case the agent is indifferent between living in x and moving to y.

Note, under the assumption of non-zero fixed costs of migration, say, search costs, $C(x,x) > 0$, the maximization problem is equivalent to the problem without fixed costs and with budget restricted by $B - C(x,x)$. Very poor agents who can not pay fixed costs will not be able to start the search, since their search set is empty $Z = \{\emptyset\}$. In the case of full information about wage in all cities and costs of moving to any city available to agents, all agents with wage below curtain threshold above $C(x,x)$ will be worse off paying fixed costs for search in the non-empty set $Z \neq \{\emptyset\}$, but not being able to find a better destination than the current point. Then it a rational decision for such agents not to do any search.

Now I describe solutions of the problem (3). In the general case there are multiple solutions to problem. The set of solutions Y* is a subset of Z, $Y^* \subset Z$. This subset is determined by the origin x and budget B in the origin. In addition, solutions depend on the cost function, the case considered below in Proposition 2. Operator A maps an element from Z into a subset of elements in Z, $A: Z \to Z$, $A(x,B) = Y^*(x,B)$. For any element y^*, $y^* \in Y^*$,

$W(y^*) - C(x,y^*) \geq W(y) - C(x,y)$ for any $y \in Z$.

This set is not an empty set, $Y^* \neq \{\emptyset\}$, since in the most simple case it consists of the origin x only, $Y^* = \{\text{x}\}$. Note, there are only one optimal returns for the entire set $Y^*(x,B)$. In the light of this observation we 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. In two propositions below I show that profit is increasing with higher budget and lower costs.

Proposition 1. The optimal profit is a non-decreasing function of budget in an origin. That is if two agents living in the same origin x have budgets such that $B_1 > B_2$, then profit for rich is at least as large as for poor $\pi^*(x,B_1) \geq \pi^*(x,B_2)$.

Proof is in Annex 2.

In order to study behaviour of the solution with respect to costs I define different types of costs function for agents living in one point. Assume there is an additional parameter θ, a type of the agent, which helps to distinguish between high and low costs individuals. Costs are increasing function of the parameter θ,

$$\frac{\partial C(x,y,\theta)}{\partial \theta} > 0 \qquad (4)$$

For this costs function the set of feasible cities, optimal set and optimal profit are functions of the parameter θ 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 in the same origin x have different costs such that $\theta_1 < \theta_2$, then profit for the lower costs agent is at least as much as for the higher costs agent, $\pi^*(x, B, \theta_1) \geq \pi^*(x, B, \theta_2)$.

Proof is in Annex 2.

2.2. 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 to migration is a positive function of budget, B, and a negative function of higher costs of a moving parameter θ

$W(y) - C(x, y, \theta) = f(B, \theta)$ (6)

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

Now, I do a series of simplifying assumptions to transform the model into a testable linear regression model. First, RHS of the equation is an additive function which consists of two linear components

$f(B, \theta) = \gamma \cdot B + \eta \cdot \theta$, where $\gamma \geq 0$ and $\eta \leq 0$ (7)

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

$B = \delta \cdot W(x)$ (8)

Forth, the costs function is not observed. It is modelled as the sum of a linear function of a type θ, and a second function, proportional to distance between x and y, $D(x, y)$.

$C(x, y, \theta) = \lambda \cdot \theta + \beta \cdot D(x, y)$ (9)

Fifth, a type θ and is also not observed. The linear function $(\lambda + \eta) \cdot \theta$ is approximated by a set of observed individual characteristics, $\sum_{i=1,\dots,M} \tau_i \cdot X_i$ or $\tau \cdot X$ in vector terms, unobserved individual characteristics, λ_i, regional dummies υ_x, and macroeconomic effect, μ_t

$(\lambda + \eta) \cdot \theta = \tau \cdot X + \lambda_i + \mu_t + \upsilon_x$ (10)

Altogether, the equation (6) is transformed with help of assumptions (7)-(10) into the empirical model

$w(y) = \alpha \cdot w(x) + \beta \cdot D(x, y) + \tau \cdot X + \lambda_i + \mu_t + \upsilon_x + \varepsilon$ (11)

In more familiar form this model has the following specification which is more relevant to panel data I am going to explore

$w_{it} = c + \alpha \cdot w_{it-1} + \beta \cdot D_{it} + \tau \cdot X_{it} + \lambda_i + \mu_t + \upsilon_{it} + \varepsilon_{it}$ (12)

where i stands for an individual, t is a year, D_{it} is a distance moved by the individual this year, and υ_{it} is a regional dummy variable.

In this model distance may well be a predetermined variable. Distance and wage are probably determined simultaneously since both are characteristics of the optimal destination. Distance is observed before a move but total costs and benefits from migration are realized with some lag in time. This model has a similar view to Mincer earning equation.

3. Empirical analysis

The econometric analysis reported here is based on micro level data. I estimate a Mincer-type equation for the individual earnings with a set of control variables which include personal characteristics common for the earnings equation including unobserved individual heterogeneity and unobserved regional characteristics. Dynamic panel data technique (Arellano-Bond method) is used to find unbiased estimates of the model. In this section, firstly, I describe the methodology of the econometric estimation technique. Secondly, I explore the HILDA migration data and provide descriptive statistics for the sample of individuals analysed. Third, I apply System GMM method for the dynamic model on panel data assuming autoregressive term to be predetermined. In the beginning I assume the variable of interest, distance, is exogenous and then do the similar regression analysis assuming distance to be predetermined.

3.1. Methodology

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

$$W_{it} = \alpha + \rho \bullet W_{it-1} + \beta \bullet X_{it} + \gamma \bullet D_{it} + \lambda_i + \theta_t + \mu_{it} + \varepsilon_{it}$$

Where W_{it} is a current year ‘t’ wage of a person ‘i’ , α is a constant term, W_{it-1} is the person’s wage at year ‘t-1’, X_{it} is a vector of individual characteristics at year ‘t’, D_{it} is the great-circle distance moved between years ‘t-1’ and ‘t’, λ_i is the individual ‘i’ fixed effect, θ_t is a time dummy (unobserved macro economic effect at year ‘t’), μ_{it} is a regional dummy (unobserved effect of a host region), and ε_{it} is an error of the model.

OLS method is not explored for the dynamic model estimation due to serial autocorrelation of errors. Another problem arising here is with the autoregressive term which appears to be predetermined; that is affected by 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 lagged error which in its turn is positively correlated with the current realisation of the error. In this case the core assumption for consistency of OLS estimate is violated.

Distance variable is assumed to be, first, exogenous and then predetermined in the second specification of the model. Exogeneity assumption may hold for some group of migrants who are moving because of non-economic reasons and are not concerned about income in a destination. For another group of economically motivated migrants exogeneity is not generally true. Instead, I assume that distance is a predetermined variable since migration decision was based not on the actual wage but rather on the expected wage in a destination. The expected wage is based on the (available information about) probability of been employed and the distribution of wages for similar workers. This decision was done at latest sometime between the current and previous waves of the survey that is well before we observe an individual's wage after migration.

In model specification with reasons of move I introduce interactive terms of distance and dummy for reasons of move. Similar to distance three interactive variables are assumed to be predetermined. In addition to that, in the final model specification I consider another predetermined interactive term for distance and job related reasons of move interacted with the autoregressive term. Another predetermined variable is the number of working hours in the model for weekly wage.

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. System GMM provides more efficient estimator as compared to 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, which can be derived for any predetermined variable including the lagged dependent variable. The moment conditions are based on the fact that correlation of level for predetermined variable and its lagged first difference is zero as well as correlation of its first difference and second lag of the level. Instruments used in the estimation are obtained from these moment conditions. The set of instruments for the equation in levels includes lagged first differences for predetermined variables while instruments for the equation in first differences consist of the second lag of predetermined variables. In addition, all exogenous variables are used as instruments for themselves. Following recommendations of Roodman (2006) in his user guide for STATA routine xtabond2, explored in my analysis, I add time dummies to the model specification since with them an assumption of no correlation across individuals in the idiosyncratic disturbances is more likely to hold. Also, I use a number of options with the routine xtabond2 such as orthogonal deviations in order to maximize sample size in a panel with gaps, two-step procedure which is asymptotically efficient and robust but may result in downward biased standard errors, and a small-sample correction to the covariance matrix estimate.

There are two tests on overidentifying restrictions imposed by these assumptions on moment conditions. The Sargan and Hansen tests verify joint validity of a list of instruments, separately for the equation in first differences and system. Arellano and Bond test is used to test autocorrelation in disturbance term. This test is used to verify validity of lags as instruments. In the models estimated below I found first order serial autocorrelation of errors but no second order serial autocorrelation what Arellano-Bond test is meant to demonstrate. Sargan test in some cases rejects the hypothesis of no overidentifying restrictions. In that cases I did robustness check reducing the number of instrument (limiting the lags used in GMM-style instruments, option gmmstyle() in xtabond2) or collapsing instruments (collapse option in xtabond2).

3.2. Data

The data analysed here is longitudinal data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The survey of a nationally representative sample of households is held from August each year to March the following year starting from 2001. For this draft of the paper I am using data for the six waves from 2001 to 2006. The main advantage of the survey is that it tracks the initial sample of individuals even in case of 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).

There is a drawback in the survey. Presumably attrition of the panel sample is to a large extent due to not followed migrants. As a result the survey is potentially not representative for migrants, and this problem is aggravated in the course of time because it does not attempt to replace these migrants by other comparable people and households. The attrition problem is reduced in the analysis below which uses individual longitudinal weights for balanced panel in waves 1 to 6.

3.2.1. 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. Since the model of interest is the earning equation, the sample is restricted on those who report wage.[6] The model studied here is a dynamic model. Therefore, in order to be in the sample an individual must report wages for at least two subsequent waves. The sample selection problem arising here is not studied in this paper because the methodology for the panel data sample selection is not very successfully realized in practise yet.[7] Two measures of earnings are considered in this paper. The first is a real weekly wage measured in the constant 2001 Australian dollars recalculated as reported current weekly gross wages and salary from all jobs (derived variable wsce in HILDA data set) divided by the state cumulative CPI.[8] The second is a real hourly wage defined as a real weekly wage over the total number of hours per week *usually* worked in all jobs (reported in variable jbhruc). Some inaccuracy in measurement of wage exists because a week preceding an interview can be not representative, that is *not usual* for wages from all other jobs.

The six waves of the HILDA survey contain weekly wage and hours of work for 42,201 respondents out of 108,035. Out of that I use 32,354 of hourly wage observations that are in a balanced panel for waves 1 to 6. In order to minimize effect of outliers, I restrict sample on individuals with the real (2001 AUD) wage above 3 and below 200 dollars per hour. This is done by exclusion of less than 1 percent of observations, 0.7 percent observations from the left tale of the hourly real wage distribution in the pooled sample and 0.07 percent observations from the right tale. After that I leave in the sample working age individuals who are in the age group above 15 and below 65 for males or 60 for females and are not full time students. Since the model studied here is a dynamic model, about 40 percent of the sample is lost due to a missing value for a wage in a previous wave. Finally, the total number of observations is 19,116. All calculations in Tables and regression analysis below are based on a sample of the individual observations using responding person longitudinal weight balanced for waves 1 to 6, variable wlra_f.

3.2.2. Independent variables: distance

[6] I am using the usual weekly wage as 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.

[7] Although there are methods in Kyriazidou (1997, 2001) atricles, they are not widely used due to poor behaviour of estimates. An alternative method which can be potentially used to attack the problem here is found in Wooldridge (2002).

[8] CPI is taken for the last quarter current year to the last quarter last year or December, assuming December 2001 quarter CPI=1. Data comes from Consumer Price Index (cat no. 6401.0) available on the Australian Bureau of Statistics website www.abs.gov.au

A common definition of migration is a change of usual place of residence or move above some specified minimum distance (UN, 1970).[9] Such a wide definition as "usual" place of residence hides measurement errors. The first source of a measurement error when migration on large distances within an administrative unit is not included in the statistics. The second source is very short moves only crossing the administrative border which are counted as migration.

Most of migration occurs within a small locality and is not associated with a change of job but rather is for housing reasons (see the literature on residential mobility versus labour migration, for example, Clark and Huang (2004)). Relatively low proportion of moves results in a change of larger administrative unit (city, region) and much lower proportions involve crossing of country borders. There are many reasons of migration and they are different for local moves and intercity / interregional migration. Personal and family reasons, followed by housing related reasons, then job and education were the most common reason for moving in the 2003 wave of the HILDA Survey (Headey et al (2006)).

I found 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 Mohlo (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 et al. (2001), Clark and Huang (2004)) a long-distance move is a move over 50 kilometres. For authors in the article Boyle et al. (2001) it seems likely that most of migrants moving above 50 kilometres are motivated by employment factors.

For practical purposes in this paper I define migration using the distance based definition. I assume migration is a move on distance over 30 kilometres as these movers are more likely to change their labour market. In the sample used for the empirical model I found 697 moves over 30 kilometres. Forasmuch as there are 19116 individual-year observations in the sample, long-distance migrations constitute 3.6 percent of all observations.[10] Only 35 moves out of 697, that is 5.0 percent, were within a city boundary if we adopt the definition of Major Statistical Region for the largest cities used in HILDA data (variable hhmsr). Nonetheless, these intra-city moves are of no particular importance, since exclusion of them

[9] 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 the later draft.
[10] Unweighted sample provides a higher proportion of long distance migrants than that proportion reported in Table 1.

from migration category does not change the results of the empirical analysis below as my additional robustness check has demonstrated.

My primary interest is in the accurate measure of distance for those individuals who have changed address and were followed up. The HILDA survey provides a distance measure on individual level beginning from the second wave in 2002 (variable hhmovek). It is the great-circle distance, measured in kilometres, between centroids of postcodes of current and previous interviews residential addresses. There is no detailed information about in-between moves if a respondent moved more than once between interviews (we know only when s/he moved to current address). Since there is no data about the area of a postcode in Australia, I assume the distance to be zero for stayers and those who have moved within a postcode.

There is additional evidence that distance may not be precisely measured since there are two types of errors found in 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.

Distance varies between 1 and 3649 kilometres for interpostcode movers. Interestingly but not surprisingly, the number of moves is a decreasing function of distance which can be described by discrete probability (Poisson) distribution function. Thus, out of 3,038 inter-postcode moves the number of moves on distance 1 km is 407, on 2 km 292 moves, on 3 km 261 moves, on 4 km 211 and so on.

Descriptive tables below are constructed from individual observations weighted by the longitudinal sample weight. Table 1 reports distribution of the sample by distance moved. Intra-postcode movers are treated as stayers. One can note a sizable proportion in the range 13-16 percent of the sample moves between any two waves. Most of these moves, 12 percent, can be classified as the 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. About 3 percent of population annually migrate on long distance over 30 kilometres,. These long distance moves represent roughly moves between labour markets which are often connected with new job. In particular, I found 11 percent of the stayers in the sample have changed an employer since a previous interview as compared to 17 percent of the short distance movers and 54 percent of those who moved on the long distance (not reported in Tables).

Table 1. Distribution of the sample by distance of move, percent

Wave	<=30km	>30km	Stayers	Total
2	12.2	3.5	84.4	100
3	13.3	3.0	83.7	100
4	11.8	3.4	84.8	100
5	11.6	3.2	85.1	100

6	9.8	2.8	87.5	100
Total	11.7	3.2	85.1	100

Analysis of mobility among different income groups shows non-linear pattern, see Table 2. Wage quartiles are defined for each wave according to weekly real wage distribution in the sample. The first, lowest, and the forth, highest, income quartiles consist of more stayers then the two other income groups. This is U-shaped income effect for stayers. The most mobile group of population is in the second income group. For short but not long distances we observe inverse U-shaped income effect. The richest quartile of population is the least mobile. This is in accordance with income opportunity hypothesis. In contrast, the poorest quartile potentially could be the most mobile because of more alternative opportunities. But in reality we observe 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.

Table 2. Distribution of the sample by wage quartiles and distance of move, percent

Wage quartile previous year	<=30km	>30km	Stayers	Total
1	11.2	3.0	85.9	100
2	13.5	3.4	83.1	100
3	12.2	3.1	84.8	100
4	9.9	3.3	86.8	100
Total	11.7	3.2	85.1	100

In the econometric analysis the measure of distance represents only long distance moves above 30 kilometres. Distance for stayers and short distance moves up to 30 kilometres is equal to zero and is the log of distance in kilometres for long distance migrations above 30 kilometres.

3.2.3. Other independent variables

Here I describe how other variables are constructed.

Reasons of move, *Job*, *Family*, and *Other*, are dummy variables derived from an answer to a question "What were the main reasons for leaving that address?". *Job* reasons include answers (here and below in diminishing frequency order): to start a new job with a new employer, work transfer, to look for work, to start own business, work reasons (NFI). *Family* reasons: to get a place of my own/our own, to get married/moved in with partner, to be closer to friends and/or family, marital/relationship breakdown, to follow a spouse or parent/Whole family moved, personal/family reasons (NFI). *Other* reasons: to get a larger/better place, property no longer available, seeking change of lifestyle, to be nearer place of work, to get a smaller/less expensive place, to live in a better neighbourhood, other (specify), temporary relocation, to be closer to amenities/services/public transport, housing/neighbourhood reason (NFI), to be close to place of study, health reasons. Since there are multiple answers to the question about reasons, I assume that job reasons dominate family reasons and family reasons dominate other reason, that is dummy for family (other) is always zero if dummy for job (family) is one.

Education is the number of years of formal education which corresponds to the maximum educational attainment (variables edhigh and edhist). I assume it to be 17 for postgraduates - masters or doctorate, 16 for those who have graduate diploma or graduate certificate, 15 for bachelors, 13 for advanced diploma or diploma, 12 for year 12 or certificate III or IV, 11 for year 11 or certificate I or II, 10 for year 10 or certificate not defined, 9 to 7 for year 9 to 7, 4 for the answer ‘did not attend secondary school but finished primary school’, and 2 for the answer ‘attended primary school but did not finish’.

Tenure is the number of years in current occupation. It is derived from the variable jbocct, which is an answer to the question “How long in total have you worked in your current occupation? Include time spent in this same occupation with previous employers or in previous businesses”.

Age is age of a person, variable hhiage.

Sex is gender of a respondent, equal to 0 for females and 1 for males. It was recalculated from the variable hgsex which is originally 1 for males and 2 for females.

Marital status, 0 for never married, separated, divorced or widowed, 1 legally or de facto married. Derived from the variable mrcurr.

Number of children below 15 years of age is defined as the number of household members less the number of adults above 15 years, variables hhpers and hhadult respectively.

Working hours are equal to the variable jbhruc, hours per week usually worked in all jobs.

Firm size is a dummy for the size of a firm for the main job. I constructed 4 categories from the variable jbmwpsz, which is the number of workers employed in current place of work. Four binary variables correspond to a firm size (1) 1-10 workers, (2) 11-50, (3) 51-500, and (4) above 500.

As *regional dummy* variables I use binary (dummy) variables for the 55 statistical divisions in Australia, variable hhsd.[11]

[11] A Statistical Division is an Australian Standard Geographical Classification defined area which represents a large, general purpose, regional type geographic area. They represent 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. They consist of one or more Statistical Subdivisions and cover, in aggregate, the whole of Australia without gaps or overlaps. They do not cross state or territory boundaries and are the largest statistical building blocks of states and territories (Census Dictionary, 2006 (Reissue) (cat no. 2901.0) available at www.abs.gov.au)

Other individual dummy variables are 37 *occupational categories* (jbmocc2) and 53 *industries* (jbmind2) based on opened questions. Both variables were originally coded in the data set using two-digit classifications.

3.3. Descriptive analysis

First, I compare individual characteristics of migrants on different distances. Table 3 reports mean values of independent variables included in econometric model. Thus, average wage weekly in 2002-2005 waves measured in constant 2001 dollars was 820, increasing from 792 in 2001-2004 waves. These numbers make sense, for comparison Australian Bureau Statistics reports average weekly earnings in November of the same years adjusted by CPI have changed from 695 to 707 and from 918 to 938 for adults working full time.[12]

The Table 3 reveals short distances movers have comparable average characteristics to long distance migrants and quite different from stayers. Thus, movers are better educated, younger, more often unmarried, have lower occupational tenure and less children. At the same time wages and working hours are different between these three groups. Though long distance migrants earn weekly wages before and after migration higher than stayers, this is mostly because of more hours of work. On the contrary, movers on short distances earn less than stayers before move and after move even though they work more time. Another observation from Table 3 and especially clear from Tables 5-5c, geographic mobility helps wages to converge. There is evidence of wage convergence equalization due to migration from macro level analysis (see literature on regional income convergence, e.g. Blanchard and Kats, 1992).

Job related reasons of move are miserable among short distance movers who explain their moves by family and other reasons. At the same time these three categories of move share equally, by one third each, the reasons of long distance migration.[13]

Table 3. Descriptive statistics, movers vs. stayers

	<=30km	>30km	Stayers	Total
Weekly wage, previous year, 2001 AUD	774	802	794	792
Weekly wage, current year, 2001 AUD	797	862	821	820
Weekly working hours, previous year	40.6	41.9	39.6	39.8
Weekly working hours, current year	41.0	42.0	39.9	40.1
Hourly wage, previous year, 2001 AUD	19.1	19.3	20.1	20.0
Hourly wage, current year, 2001 AUD	19.5	20.4	20.8	20.6
Distance, km	7	707	0	23
Education, years	12.7	12.8	12.4	12.5
Tenure, years	6.4	5.9	10.1	9.5
Age	33.8	33.8	40.9	39.9
Sex, 0/1 f/m	0.58	0.61	0.56	0.56
Marital status, 0/1	0.59	0.56	0.70	0.68

[12] Calculated from time-series spreadsheets, Average Weekly Earnings (cat no. 6302.0) and Consumer Price Index (cat no. 6401.0) downloaded from www.abs.gov.au

[13] Non-zero frequency for reasons of move among stayers is due to intra-postcode movers. I placed them in stayer category because of difficulty to separate them from stayers in the data.

No. of children 0-14 years	0.50	0.50	0.65	0.63
Reasons of move: job	0.01	0.32	0.001	0.01
Reasons of move: family	0.42	0.34	0.01	0.07
Reasons of move: other	0.57	0.33	0	0.08

I proceed to data description which can support the hypothesis on positive effect of distance on wage premium. For a moment I relax the *ceteris paribus* assumption and compare returns to distance across income groups. Table 4 compares growth of mean wage for stayers and movers by distance (columns) and initial income group (rows). Each cell of the Table shows a 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 income growth. Headey and Warren (2008) have 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 4 demonstrates, the highest proportion of gainers, 80 percent, is observed among long distance migrants in the lowest income group. The lowest proportion, 43 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 12 percent, from 68 to 80. This is also confirmed from Table 5.

Table 4. Proportion of real weekly wage gainers, by quartiles and distance moved, percent

Wage quartile previous year	<=30km	>30km	Stayers	Total
1	73	80	68	69
2	62	57	59	60
3	54	58	54	54
4	44	43	48	47
Total	58	58	57	57

One can look at income returns to distance from another angle, finding a change of mean real wage by income groups and distance comparing average income growth in particular category before and after move across all waves, see Table 5. Calculations are based on two tables (say, similar to Table 4) of average wages in particular category by income group and distance. On average real wage growth is 3.5 percent. By and large, conclusions are similar to those from the previous Table. Short distance movers do not differ markedly from stayers. Only lower paid labour with wages below median can significantly improve wage by changing labour market. This is especially true for the lowest quartile which long distance migration leads to 49 percent income growth as compared to 29 percent for stayers. Negative income growth is something of surprise. However, again report of Headey and Warren (2008) presents another evidence of downward income mobility for top income group. Using the HILDA data the authors constructed transition matrix for equivalised income, according to which 42 percent of the (top) fifth income quintile moved down into the lower quintiles between 2001 and 2005.

Table 5. Growth of mean real weekly wage, by quartiles and distance moved, percent

Wage quartile previous year	<=30km	>30km	Stayers	Total
1	32.8	49.2	29.0	30.2
2	9.1	15.4	7.4	7.9
3	2.0	9.8	3.8	3.8
4	-6.0	-4.4	-3.1	-3.4
Total	3.0	7.5	3.4	3.5

In part the positive income growth for poor and negative for rich shown in the previous Table is due to changes in both hourly wage and working hours. Indeed, Table 5a reveals a positive change in hourly wage for poor and negative for rich which mostly explains weekly wage growth in both income groups. On the other side, the poor have increased hours of work substantially while the rich have somewhat reduced them, see Table 5b. The largest growth of working hours, 18 percent, from 30.1 to 34.4 hours, is observed among migrants on long distance with the lowest income (it is worthy of note that they still are part-time employed on average), whereas highest income migrants experience the largest fall, 6 percent, from 49.1 to 46.8 hours per week. Table 5c shows, however, annual wage fall is not so dramatic for the top income group. Actually, long distance migration has led to wage growth in any income group.

Table 5a. Growth of mean real hourly wage, by quartiles and distance moved, percent

Wage quartile previous year	<=30km	>30km	Stayers	Total
1	26.9	55.4	27.9	28.5
2	5.1	9.3	6.3	6.3
3	0.9	2.6	2.6	2.4
4	-7.3	-6.4	-6.0	-6.1
Total	2.4	6.3	2.7	2.8

Table 5b. Growth of mean working hours, by quartiles and distance moved, percent

Wage quartile previous year	<=30km	>30km	Stayers	Total
1	12.0	18.3	9.5	10.1
2	1.7	-0.2	0.1	0.3
3	-0.1	-4.1	-0.9	-0.9
4	-1.7	-5.6	-1.3	-1.4
Total	1.8	-0.7	0.8	0.8

Table 5c. Growth of mean real annual wage, by quartiles and distance moved, percent

Wage quartile previous year	<=30km	>30km	Stayers	Total

1	67.0	96.7	51.0	54.6
2	9.8	10.0	7.7	8.1
3	5.3	3.9	3.4	3.6
4	-5.7	1.3	-0.2	-0.7
Total	7.3	12.7	6.2	6.5

3.4. Econometric analysis

An extract from the System GMM results for the dynamic model is given in Table 6. More detailed results are available in Table A, Annex 1. Table 6 reports only the distance elasticity of wage. At the first go-of I consider the model in which distance is assumed to be exogenous. Its results are under model (1) for weekly and hourly wages in Table. Distance of migration is found to have a positive effect on both weekly and hourly wages. In both cases coefficients are significantly different from zero, at 1 percent confidence level.

Table 6. Wage equation, system GMM results

	Weekly wage, log (W(t))			Hourly wage, log (W(t))		
	(1)	(2)	(3)	(1)	(2)	(3)
Distance, log (Dist)	0.009	-0.001	0.063	0.015	0.0002	0.059
	[0.003]***	[0.003]	[0.031]**	[0.004]***	[0.003]	[0.019]***
Dist*W(t-1)			-0.010			-0.020
			[0.005]**			[0.006]***

Standard errors in brackets
* significant at 10%; ** significant at 5%; *** significant at 1%

However, under different assumption results change. Model (2) in Table shows estimation results for predetermined distance. Hypothesis about no effect of distance on wage can not be rejected for both weekly and hourly wages. Distance is found to be insignificant for wages.

In model (3) in addition to predetermined distance I include another predetermined variable, the interaction between distance and autoregressive term of the wage equation. Both variables, distance and interactive term, are found to be statistically significant in both earning equations, on 5 percent for weekly wage and 1 percent for hourly wage. Results indicate a positive effect of distance and a negative impact of the interactive term between distance and wage. Positive effect of distance is largest for the lowest wages and then gradually disappears with higher wage levels. Absolute value of ratio of the coefficients for distance over the interactive term is equal to $0.0632/0.0095 = 6.6$ and $0.0587/0.0200 = 2.9$ for weekly and hourly wage models, respectively. Values 6.6 and 2.9 correspond to medians for (log of) weekly and hourly wage distributions meaning zero returns to migration for a median worker. Lower income half of labour is earning a wage premium for long distance migration.

The next step is to introduce reasons of move into earning equation. In model (4) I consider interactions of distance with a dummy variable for each of three categories of reasons: job, family, and other reasons. Each interaction is modelled as a predetermined variable. Results in Table 7 show for job related reasons distance significantly raises wages, while for family reasons distance significantly declines them. Also, the effect of distance for other reasons is insignificant. These results hold for both wages, per week and per hour. Hypothesis about equal (in absolute value) effect of job and family reasons on distance elasticity can not be rejected.

Table 7. Wage equation with reasons of move, system GMM results

	Weekly wage, log (W(t))		Hourly wage, log (W(t))	
	(4)	(5)	(4)	(5)
Dist*Job	0.008	0.172	0.007	0.083
	[0.003]***	[0.029]***	[0.003]**	[0.021]***
Dist*Family	-0.008	-0.008	-0.010	-0.010
	[0.004]**	[0.004]**	[0.004]**	[0.004]***
Dist*Other	-0.001	-0.002	0.0001	0.001
	[0.005]	[0.004]	[0.005]	[0.005]
W(t-1)*Dist*Job		-0.025		-0.026
		[0.004]***		[0.007]***

Standard errors in brackets
* significant at 10%; ** significant at 5%; *** significant at 1%

Finally, in model with reasons of move I introduce an interactive term between lagged wage and distance and job reasons. Results of this model (5) are also reported in Table 7. Economic migrants motivated by job reasons still have significantly positive effect of distance on wages but for them interaction of distance with wage is statistically negative. Recall from Table 6 distance elasticity of wage is decreasing with wage. Table 7 concludes that this is true only for economically motivated migration but not for family and other reasons of move.

If we take, for example, a migrant moving on distance 720 kilometres, which is the great-circle distance between Sydney and Melbourne, returns to economic migration are about 5 percent while returns to migration induced by family are negative 5 percent. Since the average log of long distance migration in the sample is 5.7 then the average wage premium is equal to 4 percent for economic migrants and negative 4 percent for family motivated migrants which are more than the effect from an additional year of education, 3 percent, according to the regressions I estimated in this paper. These average estimates for returns vary depending on an individual wage level. Poor people have higher returns to economic migration but rich have lower returns which may happen to be negative.

3.5. Robustness check

I did a number of checks to verify robustness of estimates. First, I extended the measure of distance for all movers including distances up to 30 kilometres. Now not only migrants over long distances but inter-postcode movers on short distance have a non-zero value of distance variable. Estimated coefficients for variables of interest in models (4) and (5) reported in Table 8 did not change significance from those in Table 7. For distance interacted with other reasons coefficients are not significant and not reported. Noticeably, distance elasticity in model (4) has increased approximately by 30 percent for both, job and family reasons. Say, job related move on 10 kilometres increases wage on average by 2 percent. This result is probably counterintuitive because moves on short distance should not result in significant change in earning in contrast to long distance moves.

Table 8. Wage equation with all movers, system GMM results

	Weekly wage, log (W(t))		Hourly wage, log (W(t))	
	(4)	(5)	(4)	(5)
Dist*Job	0.010	0.162	0.008	0.073
	[0.003]***	[0.030]***	[0.003]**	[0.022]***
Dist*Family	-0.012	-0.012	-0.014	-0.014
	[0.004]***	[0.003]***	[0.004]***	[0.004]***
W(t-1)*Dist*Job		-0.023		-0.023
		[0.005]***		[0.008]***

Second, I estimated the model (4) on different samples using long distance migration as in part 3.4. These include sample of males, young people with age below 36 years, young females less than 36 years of age, old people above 45 year old, poor with wage below median wage, rich with age above median wage. On other samples such as sample of females and young females system GMM estimates are not plausible and not reported here. Model (4) estimation results on different samples are collected in Table 9.

Table 9. Wage equation on different samples, model (4), system GMM results

	Weekly wage, log (W(t))						Hourly wage, log (W(t))				
	Males	Young	Young females	Old	Poor	Rich	Males	Young	Young females	Old	Poor
Dist*Job	0.004	0.01	0.016	0.019	0.014	0.006	0.006	0.009	0.019	0.036	0.011
	[0.003]	[0.004]***	[0.005]***	[0.007]***	[0.005]***	[0.003]*	[0.003]*	[0.004]**	[0.006]***	[0.007]***	[0.005]**
Dist*Family	-0.005	-0.005	-0.01	-0.002	-0.013	0.002	-0.004	-0.006	-0.006	-0.003	-0.013
	[0.004]	[0.004]	[0.006]*	[0.009]	[0.005]***	[0.005]	[0.005]	[0.005]	[0.006]	[0.006]	[0.004]***
Dist*Other	-0.007	0.003	0.007	-0.012	-0.009	0.001	-0.004	0.003	0.013	0.008	-0.004
	[0.006]	[0.006]	[0.006]	[0.008]	[0.005]*	[0.005]	[0.006]	[0.006]	[0.005]**	[0.007]	[0.005]
Observations	10120	6148	2742	7441	9546	9570	10120	6148	2742	7441	9562
No of persons	2658	2067	999	2234	3520	3034	2658	2067	999	2234	3624

Regression results obtained for males and richest half of the sample are not significant for both weekly and hourly wages. This is not surprising that economically motivated migration does not lead to gains in wage for rich if we recall results from model (5) in Table 8. In contrast to rich, results for poor are even stronger than those in original model (4). Elasticity for job and family reasons are much high for poor than for rich. Similar to part 3.4.

calculations show that economic migration of poor results on average in 8 percent wage growth which is the double effect of economic migration for an average person in the entire sample. The largest elasticity of wage with respect to distance for job reasons is found for old people, somewhat unsirprisingly because this category is known to be the least mobile.

In the third robustness check I excluded self-employed from the sample and got similar results. Forth, I used the Windmeijer (2005) small-sample correction for the two-step standard errors. Without such a correction standard errors tend to be severely downward biased. I found that new standard errors do not differ much from the original and as a result, significance of the estimates is not affected.

4. Conclusions

In this paper the hypotheses concerning the effect of geographical mobility on wage premium were studied. The costs hypothesis predicts a higher wage premium for long distance moves because of various distant dependent costs associated with a move. According to income opportunity hypothesis a poor may have positive benefits of migration increasing with distance. This is because of liquidity constraints which make a poor more likely to be in his local optimum but not in global one.

In order to test the hypotheses I explored individual level data from the HILDA, longitudinal survey of Australian households. Descriptive analysis demonstrates a relatively lower paid is awarded by large wage premium for moving on long distance whereas a relatively better paid is not. Results obtained from System GMM estimation of dynamic wage equation on panel data seem to confirm both hypotheses for economically motivated migrations. The distance elasticity of wage, weekly and hourly, is positive while the interactive term between wage and distance is negative for job related migrations. Labour with wage below median earns a premium if it migrates on long distance. Average returns to migration in Australia are estimated to be 4 percent. This number is increasing with distance and for poor.

Finally, some possible extensions of the empirical analysis I am planning to do (in the decreasing order of importance)

- Show that earning equation for migrants is similar to non-migrants using interactions with migrant status
- Explore an idea from compensating differentials theory and add individual characteristics which were observed in a sending region
- Study purchasing power of income accounting for price levels rather than real income

- Address the sample selection problem

Bibliography

Andrienko Y. and S. Guriev "Determinants of Interregional Labor Mobility in Russia." Economics of Transition 12(1), 2004: 1-27.

Arellano M. and Bond S. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equation," Review of Economic Studies 58, 1991: 277-297.

Arellano M. and Bower O. "Another look at the instrumental variables estimation of error-component models," Journal of Econometrics 68, 1995: 29-51.

Blundell R. and Bond S. "Initial conditions and moment restrictions in dynamic panel data models," Journal of Econometrics 87, 1998: 115-143.

Barrett, A. and O'Connell, P.H. "Is There a Wage Premium for Retuning Irish Migrants?" Institute for the Study of Labor (IZA), Discussion Paper No. 135, March 2000.

Bartel A. "The Migration Decision: What Role Does Job Mobility Play?" The American Economic Review, 69(5), 1979: 775-786.

Baumann, R. W. and Reagan, P. B. "Employment and Earnings of Economically Disadvantaged Whites: Understanding Why Appalachians Are Particularly Disadvantaged," (Mimeo, Ohio State University), 2002.

Becker G. "Investment in human capital: a theoretical analysis," Journal of Political Economy, Supplement 70, 1962: 9-49.

Belkar R. "Wealth Distributions of Migrant and Australian-born Households," University of New South Wales Working Paper, 2005.

Bőheim R. and M. Taylor "From the Dark End of the Street to the Bright Side of the Road? The Wage Returns to Migration in Britain," Labour Economics 14, 2007: 99-117.

Blanchard O. J. and L. F. Katz "Regional Evolutions," Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution 23 (1), 1992: 1-76.

Boyle, P.J. "Public Housing as a Barrier to Long-Distance Migration," International Journal of Population Geography 1, 1995: 147–64.

Boyle, P. J., Cooke, T., Halfacre, K. H. and Smith, D. "A cross-national comparison of the impact of family migration on women's employment status," Demography 38(2), 2001: 201–213.

Chen M. K. and Chevalier J. A. “The taste for leisure, career choice, and the returns to education,” Economics Letters 99 (2), 2008: 353-356.

Clark W. and Huang Y. “Linking Migration and Mobility: Individual and Contextual Effect in Housing Markets in the UK,” Regional Studies 38(6): 617-628.

Courchene, T. J. (1970). "Interprovincial migration and economic adjustment," *Canadian Journal of Economics* 3(4), 1970, 550-576.

Détang-Dessendre C., Drapier C., Jayet H. “The Impact of Migration on Wages: Empirical Evidence from French Youth,” Journal of Regional Science 44(4), 2004: 661-691. Détang-Dessendre, C. and Mohlo, I. “Migration and Changing Employment Status: A Hazard Function Analysis,” Journal of Regional Science 39(1), 1999: 103-123.

Fang H. “Disentangling the College Wage Premium: Estimating a Model with Endogenous Education Choices,” International Economic Review 47(4), 2006:1151-1185.

Greenwood M.J. “Research on Internal Migration in the United States: A Survey,” Journal of Economic Literature, 13, 1975: 397-433.

Greenwood M.J. “Internal Migration in Developed Countries” in Rozenzweig, M. R., Stark, O. (eds.), Handbook of Population and Family Economics, Elsevier Science, 1997.

Hayakawa K. **“**Small sample bias properties of the system GMM estimator in dynamic panel data models”, Economics Letters 95(1), 2007: 32-38.

Ham J., Reagan P., Li X. “Propensity Score Matching, a Distance-Based Measure of Migration, and the Wage Growth of Young Men,” IEPR Working Paper No. 05.13, 2005.

Headey B., Warren D., Harding G. “Families, Incomes and Jobs: A Statistical Report of the HILDA Survey,” The University of Melbourne, 2006.

Headey B. and Warren D. “Families, Incomes and Jobs, Volume 3: A Statistical Report on Waves 1 to 5 of the HILDA Survey,” The University of Melbourne, 2008.

Jasso G., Massey D.S., Rosenzweig M.R., Smith J.P. “The New Immigrant Survey Pilot (NIS-P): Overview and New Findings about U.S. Legal Immigrants at Admission”, Demography 37 (1), 2000: 127-138.

Kyriazidou, E. “Estimation of a Panel Sample Selection Model,” Econometrica 65, 1997: 1335-1364.

Kyriazidou, E. “Estimation of Dynamic Panel Data Sample Selection Models,” The Review of Economic Studies 68(3), 2001: 543-572.

Krieg R. “Occupational Change, Employer Change, Internal Migration and Earnings,” Regional Science and Urban Economics 27: 1-15.

Leigh A. and C. Ryan “Estimating Returns to Education: Three Natural Experiment Techniques Compared,” Discussion Paper No. 493, Centre for Economic Policy Research, the Australian National University, August 2005.

Lucas R.E.B. “Internal Migration in Developing Countries” in Rozenzweig, M. R., Stark, O. (eds.), Handbook of Population and Family Economics, Elsevier Science, 1997.

Miller P.W., Mulvey C. and Martin N. “What Do Twins Studies Reveal About the Economic Returns to Education? A Comparison of Australian and US Findings”, American Economic Review, 85(3): 586-599.

Magrini M.-B. “L’arbitrage coûts/benefices de la mobilité spatiale des jeunes actifs,” Groupement de Recherches Economiques et Sociales, Cahier N° 2006 – 26.

Mincer J. “Family Migration Decision,” Journal of Political Economy, 86(5), 1978: 749-773.

Mulvey C., Miller P., Martin N. “The Return to Schooling: Estimates from a Sample of Young Australian Twins,” Labour Economics 13(5), 2006: 571-587.

Rodgers, J. R. “Hourly Wages of Full-Time and Part-Time Employees in. Australia,” Australian Journal of Labour Economics 7, 2004: 231-254.

Roodman D. “How to Do xtabond2: An Introduction to “Difference” and “System” GMM in Stata”, Working Paper No. 103, 2006.

Schwartz A. “Interpreting the Effect of Distance on Migration,” Journal of Political Economy 81 (5), 1973: 1153-1169.

Sjaastad L.A. “The costs and returns of human capital,” Journal of Political Economy, Supplement 70, 1962: 80-89.

Trostel, P., Walker, I., and Woolley, P. “Estimates of the economic return to schooling for 28 countries,” Labour Economics 9, 2002: 1–16.

Watson N. and Wooden M. “Assessing the Quality of the HILDA Survey Wave 2 Data” HILDA Project Technical Paper Series No. 5/04, July 2004.

Windmeijer, F. “A finite sample correction for the variance of linear efficient two-step GMM estimators,” Journal of Econometrics 126, 2005: 25–51.

Wooldridge J. M. Econometric analysis of cross section and panel data. MIT Press, 2002. 752 pp.

Yankow J. "The Wage Dynamics of Internal Migration," Eastern Economic Journal 25, 1999: 265-278.

Yankow J. "Migration, Job Change, and Wage Growth: a New Perspective on the Pecuniary Return to Geographic Mobility," Journal of Regional Science 43(3), 2003: 483-516.

Annex 1

Table A. Wage equation, system GMM results

	Weekly wage, log (W(t))					Hourly wage, log (W(t))			
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)
Wage, lag (W(t-1))	0.154	0.15	0.13	0.147	0.156	0.168	0.154	0.155	0.148
	[0.022]***	[0.021]***	[0.020]***	[0.020]***	[0.020]***	[0.026]***	[0.024]***	[0.024]***	[0.023]***
Distance above 30km, log (Dist)	0.009	-0.001	0.063			0.015	0.0002	0.059	
	[0.003]***	[0.003]	[0.031]**			[0.004]***	[0.003]	[0.019]***	
W(t-1)*Dist			-0.01					-0.02	
			[0.005]**					[0.006]***	
Dist*Job				0.008	0.172				0.007
				[0.003]***	[0.029]***				[0.003]**
Dist*Family				-0.008	-0.008				-0.01
				[0.004]**	[0.004]**				[0.004]**
Dist*Other				-0.001	-0.002				0.0001
				[0.005]	[0.004]				[0.005]
W(t-1)*Dist*Job					-0.025				
					[0.004]***				
Hours of work, log	0.674	0.661	0.634	0.662	0.669				
	[0.033]***	[0.031]***	[0.030]***	[0.030]***	[0.029]***				
Education	0.029	0.03	0.031	0.03	0.029	0.025	0.028	0.028	0.028
	[0.003]***	[0.002]***	[0.002]***	[0.002]***	[0.002]***	[0.003]***	[0.002]***	[0.002]***	[0.002]***
Tenure	0.005	0.005	0.006	0.005	0.005	0.004	0.005	0.005	0.005
	[0.001]***	[0.001]***	[0.001]***	[0.001]***	[0.001]***	[0.001]***	[0.001]***	[0.001]***	[0.001]***
Tenure^2/100	-0.006	-0.008	-0.008	-0.008	-0.007	-0.003	-0.007	-0.008	-0.007
	[0.003]*	[0.003]**	[0.003]***	[0.003]**	[0.003]**	[0.003]	[0.003]**	[0.003]**	[0.003]**
Age	0.024	0.024	0.025	0.024	0.024	0.016	0.02	0.021	0.021
	[0.003]***	[0.003]***	[0.003]***	[0.003]***	[0.003]***	[0.003]***	[0.003]***	[0.003]***	[0.003]***
Age^2/100	-0.027	-0.026	-0.027	-0.026	-0.025	-0.017	-0.021	-0.021	-0.021
	[0.004]***	[0.004]***	[0.004]***	[0.004]***	[0.004]***	[0.004]***	[0.004]***	[0.004]***	[0.004]***
Sex, 0/1 f/m	0.132	0.125	0.134	0.122	0.118	0.077	0.08	0.079	0.078
	[0.017]***	[0.016]***	[0.015]***	[0.016]***	[0.015]***	[0.013]***	[0.010]***	[0.010]***	[0.009]***
Marital status, 0/1	0.048	0.05	0.053	0.052	0.05	0.058	0.052	0.053	0.053
	[0.009]***	[0.009]***	[0.009]***	[0.009]***	[0.009]***	[0.010]***	[0.009]***	[0.009]***	[0.009]***
No. of children 0-14 years	0.012	0.009	0.008	0.01	0.01	0.035	0.022	0.022	0.022
	[0.006]**	[0.006]	[0.006]	[0.006]*	[0.006]*	[0.007]***	[0.006]***	[0.006]***	[0.005]***
Household size	-0.014	-0.01	-0.011	-0.01	-0.009	-0.016	-0.013	-0.012	-0.012
	[0.005]***	[0.004]**	[0.004]**	[0.004]**	[0.004]**	[0.005]***	[0.004]***	[0.004]***	[0.004]***
Constant	0.84	1.778	1.132	1.727	1.499	0.105	3.953	3.148	3.16
	[3.609]	[3.557]	[3.810]	[3.515]	[3.496]	[4.159]	[3.995]	[3.972]	[3.937]

Standard errors in brackets
* significant at 10%; ** significant at 5%; *** significant at 1%
The number of observations is 19116 and the number of individuals is 5236 in each regression[14]

[14] Coefficients received for dummies for firm size, region, occupation and industry are not reported in Table but available from the author on request.

Annex 2

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 $W(\cdot)$ and costs $C(x, y)$, functions, both are defined on an everywhere dense set X. The Karush–Kuhn–Tucker conditions for the optimal solution of the maximization problem (3) are

$\frac{\partial \pi}{\partial y} = \mu \cdot \frac{\partial C(x, y)}{\partial y}$	stationarity condition
$C(x, y) \le B$	primal feasibility
$\mu \ge 0$	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

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

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 sets of cities $Y^*(x, B_1)$ and $Y^*(x, B_2)$. These two sets do not have common elements, $Y^*(x, B_1) \cap Y^*(x, B_2) = \{\emptyset\}$ otherwise the profits are be equal. Moreover, inequality $\pi^*(x, B_1) < \pi^*(x, B_2)$ means the set $Y^*(x, B_2)$, for each point of which $C(x, y) \le B_2$, is not feasible for the first agent, that is, $C(x, y) > B_1$ for any point $y \in Y^*(x, B_2)$. These two inequalities contradict the assumption he 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, I show this property for differentiable wage and costs functions. The envelope theorem together with the dual feasibility and property of costs function (4) lead to

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

Also, we can provide here a proof by contradiction in the general case. Assume $\theta_1 < \theta_2$ and the contrary to what we are eager to show

$$\pi^*(x,B,\theta_1) < \pi^*(x,B,\theta_2) \qquad (5)$$

Then the two optimal sets of cities are $Y^*(x,B,\theta_1)$ and $Y^*(x,B,\theta_2)$. It follows from inequality (6) that any point in the set $Y^*(x,B,\theta_2)$ is not feasible for the first agent, $C(x,y,\theta_1) > B$ for any point $y \in Y^*(x,B_2)$, what contradicts the assumptions $C(x,y,\theta_2) \le B$ and $C(x,y,\theta_1) < C(x,y,\theta_2)$. Q.E.D.