ESSAYS ON LABOR MARKET MATCHING, LABOR MOBILITY AND EDUCATIONAL MISMATCH

by

YAN JIANG

A dissertation submitted to the Graduate Faculty in Economics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2012

This manuscript has been read and accepted for the Graduate Faculty in Economics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Wim P.M. Vijverberg

Date

Chair of Examining Committee

Merih Uctum

Date

Executive Officer

Wim P.M. Vijverberg

Michael Grossman

David A. Jaeger

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

ESSAYS ON LABOR MARKET MATCHING, LABOR MOBILITY AND EDUCATIONAL MISMATCH

by

Yan Jiang

Advisor: Professor Wim P.M. Vijverberg

This dissertation includes three essays on labor market matching, labor mobility and educational mismatch.

Essay one is a literature survey on labor market matching. It is the first attempt in the literature to link various aspects of labor market matching. After an overview of the structural equilibrium search models, I elaborate on the micro-foundations of the matching function, a major modeling tool to capture the influence of frictions on equilibrium outcomes. Issues such as turnover theory, mismatch and dynamic income processes are also examined in this survey.

Essay two considers the effect of voluntary job mobility on worker well-being. Using data from National Longitudinal Survey of Youth 79 (NLSY79), I construct measures of worker well-being that take into account various ingredients likely to factor into a worker's utility at workplace. I adopt a difference-in-differences matching strategy to uncover the otherwise unobservable potential outcomes of not changing jobs and identify the effect of voluntary labor mobility on worker well-being. The result shows that voluntary turnover increases the well-being at workplace for movers who are in the early stage of their career and conduct complex job changes involving different types of job. However, the positive effect of job

mobility is insignificant and much smaller for movers taking simple job changes. This is in contrast with the fact that complex job movers actually experienced insignificant wage gains from the mobility. This result highlights the role of non-pecuniary job rewards in triggering voluntary turnover.

Essay three considers the wage effects of educational mismatch using data from the 2003 wave of National Survey of College Graduates (NSCG 2003). I find that the average wage loss associated with educational mismatch is significant and persistent. Graduates who are mismatched for involuntary reasons incur greater wage penalty compared to those for voluntary reasons. In addition, graduates with advanced degree suffer more from mismatch relative to those with only bachelor's degree. Lastly, there are considerable amounts of variations in the distributional impacts. The wage penalty is quite large at lower quantiles and decrease sharply towards higher quantiles.

ACKNOWLEDGMENTS

First of all, I should thank my advisor Wim Vijverberg for his warm encouragements, insightful criticism, time and attention when I was working on this thesis.

I am also grateful to my committee members Michael Grossman and David Jaeger for their generous help and thoughtful comments throughout the years.

Furthermore, I am indebted to Professor Richard Freeman for his inspiring suggestions about the data.

Last but not least, I would like to thank my loving family members: my father Guozhen Jiang, my mother Zhuoji Hu, my husband Wei Zhang and my daughter Claire Zhang. Without their support, I would not have weathered this long journey to fulfill my dream.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

LABOR MARKET MATCHING: MODELS AND IMPLICATIONS

1.1 Introduction

One aspect of labor economics concerns the matching of workers to jobs so that mutually beneficial relationships develop over time. Modern matching theory, evolved from an earlier framework of "search theory", offers a way of modeling markets in which frictions prevent instantaneous adjustment of the level of economic activity. Section 2 thus begins with a survey of structural equilibrium search models. The major modeling tool to capture the influence of frictions on equilibrium outcomes is the matching function. Section 3 then elaborates on the microfoundations underlying the function and discusses possible variables that are likely to be influential in empirical work. One important implication of the matching theory is that labor turnover is a dynamic process in which job-worker matches are improved over time. Hence Section 4 discusses different theories of labor turnover. Section 5 deals with issues of mismatch with a focus on the skill dimension. Section 6 sketches different specifications of dynamic income processes and the associated assumptions. Finally, a brief summary concludes the paper.

1.2 Modeling Matching Process

A worker-job match forms when a qualified worker and a sufficiently attractive vacancy meet and create a cooperating coalition. One way to characterize the matching process is the structural equilibrium search model initiated by Eckstein and Wolpin (1990) and featuring search and recruiting friction and the reallocation of workers from time to time across alternative productive activities. In a typical structural equilibrium search model, job seekers conduct sequential job search while the employers set optimal wages to take account of the responses by job seekers and other firms. Solutions are dynamic stochastic equilibria in the sense that time and uncertainty are explicitly modeled, expectations are rational, private gains from trade are exploited and the actions taken by all agents are mutually consistent (Mortensen and Pissarides, 1998).

Following is a theoretical framework of equilibrium search model developed by van den Berg and Ridder (1998). There are continua of workers and firms with measures m and 1, respectively. Workers receive job offers at rate λ_0 if unemployed and λ_1 if employed. A job offer is an *i.i.d.* drawing from a wage offer distribution with *c.d.f.* $F(w)$, while the *c.d.f* of the distribution of wages received by workers who are currently employed is denoted as $G(w)$. An offer has to be accepted or rejected upon arrival. During tenure of a job, the wage is constant. The utility flow of being employed at a wage w equals w. Job-worker matches break up at rate δ. If this happens, the worker becomes unemployed. The utility flow of being unemployed is b. Firms have a linear production function and the marginal revenue product is p. A firm pays all its workers the same wage w. Workers maximize their expected wealth and firms maximize their expected steady-state profit flow. The firms cannot set their wage below the mandatory minimum wage $\underline{w}_L$.

Under those assumptions, the optimal strategy of an unemployed individual has the reservation wage property and the reservation can be shown to be:

$$r = b + (\lambda_0 - \lambda_1) \int_r^{\infty} \frac{\overline{F}(w)}{\delta + \lambda_1 \overline{F}(w)} dw \tag{1}$$

where $\overline{F} = 1 - F$. Further, an employed individual accepts a wage offer if and only if it exceeds his or her current wage. In a steady state, the outflow of employed workers to jobs with a wage that exceeds w must be balanced by an inflow from unemployment, so the following relation between the earnings distribution and the wage offer distribution holds:

$$G(w) = \frac{\delta F(w)}{\delta + \lambda_1 \overline{F}(w)} \tag{2}$$

On the firm side, the steady-state workforce l that is available to the firm depends on w, r and the distribution F of wages set by the other firms competing for the same workers. Each firm sets a wage w to maximize its steady-state profit flow π, which equals $(p-w)l(w; r, F)$.

A noncooperative steady-state equilibrium solution consists of reservation wage r and a wage offer distribution F such that r satisfies (1) given F, and every w in the support of F maximizes the steady-state profit flow $(p-w)l(w; r, F)$. It can be shown that the equilibrium F is absolutely continuous. Further, F and G have probability density functions f and g with support $[\underline{w}, \overline{w}]$ satisfying $\underline{w} = max(\underline{w}_L, r)$ and $\overline{w} < p$.

The equilibrium wage offer distribution can be shown as:

$$F(w) = \frac{\delta + \lambda_1}{\lambda_1} (1 - \sqrt{\frac{p - w}{p - \underline{w}}}) \tag{3}$$

Notice that match rents are divided between firm and worker by the wage rule. In this model, firms set wages and there is no bargaining over the wage. A match between a worker and a firm has a net revenue flow of $p-b$. At the prevailing wage w, the firm receives the part $p-w$ of this flow and the worker receives $w-b$. The way in which the value of the match is split between the worker and the firm varies randomly across firms, but the corresponding wage offer distribution shifts upward if the bargaining power of workers is increased by the way of a decrease in search frictions for the employed. Alternative assumptions about wage determination will allow for more complicated strategic behavior on both sides of the market.

Another innovation to the equilibrium search model is to bring in heterogeneity on both sides. Postel-Vinay and Robin (2002) construct and estimate an equilibrium search model with double productive heterogeneity and on-the-job search. In their model, unobserved heterogeneity among workers is introduced in the form of cross-worker differences in a "competence" parameter, while firms are also heterogeneous with respect to their marginal productivity of labor. Unemployed workers search for a job and employees search for a better job. Both workers and firms are imperfectly informed about the location of the other party's types until they meet. Employers have all the bargaining power. They make take-it-or-leave-it wage offers conditional on workers' characteristics and can respond to the outside job offers received by their employees. On-the-job search allows employees to locate alternative employers and brings in Bertrand price competition with the employee's current employer. This competition either results in a wage rise or in job mobility. It is noteworthy that the poaching employer might pay the worker a wage that is less than his/her current wage if the option value of turning down the best offer that the current employer can make in exchange of a greater potential best offer is large enough. As a result, their model not only generates tenure effects but also job-to-job mobility with wage cuts.

1.3 Microfoundations of the Matching Technology

One key element of the equilibrium search model is the matching technology, defined as the relation between inputs, search and recruiting activity, and the output of the matching process, the flow rate at which unemployed worker and vacant jobs form new job-worker matches (Mortensen and Pissarides, 1998). For example, some empirical equilibrium search models make the convenient consumption of random matching by which all firms have the same probability of being contacted by searching workers regardless of their recruiting efforts. In this section, I focus on the anatomy of the complex process through which workers and firms are matched together. The general way to summarize matching technology in the literature is to invoke a well-behaved matching function of the following form:

$$M = m(U, V) \tag{4}$$

where M is the number of jobs formed during a given time interval, U is the number of unemployed workers looking for work and V the number of vacant jobs. This function is increasing in U and V and homogeneous of degree one. In particular, this section looks at the theoretical foundations of the matching function at the micro level and discusses some of the variables that are likely to be influential in empirical work.

In the simplest case of matching process summarized by Petrongolo and Pissarides (2001), U workers know exactly the location of V job vacancies and send one application to each. If a vacancy receives more than one applications, it selects one application at random and forms a match. The other applicants are returned to the pool of unemployed workers to apply again. In this case, the probability that a given vacancy will not receive any applications is

$(1-1/V)^U$ and the matching function has the following form for large U and V:

$$M = V[1 - e^{-U/V}] \tag{5}$$

This matching function satisfies constant returns to scale. However, it is far away from a good approximation to matching in real labor markets. For example, if the level of unemployment and vacancies is the same, the mean duration of unemployment is 1.58 periods. But if the level of unemployment is three times as high as that of vacancies, mean duration is 3.16. In actual labor markets, duration will rise by more than this prediction if the level of unemployment is higher (Petrongolo and Pissarides, 2001).

To bring the model closer to the data, the assumptions underlying (5) are therefore relaxed. In the first case, workers do not know the firms with the vacancies and choose at random one firm to apply. Denote the labor force size as L, then the matching function becomes:

$$M = V[1 - e^{-U/(L-U+V)}] \tag{6}$$

However, although it satisfies constant returns to L, U and V, the matching function (6) exhibits increasing returns to scale in U and V and even fails the assumption of diminishing returns to unemployment.

In the second extension, not all workers are suitable for the vacancies available but the worker does not know which vacancies are suitable. Let K be the fraction of workers who are suitable employees for a randomly selected vacancy, the matching function generalizes to

$$M = V[1 - e^{-KU/V}] \tag{7}$$

Worker heterogeneity can also be introduced into matching function (5) by making the assumption that the intensity of search is a choice variable. Here the intensity of search is defined as the number of “units” of search supplied by a given individual. Let s be the average number of search units supplied by an unemployed person. Then the total number of search units supplied is sU and the matching function is:

$$M = V(1 - e^{-sU/V}) \tag{8}$$

Both (7) and (8) have constant returns to U and V for given K and s. Further, search units are supplied at a cost, which is normally increasing, and they are chosen optimally to maximize the net returns from search. (Pissarides 2000, ch. 5) Therefore, different individuals will choose different number of search units, depending on their search costs, the cost of unemployment, and the expected returns from employment. Under model (8), the hazard rate for an individual who supplies s_i units of search is $s_i m(sU, V)/sU$.[1] Econometric estimation of the hazard function thus depends on individual characteristics. For example, they can be demographic variables that influence the intensity of search or other variables that influence the cost of search such as unemployment insurance.

In what follows, I elaborate on two more alternative assumptions of matching and their implications on the matching function. The matching function used by Blanchard and Diamond (1994) imposes the assumption that firms receive many applications at a time and have preferences over job applicants. Specifically, there are two types of unemployed workers, the short-term unemployed and the long-term unemployed. Let the number of short-term unemployed be U^S and the number of long-term unemployed be U^L. The firms prefer short-term unemployed to long-term unemployed. Therefore, the long-term unemployed only get jobs for which there are no short-term unemployed applicants. In this case, the aggregate

[1] s denotes the average number of search units at the aggregate level, while s_i denotes the number of search units for individual i.

matching function is

$$M = m(U^S + U^L, V) \tag{9}$$

However, the hazard rate for the short-term unemployed is $m^S(U^S, V)/U^S$ and for the long-term unemployed $m(U^S + U^L, V)/U^L - m^S(U^S, V)/U^L$. It can be shown that if the matching functions are identical, the hazard rate of the short-term unemployed is always higher than the hazard rate of the long-term unemployed.

In the job search literature, it is usually assumed that job seekers have no information on the price and location of potential matches. They take a vacant job at random and apply for it. On the other extreme, Coles and Smith (1998) consider an assumption that job seekers have complete information about the available job vacancies through marketplaces such as job centers and apply simultaneously to all the vacancies they think are likely to be acceptable. The essential idea is that job seekers and employers can contact each other costlessly through marketplaces. Job seekers have heterogeneous preferences. For simplicity, assume that with probability λ the job seeker likes the job and with probability $1 - \lambda$ the job seeker does not like the job. At each matching round, job-worker pairs make contact and a market equilibrium is characterized by immediate trade. Given the assumption of full information and positive time discounting, it is not surprising because deferring trade makes both agents strictly worse off. Therefore, those workers who remain unmatched do so because there are no vacancies suitable for them among the existing pool. Also, no job vacancy or unemployed worker who has been through one round of matching will attempt to match again with a pre-existing job seekers or vacancy in the subsequent matching rounds. Hence, the stock of unmatched workers must be waiting to match with the new flow of vacant jobs. Similarly, the stock of unmatched vacancies must be waiting to match with the new flow of workers. This is referred to as "stock-flow" matching.

At each period, there is a stock of "old" unemployed workers and "old" vacant jobs, denoted as U and V respectively. New unemployed workers and vacancies enter this market according to independent poisson processes, where u and v are arrival rates of the unemployed workers and vacancies, respectively. Therefore, the probability that a new vacancy is matched upon entry is $1-(1-\lambda)^U$, so the matches due to new vacancy creation are $v\left[1-(1-\lambda)^U\right]$. The probability that a new worker is matched upon entry is $1-(1-\lambda)^V$, so the matches due to new entry of workers are $u\left[1-(1-\lambda)^V\right]$. Because flows match with stocks and there is no interaction between stocks, the matching function is the sum of the two matches:

$$M = v\left[1-(1-\lambda)^U\right] + u\left[1-(1-\lambda)^V\right] \tag{10}$$

The matching function in (10) exhibits increasing returns to scale in the stocks and the flows. The reason is that job seekers apply to all the available job vacancies simultaneously.[2] If the number of job vacancies and unemployed workers doubles, the applications of each and every job seeker double.

1.4 Job Matching and Theories of Turnover

One important implication of the matching process discussed earlier is that the existence of better alternatives for both employers and workers involved in a job-worker match motivates search on both sides. Permanent separation at some future date occurs when either side

[2] Applying to more than one vacancy at a time is a realistic feature of the application process but it depends on a constant $1-\lambda$, the probability that any given unemployed worker does not apply to a given vacancy. However, this is a fairly strong assumption because if U is large and λ is not tiny, then $(1-\lambda)^U$, the probability that nobody in the entire pool of unemployed applies to any given vacancy, becomes a really small number. Yet, some employers can't find workers immediately; they have vacancies posted for a while.

finds a better match. Labor turnover thus characterizes a dynamic process in which job-worker matches are improved. Imperfect information underlies most theoretical models of turnover. As summarized in Jovanovic (1979), there are two categories of models for job switching. The first category includes models treating turnover as a result of the arrival of information about the current job match. In those models, the only way to determine the quality of a particular match is to form the match and "experience" it. In the second category are "pure-search-good" types of job change. In those models, jobs are pure search goods, and matches dissolve because of the arrival of new information about an alternative prospective match. In this section, I take up the second type of model first because it is a natural extension of the standard search model. Then I review the model treating jobs as "experience goods", followed by a brief discussion of the inclusion of non-pecuniary rewards into models of labor turnover.

Jovanovic (1979) elaborates on the standard search model and uses the job-matching approach to give job turnover an equilibrium interpretation. In the model, the quality of match differs across prospective matches. There are no "good" workers and "good" employers, but only good matches. Employer can contract with workers individually so that the wage paid reflects the quality of the match. Each period the worker contacts a new employer. Upon contact, the quality of match is immediately ascertained and the employer makes a wage offer to the worker based on the quality of the match. All search in the model is assumed to be done by the worker. The worker's intensity of search determines the rate of arrival of new wage rates.

Let $\lambda(t)\Delta t$ be the probability that an offer will arrive during the time interval $(t, t + \Delta t)$. Wage offers are drawn independently from the wage-offer distribution characterized by the *c.d.f.* $F(w)$. Since $F(w)$ is known, the worker's optimal policy is characterized by a reservation wage $\theta(t)$. Then the probability that an acceptable offer arrives on $(t, t+\Delta t)$ is $h(t)\Delta t$ and $h(t)$ is given by:

$$h(t) = \lambda(t)(1 - F(\theta(t))) \tag{11}$$

Assume that a fraction, $s(t)$, of the worker's time is devoted to on-the-job search and that another fraction, $\phi(t)$, is devoted to on-the-job training. Then $0 \leq s(t) + \phi(t) \leq 1$. Let $x(t)$ be the worker's productivity on a particular job and it can be decomposed as:

$$x(t) = \mu + k(t) \tag{12}$$

where μ is quality of the employer-worker match, $k(t)$ is the human-capital stock accumulated through training on the job and t is interpreted as job tenure. Assume the law of evolution for $x(t)$ on the current job is:

$$\frac{dx(t)}{dt} = g[\phi(t)x(t) - \delta x(t)] \tag{13}$$

where $x(0) = \mu$, which means at the time $(t = 0)$ when the worker begins working on a particular job, his or her productivity is equal to μ. Growth in productivity can be achieved by devoting time to on-the-job training. If no time is devoted to investment, productivity depreciates at the rate δ. The amount actually produced by the worker is:

$$[1 - \phi(t) - s(t)]x(t) = w(t) \tag{14}$$

Form (12) and (13), it can be inferred that a job change occurs only if $\mu' > \mu + k(t) = x(t)$, where μ' is the quality of the match with a prospective new employer.

Three results can be derived from the equilibrium of the above model. First, $\frac{ds(t)}{dx(t)}|t < 0$, which means that the amount of time devoted to searching for alternative employment decreases with $x(t)$ holding t fixed. Hence those who are better matched and those who have more specific human capital spend less time searching. Second, let $x(t, \mu)$ and $s(t, \mu)$ denote the optimal trajectories of human capital and search activity on a job characterized by quality of match μ. Given that tenure t has been attained, the density of job separation over the tenure interval $(t, t + \Delta t)$ is

$$h(t, \mu)) = \lambda(t)(1 - F(x(t, \mu)) \tag{15}$$

It can be shown that $h_\mu(t, \mu) < 0$, which shows that separation probabilities regarded as a function of job tenure are uniformly lower for those who are well matched. Third, $h_t(t, \mu) < 0$ as long as $\dot{x} > 0$, implying that the separation probability will decline with tenure on a given job as long as the stock of specific human capital is increasing.

As an alternative to this "pure-search-good" model of job change, Jovanovic (1979) builds a model of job turnover when a job match is treated as a pure experience good. In this model, for each worker a non-degenerate distribution of productivities exists across different jobs. The problem is to optimally assign workers to jobs. Employers can contract with workers on an individual basis. Individual contracting creates a structure of rewards that provides proper signals for the attainment of optimal matches. Further, imperfect information exists on both sides of the market about the exact location of one's optimal assignment. The job-matching thus generates turnover as the phenomenon of optimal reassignment caused by the accumulation of better information with the passage of time.

Let $X(t)$ be the contribution by a worker to the total output of the firm over a period of length t and

$$X(t) = \mu t + \sigma z(t) \tag{16}$$

where $z(t)$ is a standard wiener process. σ is the same for each firm-worker match while μ differs across matches. μ can be interpreted as a measure of the quality of the match. A good match is one possessing a large μ. When the match is formed, μ is unknown. As the match continues, further information in the form of output as given by (16) is generated and this information is completely match specific. Therefore, the only way to learn about μ is to observe the worker on the job for a period of time.

Firms are assumed to be risk neutral. They compete for workers by offering wage contracts characterized by a wage function $w[X(t), t]$. This is the wage paid to the worker with tenure t if his cumulative output contribution is equal to $X(t)$. If the firm wants to fire a certain worker, it will lower the wage by an amount sufficient to induce the worker to quit. Workers are assumed to live forever. As long as the worker remains with the firm, he or she receives payment according to the wage function $w(\cdot)$. The worker has the option of quitting at any time and the present value of quitting a job and then pursuing the best alternative is denoted as Q. Let $\pi(Q, [w])$ be the discounted expected net revenue from the employment of a given worker who is offered the contract $[w]$ and who has a present value of quitting equal to Q. Therefore, the firm's problem is to maximize $\pi(Q, [w])$ over functions $[w]$, treating Q as given because Q is determined by the wage policies of other firms.

The equilibrium results given by the model predict that workers remain on jobs in which their productivity is revealed to be relatively high and that they select themselves out of jobs in which their productivity is revealed to be low. Since wages always equal expected marginal products for all workers, the model generates wage growth as tenure increases. Also, the model shows that each worker's turnover probability is a decreasing function of his tenure because mismatch between a worker and his employer is likely to be detected early on rather than later.

While most of the studies on labor turnover focus on the pecuniary returns of a job, it is noteworthy that workers are willing to switch jobs for non-pecuniary rewards as well. Akerlof et al. (1988) construct a Keynesian model of labor turnover in a non-clearing market with job rationing. A characteristic of the equilibrium of this model is that some individuals covet jobs held by others who are no better qualified. When wages are sticky, people cannot obtain jobs they desire by offering to work for lower pay. As a result, the autonomous departure of an individual from a job creates a sequence of opportunities that is called a vacancy chain.

The idea of a vacancy chain can be illustrated in the following way. Suppose that an employee (A) autonomously withdraws from the labor force, creating a job vacancy. This vacancy provides an opportunity for workers who covet A's job. Suppose that B takes A's old position. If B is employed, B's quit creates a further vacancy and the vacancy chain continues. But if B is unemployed, the chain of vacancies, which began with A's departure from his or her job, ends. Therefore, vacancy chains are triggered by the creation of autonomous vacancies through new job creation, withdrawal of workers from the labor force and voluntary quits into unemployment. Vacancy chains end only when a vacancy is filled by an individual who is unemployed.

The logic of the vacancy chain relies on the idea that some employed workers are ready to switch jobs for either higher wages or higher non-pecuniary rewards. Akerlof et al. (1988) then continue to empirically investigate the importance of non-pecuniary reward in quit decisions. Econometric tests confirm the hypothesis that proportionate changes in pecuniary and non-pecuniary rewards are of equal importance in affecting workers' propensities to quit. However, because non-pecuniary rewards are so variable, most job-related quits are motivated by non-pecuniary rewards.

1.5 Mismatch

Mismatch is an empirical concept that measures the degree of heterogeneity in the labor market across a number of dimensions, usually restricted to skills, industrial sector and location (Petrongolo and Pissarides, 2001). Skill mismatch refers to various types of gaps or imbalances in skills, knowledge or abilities that may be qualitative or quantitative in nature. Skill match may lengthen the time that it takes to match a given group of workers to a given group of firms. Industrial sector influences matching because some industry-specific skills may not be picked up by generally available measures of skills. Location matters in matching because of imperfect labor mobility. This section focuses on the skill dimension of mismatch and reviews the literature on two aspects of skill mismatch: its measurement and its determinants.

Measuring skill mismatch requires a good understanding of what skills are and what skill mismatch constitutes. According to Allen and van der Velden (2005), in principle there are five ways to assess skill levels. Assessment and testing are two objective measures, while supervisor rating, individual self-assessment of skills level and a job's skill requirements as a proxy for possessed skills are subjective measures. With skill measures in hand, the measurement of skill mismatch should be easy, but actually it is not straightforward. Mismatch is a complex phenomenon with different dimensions. One important aspect is whether the mismatch refers to vertical or horizontal skill mismatch. Vertical skill mismatch refers to a situation where the level of skills a worker possesses is higher or lower than what is required on the job. Horizontal skill mismatch refers to a situation where workers have the appropriate qualification level but different skills from what is required for the job they occupy.

There is large literature on over- and under-education that focuses on the vertical dimension of educational mismatch. Graduates who possess more schooling than their job requires are overeducated, while those with less schooling than required are undereducated. Notice that although there is a clear relation between educational mismatch and skill mismatch, these two concepts are not interchangeable. The empirical research by Allen and van der Velden (2007) shows that the wage effects of educational mismatch, particular the effects of working below one's level, are much stronger than that of skill mismatch. According to Hartog (2000), required schooling has been measured in the literature in three different ways. The first way is job analysis where job analysts specify the required level of education for the various job titles in an occupational classification. This is a very attractive method of defining job requirement because of its explicit goal of objectivity, clear definitions and detailed measurement instruction. However, the required careful, systematic work may be too expensive to carry out on a large scale. The second way is worker self-assessment that relies on the worker's specification of the education required for the job. This way has the advantage of drawing on all local, up-to-date information, but lack of rigorous instruction may be its major drawback. The third method is realized matches, where required education is derived from what workers in the respondent's job or occupation usually have attained. This measure is different from job analysis and worker-assessment in that it relies on the realized equilibrium determined by the interplay of labor supply and demand while the other two act as demand curve location parameters.

Hartog (2000) interprets educational mismatch as a temporary status in a worker's career development resulting from imperfect information and deliberate searching and matching.

Here the job search is characterized by the process in which a worker with a given level of education seeks to improve the job-worker match by searching for a higher job level. The job level is defined by the required education level of the job. Given the positive effect of required education on wages, this is equivalent to maximizing wages through labor mobility. The incidence of over-education falls with experience because with voluntary job mobility experience will bring about an increase in job level. This happens because individuals switch jobs only if the job level increases, given that their current job level is the highest they were able to obtain. However, the effect of tenure on job level is ambiguous. Assuming random arrival of job level offers, there are two opposing forces. If the present job results from a particularly high offer, the worker will tend to stay. Hence long tenure is characterized by high job levels. On the other hand, workers observed with short tenures are those who have just moved due to a high outside offer taking them away from the previous job. Then low tenure is related to high job levels. Without specific distributional assumptions, it is not clear which factor dominates.[3]

As for skill mismatch in general, there are aggregate forces unrelated to individual search decisions that explain the phenomenon. For example, skill obsolescence has emerged as an important explanation of skill mismatch due to fast technological advances. De Grip and Van Loo (2002) distinguish between technical and economic obsolescence. Technical skill obsolescence affects the stock of human capital a worker possesses. A classical example is the declining physical capacity of a bricklayer whose skills wear out in the course of the career. Economic obsolescence affects the human capital a worker possesses due to external

[3]Self-selection will probably make the job level increasing with tenure, so the chance that a job level is high is greater when tenure is longer than when tenure is shorter.

developments. A typical example is the decline of the value of traditional typewriting skills when word processing was introduced. To empirically examine the effect of skill obsolescence on skill mismatch process and dynamics, it is important to quantify how the half-time of skills varies across occupations. The half-time of a professional is the time after completion of professional training when practicing professionals have become roughly half as competent as they were upon graduation to meet the demands of their profession due to new developments. For example, Lukasiewicz (1971) shows that the half-life of engineering graduates decrease sharply from 12 years for 1940 graduates to five years for graduates at the end of the 1960s.

1.6 Dynamic Income Processes

A logical question following the above discussion of labor market matching is how this dynamic process affects the way wages are expected to evolve over time for individuals. Using the longitudinal employee-employer data (LEED), Topel and Ward (1992) separate wage growth at job changes from that within jobs. Their findings suggest that during the first ten years in the labor market, the typical young worker holds seven full-time jobs and wage gains at job changes average about one third of total wage growth. Therefore, it is essential to take account of job turnover and other uncertainties resulting from dynamic matching when modeling income processes. This section thus discusses different specifications of dynamic income processes and the underlying assumptions.

The commonly adopted specification for log labor income of individual i with t years of labor market experience is of the following sort:

$$y_t^i = f(\boldsymbol{\theta}, \mathbf{X}_t^i) + g(\boldsymbol{\theta}^i, \mathbf{X}_t^i) + u_t^i + e_t^i \tag{17}$$

The first function represents the part of life-cycle variation that is common to all individuals with observable characteristics, so $\boldsymbol{\theta}$ is not individual specific. The second function captures potential individual-specific differences implied by a human capital model with heterogeneity in learning ability. It could be as simple as α^i, which denotes person-specific effect in levels. Alternatively, it could be specified as $\alpha^i + \beta^i t$, denoting heterogeneity in both levels and income growth rates. The last term captures the stochastic component of income with u_t^i denoting permanent effect and e_t^i denoting transitory effect. u_t^i is generally assumed to follow an ARMA(p, q) process such as $u_t^i = \rho u_{t-1}^i + m_t^i$. The transitory error component e_t^i represents a mix between a transitory shock and measurement error, and it can not be foreseen by the individual. Individual uncertainty on the permanent component, however, can be partly foreseen, allowing the individual to have a better prediction than an outside observer.

Debates over income dynamics generally focus on the second and third parts of the above specification. Guvenon and Smith (2010) divide current existing studies in the literature into two groups regarding different beliefs over σ_β^2 (the variance in the individual specific growth rate of income, β_i) and ρ (the rate of persistence in income shocks u_t^i). The first group, referred to as "Restricted Income Profiles", holds the view that individuals are exposed to large and persistent income shocks, while facing similar life-cycle income profiles. In other words, they are imposing the restriction of "$\sigma_\beta^2 \equiv 0$" and "ρ close to 1" on the above specification. The second group, named as "Heterogeneous Income Profiles", insists that individuals are subject to income shocks with modest persistence, while facing individual-specific profiles, so they do not impose any restrictions on σ_β^2 and ρ. Both groups have their supportive evidence; therefore, no consensus has yet been reached in the literature regarding which model matches real data the best.

Sometimes, the main purpose of modeling an income process is to quantify the income risk associated with risky decisions such as human capital investment and choosing a career path. It is the perceived income risk that people act upon, rather than the *ex post* income fluctuation. Therefore, a proper econometric model for this purpose should retrieve the perceived risk that individuals base their decision on. Unobserved heterogeneity and risk could easily confuse an outside observer, even with panel data. There are at least two sources for the confusion. On one hand, to what extent do individuals know their θ^i of the second component of income specification in (17) at different points of their life-cycle? On the other hand, if the permanent component u_t^i is imperfectly predicted by individuals at different stages, then how can the outside observer differentiate between "uncertainty" on the part of individuals and "unknown heterogeneity"?

Guvenen and Smith (2010) aim to answer the first question by taking a stand on what individuals know about their β^i. In their model, individuals enter the labor market with full knowledge of their own α^i and some prior belief about their β^i and then update their beliefs over time in a Bayesian fashion. Uncertainty about β^i at time zero is measured as $\widehat{\sigma}^2_{\beta|0}$ and the learning process is cast as a Kalman filtering problem. Let λ be the fraction of population dispersion in income growth rates that represents uncertainty on the part of individuals at the time they enter the labor market. The value of λ is estimated to be 0.191 using the Panel Study of Income Dynamics (PSID), which reveals only a small amount of prior uncertainty regarding individuals' growth rate. The results of this empirical research show that individuals have much better information about their own income growth rates than what can be predicted by some observable variables available to the econometrician.

As to the second question, Cunha, Heckman and Navarro (2005) implement a method to retrieve the information set of individuals at the stage of their life cycles when they make their college-going decision. Instead of ARMA process, they postulate the u_t^i term as a factor structure: $u_t^i = \mu^i \gamma_t$, where μ^i is a vector of skills (e.g., ability, initial human capital, motivation, and the like) and γ_t is a vector of skill prices. This specification is more directly interpreted as a pricing equation for human capital analysis. This specification also allows them to test which information structure characterizes the data by exploring the covariance between schooling and realized earnings that arise under different agent information structures. The results indicate that for a variety of market environments and assumptions about preferences, over 50% of the *ex post* variance in the returns to schooling are forecastable at the time students make their college choices.

In most studies concerning income dynamics, income risks are identified through properly defined moments of u_t^i and e_t^i in (17) under the assumption of incomplete insurance market. For an outside observer, however, it is hard to distinguish between "unexpected shock" and "anticipated changes". For example, consider a female worker who decide that she will work up to a certain point of her career, and then quit her job to raise a child. The large fall in her income might appear as a permanent shock while in fact this is anticipated beforehand. Yet, for another woman, the pregnancy was a surprise and the fall in income was unexpected. One solution is to utilize the joint data of labor income and consumption. If consumption smoothing holds, anticipated changes in income could be revealed in consumption (and saving) data. This approach has been explored in Hall and Mishkin (1982), Deaton and Paxson (1994), Blundell and Preston (1998), Blundell, Pistaferri and Preston (2008) and Govenen and Smith (2010).

Moreover, people might respond to uninsurable risks by taking actions (such as changes in labor supply or job mobility), so it is necessary to disentangle exogenous shocks from the effects of endogenous choices. Low, Meghir and Pistaferri (2010) build a structural life-cycle model of consumption, labor supply and job mobility with partial outside insurance. Grounding the model of the income process in a utility maximization framework allows individual's precautionary response to shocks to be examined and different sources of shocks to be analyzed separately. Their findings suggest that the estimated variance of the permanent innovation to wage doubles if mobility is ignored. Another approach to incorporate individual's endogenous decisions is to use a simultaneous model of earnings with little attention to an underlying theory of household decisions and constraints. The equations could be viewed as an approximation to the decision rules in a structural model with intertemporal utility maximization. The advantages of this model are less computational constraint and more flexibility to extend the analysis to include other important economic risks such as changes in family structure. Altonji, Smith and Vidangos (2009) use indirect inference to estimate a joint model of earnings, employment, job changes, wage rates, and work hours over a career. They find that shocks associated with job changes and unemployment make a large contribution to the variance of career earnings.

1.7 Summary

This survey is the first attempt in the literature to link various aspects of labor market matching. After an overview of structural equilibrium search model, I elaborate on the micro-foundations of the matching function, a major modeling tool to capture the influence of frictions on equilibrium outcomes. An important implication of the equilibrium model is the role of labor turnover as a way to improve the job-worker matches, so I look at the literature on theories of labor turnover and pay a special attention to the non-pecuniary rewards of job switching. Mismatch, an empirical concept that measures the degree of heterogeneity in the labor market, will hamper or slow down the process of matching. Therefore, I lay out the conceptual framework of mismatch and emphasize skill mismatch as an important dimension of mismatch. Finally, dynamic labor market matching brings about income uncertainties from the perspective of life cycle income maximizing individuals. The last part of the survey thus deals with different specifications of income processes and the underlying assumptions.

CHAPTER II

VOLUNTARY LABOR MOBILITY AND THE GROWTH OF WORKER WELL-BEING

2.1 Introduction

One of the most profound impacts of technological advance and globalization is the increasing levels of turbulence in the labor market. Job creation and job destruction speed up, accelerating job mobility and rendering the notion of lifetime job obsolete. A recent study on job mobility in the European Union reveals that as many as 22.9% of the employed in the UK experienced a change of job during past year in 2005. Therefore, job mobility is a theme that suits current trends.

Within these broader trends, this study focuses on the job mobility of young men. Apart from market-side factors, job mobility is virtually a necessity for those novice career builders. Choosing a career is no doubt a decision under uncertainty. A good match between capacity and job can only be obtained through active searching and trying. This "try and try again" process therefore highlights the role of job mobility as a principal instrument to more productive employment relationships. Just as noted in Topel (1992), the first ten years of career will account for almost 70% of lifetime job changes.

The type of job mobility I target in this paper is voluntary job change. Rather than merely measuring the effect of labor mobility on pecuniary outcomes, my aim is to assess whether voluntary job mobility contributes to improving the well-being of young workers at the workplace, while being mindful of the fact that potential outcomes of not changing jobs can not be observed. For a typical person, work is an important part of life and its meaning is far more than a dollar-and-cents issue. Work is, above all, an activity through which an individual fits into the world, creates new relations, uses his talents, learns and grows develops his identity and a sense of belonging (Morin, 2004). Therefore, what I am interested in here is how much better off job mobility makes workers feel. After all, economics is not about wealth — it's about the pursuit of happiness (Krugman, 1998).

To answer the question of whether and to what extent young men benefit from job mobility, I use data for a cohort of young men who were born between 1957 and 1964 and were working full time without gaps between 1982 and 1988 from National Longitudinal Survey of Youth 79 (NLSY79) to examine the relationship between mobility patterns and labor market outcomes.

As a first step, I construct measures of worker well-being, taking into account various ingredients likely to factor into a worker's utility at workplace. I define worker well-being as the extent of discrepancies between what they desire and what they have actually experienced in all job aspects, reflected in measures of job satisfaction. However, the overall job satisfaction has been criticized for its subjective nature, whereas a series of facet-specific job satisfaction convey more accurate information at the cost of bringing too many variables into analysis. The measure in this paper, however, encapsulates all relevant job satisfaction measures, and presents them in a compact fashion. The resulting variable of interest treats worker well-being as a matter of degree, ranging from 0 (worst) to 1 (best).

Following the terminology of the literature of quasi-experiment, I treat labor mobility as a type of intervention in the early career. In order to evaluate the impact of this treatment on young men, I adopt a difference-in-differences (DID) matching strategy, developed in Heckman, Ichimura and Todd (1997, 1998) and Heckman, Ichimura, Smith and Todd (1998). Conventional DID provides insightful results by laying out the difference between movers and stayers in the before-and-after differences in outcomes. However, failing to satisfy the key identifying assumption of same temporal trends in the absence of treatment for both groups disqualifies DID from carrying a clear interpretation in this context. The solution here is to extend DID by constructing differences conditional on observable characteristics. Throughout the analysis, I differentiate two types of job moves: complex job change and simple job change. Complex job change occurs when workers not only change employers but also change tasks, while simple job change refers to firm-level separations when workers change employers but continue doing the same line of work (Neal, 1999). The DID matching result shows that worker well-being is indeed boosted through complex job mobility in the sense that those movers are significantly happier at their workplace than they otherwise would be if they stayed with their previous employers. However, for movers taking simple job changes the positive effect of job mobility is insignificant and much smaller.

This paper proceeds as follows. Section 2 begins with a discussion of existing related research. Section 3 describes the dataset and presents the pattern of labor mobility exhibited in the sample. In Section 4 I define and construct measures of worker well-being at the workplace. Section 5 then lays out the empirical strategy and Section 6 follows with results. Section 7 provides some interpretation of the findings and concludes.

2.2 Literature Review

Worker well-being is connected to labor mobility through a matching-view of labor turnover. In the vast literature of labor mobility, a number of theoretical models have been proposed, generating mixed predictions about the direction of the effects of labor mobility on worker well-being. For example, in Jovanovic (1979)'s job matching model, jobs are treated as "search goods" (Nelson, 1974), the quality of worker-employer match being revealed ex ante. That mobility reflects worker's voluntary move to increasing high-quality matches predicts a positive effect of job mobility in terms of worker well-being. At the other extreme, jobs are "experience goods" in that match quality is learned over time as jobs are "experienced" (Jovanovic, 1979). Poor matches terminate while good matches tend to be preserved. The possibility of a sequence of "bad" matches leads to an ambiguous prediction of the effects of job mobility. Neal (1999) expands the notion of job match by distinguishing between employer match and career match. Assuming no learning about job matches or investment in matches, the optimal policy of workers is to locate a good career match first, and then search for a good employer match within the chosen career. Hence the first-stage job mobility is associated with increasing career-wide outcomes, and the second-stage mobility improves employer-specific outcomes.

Although the connection between job mobility and wage dynamics has attracted much attention in the empirical literature (for example, see Bartel, 1980; Borjas and Rosen, 1980; Bartel and Borjas, 1981; Borjas, 1984; Light and McGarry, 1998; Lillard, 1999; Davia, 2010), the relationship of job mobility and worker well-being has largely remained unexplored in the field of economics. The difficulty of measuring worker well-being at work place is possibly the main explanation for the scarcity of literature. The other reason is that the capacity of job satisfaction, the closest barometer of well-being, to serve as an indicator of the match quality has been questioned by economists, perhaps because it measures "what people say" rather than "what people do".

Among the existing studies, a number of studies have demonstrated that job satisfaction is a major determinant of labor market mobility. Hamermesh (1977) builds a testable model of overall job satisfaction and finds that job satisfaction predicts future quits. Freeman (1978) uses two sources of panel data, National Longitudinal Survey (NLS) and Panel Survey of Income Dynamics (PSID), to show that subjective expressions of job satisfaction, which captures unobservable aspects of the work place, are significantly related to future overt behavior. Akerlof et al. (1988) use changes in job satisfaction to demonstrate that non-pecuniary considerations motivate most "job-related" quits. Shields and Ward (2001) look at intention to quit and establishes the importance of job satisfaction in determining nurses' intentions to quit the NHS (British National Health Service). Clark (2001) uses the first seven waves of the British Household Panel Survey (BHPS) to illustrate that job satisfaction data are powerful predictors of both separations and quits. The findings of Kristensen and Westergard-Nielsen (2004), using European Community Household Panel (ECHP) data, also confirm that low overall job satisfaction significantly increases the probability of quit.

As for the effect of job mobility on job satisfaction, Bartel and Borjas (1981) show that people tend to exhibit increased satisfaction with the next job in the case of voluntary mobility. Akerlof et al. (1988) demonstrate that time-varying non-pecuniary rewards explain the fact that a significant fraction of quitters realize insignificant or negative wage changes and yet achieve significant gains in overall job satisfaction. Fasang et al. (2007) use the detailed information about labor market mobility and job satisfaction contained in a Eurobarometer survey to investigate the effects of job mobility on various aspects of job satisfaction: previous voluntary job change leads to increased satisfaction with objective work arrangement and quality of position, especially if the mobility is connected to a low number of unemployment spells and the application of same or more skills in the current job.

2.3 Data and Labor Mobility Patterns

I implement this empirical study using data from National Longitudinal Survey of Youth 79 (NLSY79). The original sample captures work, schooling, and many other experiences of 12686 young persons who were born between 1957 and 1964. To address the topic of early career job mobility, I examine the cohort under a six-year window, that is, from the 1982 survey round to the 1988 survey round.[1] Weekly job history data are used to construct people's job experience and related information up to the survey week of 1988. I keep in the sample only those who were working full time without gaps during the six years. Considering only continuously-working workers is a major step towards including only voluntary job switches in the analysis. The rationale is that involuntary job separations usually involve certain spell of unemployment, despite the fact that people can also quit into unemployment or start new jobs right after being laid off. This approach is a compromise with the difficulty of distinguishing between voluntary and involuntary job separations in NLSY79. In addition, I exclude respondents who did not report job satisfaction-related information in either 1982

[1] There are three reasons for selecting the time frame of 1982-1988. First, complete sets of job satisfaction measures, upon which worker well-being is based, are only available in four years: 1979-1982 and 1988. Second, starting from 1982 ensures that all workers were over 18 when they entered the sample. Finally, shortening time frame is a trade-off for larger sample size because the longer the observation window, the more probable that respondents are excluded for missing chains of work history.

or 1988 round of survey. The final sample size after deletion is 1811, with a mean age of approximately 22 in 1982.[2]

Table 1 reports the mean and standard deviation of key variables for my sample. The left panel reports summary statistics for time invariant variables. The education of workers is defined by the highest grade completed as of May 1 according to the 1982 round of survey. Since workers in my sample were working full time between 1982 and 1988, education is a time-invariant variable by construction. The average amount of schooling is less than 13 years, indicating that a significant portion of the workers are not college educated. The right panel shows that the percentage of being married rises immensely from 24% to 59% during this period, which is as expected considering the age range of these workers. Hourly rate of pay, measured in 1983 dollars, increases from 5.69 dollars per hour to 8.39 and the total hours at work increase as well. Therefore, the total income for an average worker increases from 1982 to 1988.

I construct work history related variables using the weekly work history data in NLSY79. I define the total number of employers within the observation window as the counts of all employers for whom the worker has worked over 30 hours per week for over one year from 1982 to 1988. I also count the total number of different types of job ever taken during this time. Theoretically, changing job types is best captured by changes in occupations because workers might be doing the same job across different industries. However, the occupation codes in NLSY79 contain many errors that imply false changes (Neal, 1999). To avoid exaggerating the frequency of occupation changes, I define job change involving both different occupations

[2]To adjust the probability of selection, I rescale the initial weights so that the weights for each demographic group in my sample add up to the total weights for the same demographic group in the whole NLSY79 sample. The adjusted weights are applied to all descriptive statistics and estimation results.

Table 1: Mean and Standard Deviation of Key Variables for the Whole Sample (N=1811)

Variable (time-invariant)		Variable (time-variant)	1982	1988
Age in 1982	21.81	Urban	0.80	0.80
	(2.10)		(0.40)	(0.40)
Male	0.51	Northeast	0.23	0.23
	(0.50)		(0.42)	(0.42)
Hispanic	0.05	South	0.36	0.36
	(0.23)		(0.48)	(0.48)
Black	0.09	West	0.14	0.15
	(0.29)		(0.35)	(0.35)
Education [a]	12.42	Married	0.24	0.59
	(1.62)		(0.43)	(0.49)
		Hourly rate of pay[b]	5.69	8.39
			(2.43)	(3.67)
		Hours per week	39.71	43.36
			(8.01)	(7.86)

Notes: Standard Deviations are in parenthesis.
[a] The education of workers is defined by the highest grade completed as of May 1 according to the 1982 round of survey. Since workers in my sample were working full time between 1982 and 1988, education is a time-invariant variable by construction.
[b] Hourly rate of pay in 1982 and 1988 are measured in 1983 dollars.

and different industries as the job type change. I define movers as those who have worked for more than one employer between 1982 and 1988. Among movers, complex job change occurs when workers not only change employers but also change tasks.[3]

I then couple the job history data between 1982 and 1988 with pre-1982 work history information to construct full-time working experience of workers when they entered the observation window, as well as the tenure with the current job at the 1982 and 1988 interviews. To construct the experience variable in 1982, I keep a running tally of all reported working weeks until the survey week of 1982, and then divide the cumulative number of weeks by 49, that is, the number of working weeks in a calendar year. To construct the tenure variables,

[3]I define movers with complex job change as those who have reported more than three industry codes and three occupation codes between the 1982 wave and the 1988 wave of the survey to minimize the effect of coding errors.

Table 2: Percentage of the Number of Employers and Complex Job Change by Prior Experience

	Experience<4	Experience=4	Experience>4
# of employers[a]			
1	21.77	38.40	52.38
2	40.87	35.73	30.19
3	29.40	20.83	13.78
4+	7.95	5.04	3.65
Complex change[b]	55.76	55.34	56.04

Notes: The experience is constructed using NLSY79 work history data. I keep a running tally of full-time working weeks for each respondent in the sample until the survey week of 1982 and then convert the results into years by dividing total number of weeks by 49. It is called prior experience because this is the accumulated experience at the time when respondents entered the observation window in this study. A majority of the respondents have a prior experience of four years, while the maximum reaches nine.

[a] The percentage reflects the proportion over all workers within the same experience category.

[b] The percentage reflects the proportion over movers within the same experience category.

I trace the job back to its starting week, count the total full-time working weeks in that job, and convert the results into years.

Distributions of the total number of major employers and types of job change, broken down into three experience categories, are displayed in Table 2. The occurrence of job separations, that is, the total number of reported employers equalling two or larger, decreases from 78.23% for the junior group, then to 61.60% for the next, and finally down to 47.62% for the senior group, highlighting a negative relationship between job mobility and experience. Further, a significant portion of the job changes involve changes in both employers and occupations because for all three experience groups, the percentages of complex change are over 50%.

Table 3 shows the direction of occupation changes. For ease of exposition, I group the original 1970 three-digit occupation code into six broad categories (Flyer, 1997): crafts and

Table 3: Distribution of Occupation Movements from 1982 to 1988 among Movers

1982 Job	1988 Job: Crafts, Specialty Skills	Business, Managerial Professions	Others
Craft, Specialty Skills	58.16	17.07	4.92
Business, Managerial Professions	5.20	7.73	0.94
Others	1.78	1.44	2.76

Notes: Numbers represent the weighted percentages for each category of job mobility. Following Flyer (1997), I group the original 1970 three-digit occupation code into six broad categories: crafts and specialty skilled workers; business and managerial professionals; engineering-mathematics-hard sciences; social sciences; health care; and humanities, arts and entertainments. Upon inspection of the sample, large concentration in the first two groups allows me to further simplify the classification by combining the other four categories.

specialty skilled workers; business and managerial professionals; engineering-mathematics-hard sciences; social sciences; health care; and humanities, arts and entertainments. Given that majority of the observations are concentrated in the first two groups, I combine the rest four groups. Even for such broad occupational groupings, over 31% of the movers switched occupations from 1982 to 1988, with a general tendency of moving towards managerial professions.

Table 4 shows the differences between movers and stayers prior to the job change. Movers as a group were younger on average, more likely to be single and had a shorter employment history than the stayers in 1982. With respect to the current job, movers earned less and had shorter tenure than stayers. 70% of the movers had employment-based health insurance, in comparison to 85% for stayers. The propensity to leave the current job is measured by a survey question phrased as: "If given the opportunity, would you take a different job?". 70% of the movers expressed the willingness to take another job when given a chance while only 50% of the stayers did so. In addition, a series of questions regarding characteristics of the current job such as amount of variety were also delivered in the survey and the respondents were asked to rate each characteristic using five scales: 1=minimum amount, 2=not too much, 3=moderate amount, 4=quite a lot, 5=maximum amount. There are only marginal differences between the two groups regarding most aspects of characteristics except self-assessed job significance and perhaps job autonomy. As expected, stayers rated the significance and autonomy of their job higher than movers.

2.4 Constructing Measures of Worker Well-being

Self-reported job satisfaction is the traditional gauge of worker well-being at workplace. As yet, relatively few economists have dealt with these subjective variables, despite of a vast literature on job satisfaction in other fields such as sociology and industrial psychology. Although subjective measures should be treated with a bit of caution, they provide direct means for the assessment of preferences. Well-being is, to some extent, subjective in that what is required in order to be happy depends on mental state and thus differs across individuals; subjective measures such as job satisfaction convey valuable information regarding the result of workers weighing their own mind of the job with current situations.

In this section, I measure worker well-being by piecing together a jigsaw puzzle reflecting worker's appraisal of the entire panoply of job characteristics. Overall job satisfaction serves as a barometer of worker's overall well-being at work place; yet it is the resultant of an unknown and potentially diversified process of evaluation formation process due to different job attitudes and beliefs. To conquer this drawback of overall job satisfaction, I use satisfaction regarding different facets of job to provide a more clearly-defined measure of worker well-being. The facets of job satisfaction considered in the following analysis include: pay,

Table 4: Means of Key Variables for Movers and Stayers, Prior to the Job Change

Variable	Range	Movers	Stayers	p [a]
1. *General*				
Male	–	0.53	0.47	0.072
Age	–	21.44	22.56	0.000
Hispanic	–	0.05	0.06	0.074
Black	–	0.08	0.12	0.008
Married	–	0.21	0.31	0.001
Education	–	12.40	12.46	0.489
Urban	–	0.81	0.79	0.506
South	–	0.36	0.38	0.541
West	–	0.16	0.10	0.001
Northeast	–	0.24	0.21	0.254
Aptitude score	0-100	51.86	48.89	0.054
Years of full-time experience	–	3.31	4.09	0.000
2. *Job at the time of the survey in 1982*				
Hourly rate of pay	–	5.36	6.36	0.000
Hours per week	–	39.34	40.45	0.018
Tenure	–	1.54	2.31	0.000
Employment-based health insurance	–	0.70	0.85	0.000
Propensity to leave another type of job	–	0.70	0.59	0.000
Learning general skills at work	–	0.48	0.49	0.627
Job Variety	1-5	3.17	3.29	0.106
Chance to deal with other people	1-5	3.73	3.74	0.893
Job autonomy	1-5	3.26	3.39	0.063
Chance to develop friendships at job	1-5	3.50	3.55	0.446
Chance to do a complete task	1-5	3.90	3.89	0.833
Self-assessed job significance	1-5	3.38	3.63	0.000
Amount of feedback from work	1-5	3.79	3.84	0.459

Notes: See text for definitions of variables.

[a] p-value for the equality test of means for the movers and stayers.

whether given the change to do what one is good at, job security, promotion opportunities, physical surroundings of the job, relationship with colleagues and whether the supervisors are competent and supportive.

2.4.1 Methodological Framework

Questionnaires designed to measure job satisfaction typically use Likert scales as a response format; that is, respondents are asked to rate their opinions on a scale consisting of a given number of ordered categories. Data arising from Likert-type items are often analyzed as multivariate normal outcomes in models such as confirmatory factor analysis. However, assigning successive integer values to scale categories often produces distributions for which the multivariate normal assumptions are not satisfied. For example, people's tendency of avoiding strongly negative ratings will drive most of the observation towards the positive end of the scale, so the mass of the actual distribution is concentrated on the right. In light of the aforementioned problems, I start with transforming Likert scale data into more realistic and normally distributed data for subsequent analysis using a method proposed by Vijverberg (2004).

Denote L_j as the respondent's answer to question j ($j = 1, 2, \ldots, P$) and $L_j = l$ ($l = 1, 2, \ldots, M$) if the respondent selects category l on a Likert scale. It is assumed that L_j is the revealed expression of the respondent's latent opinion x_j about job aspect j and $L_j = l$ corresponds to an interval x_{jl-1} to x_{jl} on the latent scale. The relative location of cut-off points x_{jl} reflects the way respondents process each Likert-type scale category; hence the set of x_{jl} is likely to vary across different groups of people.

Denoting the underlying distribution function by $P(x) = Pr(X \leq x)$, then the probability of observing category l can be represented as:

$$Pr(L_j = l) = P(x_{jl}) - P(x_{jl-1}), \tag{18}$$

Thus the sample counterpart of $P(x_{jl})$, denoted as $\hat{P}_{jl}$, can be calculated from the observed proportions for each category. Further, if the latent distribution is standard normal, then interval boundaries can be inferred via:

$$x_{jl} = \Phi^{-1}(\hat{P}_{jl}). \tag{19}$$

Values contained in the estimated class boundaries represent all possible underlying continuous measurement of job satisfaction for each reported scale, and I take the conditional mean as scores, so that

$$s_{jl} = E[X|x_{jl-1} < X < x_{jl}] = \frac{\phi(x_{jl-1}) - \phi(x_{jl})}{\Phi(x_{jl}) - \Phi(x_{jl-1})} \tag{20}$$

where

$$x_{j0} = -\infty \quad x_{jM} = \infty.$$

With the value of s_{jl} being identified, I suppress the subscript l when referring to job satisfaction scores in the following analysis for brevity of representation. I next posit that a group of latent variables ξ_k $(k = 1, \ldots, H)$ with a restricted number of dimension H $(H < P)$ exists that identifies worker well-being. Although ξ_k is unobserved, it is manifested by various clusters of job satisfaction scores, s_j. The underlying structure between latent variable ξ_k and observed variable s_j can be written as:

$$\mathbf{s_i} = \boldsymbol{\mu} + \boldsymbol{\Lambda\xi_i} + \boldsymbol{\delta_i}, \tag{21}$$

where subscript i denotes respondent i $(i = 1, 2, \ldots, N)$; $\boldsymbol{s_i}$ is $P \times 1$ vector of job satisfaction scores; $\boldsymbol{\mu}$ is $P \times 1$ vector of intercepts; $\boldsymbol{\Lambda}$ is $P \times H$ matrix of factor loadings; $\boldsymbol{\xi_i}$ is $H \times 1$ vector of latent variables; and $\boldsymbol{\delta}_i$ is $P \times 1$ vector of measurement errors. The measurement errors, $\boldsymbol{\delta}$, are assumed to be independent of the factors, $\boldsymbol{\xi}$. Job satisfaction scores purported to be manifestations of the same ξ_k load on the same factor, so $\boldsymbol{\Lambda}$ has a strict structure with zeros in several places. In the context of this paper, $\boldsymbol{\Lambda}$ has the following block structure because each observed variable loads on only one factor:

$$\boldsymbol{\Lambda} = \begin{pmatrix} \boldsymbol{\Lambda_1} & 0 & \ldots & 0 \\ 0 & \boldsymbol{\Lambda_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \boldsymbol{\Lambda_H} \end{pmatrix}.$$

Without a priori knowledge of the structure, I perform an exploratory factor analysis (EFA) to intuit the number of factors, as well as the group of variables to load on each factor before proceeding to estimation of equation (21). Then I employ confirmatory factor analysis (CFA) to identify the model and check its adaptability to my data. The reason why estimation can not be done with standard regression methods is that $\boldsymbol{\xi}$ is not observable. Factor analysis is the best choice because it works with the covariance matrix of observed variables, which can be written as: (Bollen 1989)

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda E(\xi\xi')\Lambda'} + \boldsymbol{E(\delta\delta')}.$$

The structural model is evaluated using a STATA program written by Kolenikov (2009). With factor structure identified and verified, satisfaction scores can be divided into H groups, so that scores in the same group are highly correlated with each other and originate from the same factor. Each factor represents one dimension of worker well-being. Following Betti, D'Agostino and Neri (2010), I then construct worker well-being of respondent i in regard to facet k as:

$$w_{ik} = \frac{\sum_{j \in \Xi_k} \lambda_{jk} s_{ij}}{\sum_{j \in \Xi_k} \lambda_{jk}}. \tag{22}$$

where

$$\Xi_k = \{j | \lambda_{jk} \neq 0\}$$

To yield a more interpretable result, I rescale w_{ik} so that the values span $[0, 1]$ using:

$$w_{ik}^{*} = \frac{w_{ik} - w_k^{min}}{w_k^{max} - w_k^{min}}, \tag{23}$$

where w_k^{min} and w_k^{max} are the minimum and maximum values of w_k across the observations. Workers with the highest level of well-being will be scored 1, while those with lowest level scored 0. Values between 0 and 1 measure the relative degree of happiness at the workplace.

2.4.2 Implementation of the Factor Analysis

The use of a two-year panel for my NLSY79 sample raises a rarely addressed yet essential issue as far as inter-temporal comparison is concerned: longitudinal measurement invariance.[4] In the application of this study, the measurement of worker well-being is subject to changes in the construct, which may involve the total number of dimensions, the group of variables underlying each dimension of latent well-being, as well as the loadings for each

[4]Chan (1998) outlines three types of change that may be encountered in repeated measurements: alpha, beta, and gamma change. Alpha change refers to true score change in a construct given a constant conceptual domain and constant measurement. Beta change occurs in instances where the construct of interest remains constant, but the measurement properties of the indicators of the construct are temporally inconsistent. Gamma change occurs when the meaning of the construct changes over time.

variable. Actually, as workers grow more mature and gain more experience, the aspects of job they value will switch, leading to changes in the well-being construct. Therefore, when measurement is not guaranteed to be invariant over time, it is misleading to analyze and interpret the temporal changes in worker well-being as a result of mobility; it can not be determined whether the observed change is due to true change or merely to changes in the construct of measurements of worker well-being. To accommodate any possible variations in the construct, I apply the well-being construct in 1982 to both years, so the temporal differences in well-being will not be attributed to changes in the construct.

During the 1982 and 1988 rounds of the survey, questions concerning job satisfaction were delivered through statements of the type “the pay is good” and workers were asked to choose from one of the four ordered categories which are phrased as “not true at all”, “not too true”, “somewhat true” and “very true”. Seven aspects of job satisfaction are investigated in this study: pay, whether given the chance to play to one’s strength, job security, promotion opportunities, physical surroundings, relationship with colleagues, and the competency of supervisors.

To convert those Likert-type scales, I start by dividing the sample into four groups defined by age and gender. Specifically, I define the groups “young men” and “young women” as male and female respondents who were between 17 and 21 as of 1982, the groups “older men” and “older women” as male and female respondents who were between 22 and 27 as of 1982, respectively. The rationale for computing scores separately within each subgroup is that there might be systematic differences in the manner workers respond to the questions that will potentially bias the estimates. These differences are assumed away within each subgroup. Table 5 presents the scores for each category using the aforementioned methods. s_1, s_2 and

s_3 are mostly below zero, showing that majority of the workers answer those questions quite positively. Moreover, unlike arbitrary assigned successive integer scales, these scores based on the data are not equally spaced; differences from one lower category to the next one increase progressively; distances from s_3 to s_4 are almost twice as large as those from s_1 to s_2.

The following analysis applies to the whole sample instead of the four subgroups separately. Results of exploratory factor analysis in Table 6 show that the eigenvalue for the first factor is 2.0468, while that of the second factor rapidly drops to 0.4570. At the same time, the first factor explains almost 70% of the total variance. Therefore, I retain one factor according to the Kaiser criterion, which suggests dropping all components with eigenvalues below one.

I proceed to estimate the parameters of this one-factor model via the maximum likelihood method traceable to the Jöreskog (1967)'s pioneering work. Table 7 reports the estimates of factor loadings for the specified model. For the purpose of identification, one of the loadings is set to one. Every freely estimated parameter is statistically significant. Moreover, there are no unreasonable parameter estimates such as Heywood cases; the direction and size of estimates are in accord with prediction. To facilitate interpretation, I standardize the estimates by

$$\lambda_j^* = \lambda_j \times \sqrt{\frac{Var(y_j)}{Var(\xi)}},$$

where λ_j^* and λ_j are standardized and unstandardized loadings for satisfaction score s_j respectively. The standardized loadings can be interpreted as standardized regression coefficients. For example, a one standardized score increase in well-being is associated with a

Table 5: Category Scores for Each Job Satisfaction Item

Item	s_1	s_2	s_3	s_4
1. *Young men*[a]				
Job playing to strength	-2.10	-1.33	-0.39	0.92
Pay	-2.04	-1.12	-0.06	1.23
Job security	-2.12	-1.34	-0.43	0.86
Promotion opportunities	-1.86	-0.96	-0.1	1.1
Physical surroundings	-2.19	-1.37	-0.35	0.98
Relationship with colleagues	-3.11	-2.35	-1.1	0.48
Competency of supervisor	-2.32	-1.62	-0.73	0.67
2. *Young women*[b]				
Job playing to strength	-2.43	-1.57	-0.46	0.91
Pay	-1.99	-1.20	-0.11	1.25
Job security	-2.25	-1.51	-0.59	0.75
Promotion opportunities	-1.68	-0.78	-0.02	1.14
Physical surroundings	-2.25	-1.49	-0.57	0.77
Relationship with colleagues	–	–	-1.09	0.46
Competency of supervisor	-2.24	-1.57	-0.87	0.51
3. *Older men*[c]				
Job playing to strength	-2.27	-1.51	-0.50	0.85
Pay	-2.07	-1.29	-0.21	1.14
Job security	-2.17	-1.37	-0.42	0.89
Promotion opportunities	-1.76	-0.87	-0.04	1.14
Physical surroundings	-2.22	-1.40	-0.35	0.98
Relationship with colleagues	-2.96	-2.31	-1.03	0.54
Competency of supervisor	-2.23	-1.57	-0.72	0.67
4. *Older women*[d]				
Job playing to strength	-2.20	-1.52	-0.49	0.88
Pay	-1.89	-1.11	-0.08	1.25
Job security	-2.24	-1.52	-0.58	0.77
Promotion opportunities	-1.61	-0.67	0.17	1.28
Physical surroundings	-2.25	-1.42	-0.50	0.81
Relationship with colleagues	-2.76	-2.15	-0.95	0.58
Competency of supervisor	-2.41	-1.66	-0.72	0.67

[a] Male respondents who were between 17 and 21 as of 1982.
[b] Female respondents who were between 17 and 21 as of 1982.
[a] Male respondents who were between 22 and 27 as of 1982.
[a] Female respondents who were between 22 and 27 as of 1982.

Table 6: Exploratory Factor Analysis

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor 1	2.0468	1.5898	0.6999	0.6999
Factor 2	0.4570	0.2469	0.1563	0.8562

Notes: According to Kaiser criterion, only one factor is retained because only the first factor is greater than one. This factor explains almost 70% of the total variance.

0.5148 standardized score increase in satisfaction score towards pay. Moreover, since there is only one factor in this case, the standardized factor loadings can also be interpreted as the correlation of the satisfaction score with underlying worker well-being. Satisfaction regarding promotion opportunities has the highest loadings, underscoring the role of career prospects as a major determinant of worker well-being at this stage of career. However, the goodness-of-fit indices do not provide satisfactory results, perhaps due to large sample size. For example, both RMSEA and CFI fall outside the rule-of-thumb range of a good fit, while the χ^2 test is significant.[5]

2.5 Framework for Assessing Effects of Labor Mobility

In this section, I lay out a framework for assessing the influence of voluntary labor mobility on worker well-being. There are two sources of complication:

1. unobserved factors that are relevant and known to the individual rather than the researcher. Examples include unobserved taste for job, productivity, proportion of skills transferable across jobs, and reservation level of happiness.

2. unknowable counterfactual outcomes, that is, the labor market outcomes of movers

[5] To investigate the extent of measurement invariance, I repeat the same analysis using the 1988 round of data. The result of EFA also suggest a one-factor model. Further, confirmatory factor analysis shows that the chance of promotion is still the most influential element in determining worker well-being at the workplace.

Table 7: ML Standardized Estimates of Factor Loadings: NLSY79 (1982)

	Factor loadings	
Indicators	Unstandardized	Standardized
Job playing to strength	1.0000**	0.5159
	(–)	
Pay	1.0558**	0.5288
	(0.0723)	
Job security	1.0940**	0.5689
	(0.0730)	
Promotion opportunities	1.1627**	0.5896
	(0.0741)	
Physical surroundings	0.9654**	0.5190
	(0.0646)	
Relationship with colleagues	0.6905**	0.4247
	(0.0546)	
Competency of supervisor	0.8707**	0.4898
	(0.0614)	
$\chi^2(14)$[a]	212.047	
RMSEA[b]	0.089	
RMSEA 90% CI	0.079 – 0.100	
CFI[c]	0.8739	

Note: Standard errors for unstandardized loadings are in parenthesis. Significance levels: † : 10% $*$: 5% $**$: 1% .

[a] The most common kind of absolute fit index. It evaluates the reasonability of the hypothesis of $S = \Sigma$, where S is the sample covariance matrix (Brown, 2006).

[b] Root mean square error of approximation (Seitger and Lind, 1980). It assesses the extent to which a model fits reasonably well in the population. Browne and Cudeck (1993) propose that RMSEA less than 0.08 suggest adequate model fit.

[c] Comparative fit index (Bentler, 1990). It evaluates model fit against a solution positing no relationships among the variables. CFI values in the range of 0.90–0.95 may be indicative of acceptable model fit.

had they not gone through mobility.

I tackle both problems with difference-in-differences matching strategy, developed in Heckman, Ichimura and Todd (1997, 1998) and Heckman, Ichimura, Smith and Todd (1998). This section describes why and how the strategy is applied in the context of this research.

2.5.1 The Rationale for the Difference in Differences Matching Estimator

Following the terminology of quasi-experiment literature, I treat voluntary labor mobility as a type of binary treatment denoted by D_i. D_i equals one if worker i switch jobs within the observation window. The outcome of interest, worker well-being, is denoted by w_i. It is assumed that there are two potential outcomes of w_i that are defined as:

$$\text{Potential Outcomes} = \begin{cases} w_{1i} & \text{if } D_i = 1 \\ w_{0i} & \text{if } D_i = 0 \end{cases}$$

Since workers can only be in one state at a time, only one of the potential outcomes is observed. w_i can be written as: $w_i = D_i w_{1i} + (1 - D_i) w_{0i}$. The question of interest here is the average effect of labor mobility on movers, namely, the average treatment effect on the treated (ATT).

Applying traditional difference-in-differences (DID) method yields the difference between stayers and movers in the before-and-after difference in outcomes. DID is a version of fixed effects estimation, which solves unobserved heterogeneity problem if the unobserved confounders are time invariant (Angrist and Pischke, 2009). However, interpreting DID estimates as ATT requires the overly strong assumption that movers' outcome would have followed the same trend as stayers in the absence of treatment.

The link between ATT and DID can be conveniently represented by the equation:

$$\underbrace{E[(w_{1t} - w_{0t'})|D = 1] - E[(w_{0t} - w_{0t'})|D = 0]}_{\text{DID}} = \tag{24}$$

$$\underbrace{E[(w_{1t} - w_{0t'})|D = 1] - E[(w_{0t} - w_{0t'})|D = 1]}_{\text{ATT}}$$

$$+ \underbrace{E[(w_{0t} - w_{0t'})|D = 1] - E[(w_{0t} - w_{0t'})|D = 0]}_{\text{Selection bias}},$$

where t' and t are time periods at the beginning and end of the observation window, respectively. The ATT term captures the average difference between the outcome of movers, $E[(w_{1t} - w_{0t'})|D = 1]$, and what would have happened to them had they not switched jobs, $E[(w_{0t} - w_{0t'})|D = 1]$. The DID estimator, however, add to this causal effect a term called selection bias. DID equals ATT only if there is no selection bias.

The ideal way to solve the selection problem is randomly assigning workers to mover and stayer groups. However, workers do not change jobs randomly; those who foresee stagnant income growth with the current employer are more likely to seek alternative opportunities, making the selection bias negative in this case. However, selection bias could go the other way, if the mobile individuals belong to a group with advantageous skill set and self-select into better jobs (Fasang et al., 2007). Given that matching can be utilized to infer counterfactual outcome distributions for movers when random assignments are infeasible, I implement the DID matching estimator by defining DID conditional on X, so that:

$$E[(w_{0t} - w_{0t'})|X, D = 1] = E[(w_{0t} - w_{0t'})|X, D = 0], \tag{25}$$

where X is a vector of observable individual characteristics. This condition asserts that the labor market outcome of movers would have evolved from pre-treatment to post-treatment in the same way as it would have for an observably similar stayer had he (or she) not switched jobs. Accordingly, DID matching estimator converges under standard conditions to ATT, which is:

$$E[(w_{1t} - w_{0t'})|X, D = 1] - E[(w_{0t} - w_{0t'})|X, D = 1].$$

2.5.2 Key Identifying Assumptions

Two key assumptions about treatment assignment are required for the validity of DID matching estimator. The first assumption states that conditional on X, $w_{0t} - w_{0t'}$ and D are independent.

$$(w_{0t} - w_{0t'}) \perp D|X, \tag{A-1}$$

where "$\perp$" denotes independence.[6] This is a DID version of Condition Independence Assumption (CIA) for the controls. If (A-1) is true, then conditional on X outcomes of stayers have the same distribution as movers would have experienced if they had not switched jobs. As a consequence, (25) follows if mean exists. In addition, I make a second assumption regarding the joint distribution of treatments and covariates:

$$Pr(D = 1|X) < 1. \tag{A-2}$$

(A-2) guarantees the possibility of selecting a control group from the pool of stayers. Hence, (A-2) restricts matching to be performed over a common support region.

The DID version of CIA allows for selection on unobservables as long as those confounders are time-invariant. For example, workers who are easily fulfilled tend to report job satisfaction more positively and are less likely to change jobs than those who are critical. This type of attribute is the part of personality that will not alter much so I can safely treat it as

[6] Although the mean independence assumption given in (25) is weaker and sufficient in my use, it is rare that a convincing case is made for this weaker assumption without being equally strong for the stronger version (A-1) (Imbens, 2004).

time-invariant. Time-variant unobserved covariates, on the other hand, can also be accommodated in the model without invalidating (A-1) if the optimization objective of the worker is distinct from the outcome that is of interest to the econometrician. If this is true, then two agents with the same values for observed characteristics may differ in treatment choices if the difference is driven by unobserved characteristics that are unrelated to the outcomes of interest (Imbens, 2004). Lastly, imposing restrictions on the way individuals form their expectations about the unknown potential outcomes provides another way to justify (A-1) (Heckman, Lalonde and Smith, 1999). However, (A-1) is still a strong assumption in that it precludes selection on time-variant unobservables that affect both $w_{0t} - w_{0t'}$ and D. It is therefore demanding of the data in that a rich set of X is required to ensure (A-1) is satisfied.

Implementing matching requires choosing a set of variables that satisfy both (A-1) and (A-2). The ideal set of X should include all of the key variables so that the exposure to treatment is independent of $w_{0t} - w_{0t'}$ among individuals with the same values of X. This requirement highlights the advantage of using NLSY79 because the data set provides a large reservoir of potential conditioning variables. On the other hand, the set of X should not be so inclusive that (A-2) is violated due to lack of common support. This requirement can be statistically tested by estimating a logit model for labor mobility and inspecting the estimated distribution of $\widehat{Pr}(D = 1|X)$ for both groups to detect any lack of common support. Further, variables that have no or only weak correlation with either D or $w_{0t} - w_{0t'}$ shouldn't be adjusted for because they may reduce precision. A particular concern in the literature is the inclusion of covariates that are themselves affected by the treatment (Imbens, 2004). This problem can be easily avoided in NLSY79 in that information is collected year by year rather than retrospectively.

2.5.3 DID Matching Estimator

I use the covariate matching method proposed in Abadie and Imbens (2006) and implemented through Abadie, Drukker, Herr and Imbens (2004). The basic idea is to impute the missing potential outcomes using the outcome of nearest neighbors of the opposite treatment group. Specifically, for each treated i, the matching algorithm finds all untreated units j so that:

$$\sum_{j:D_j=0} \omega_j \mathbf{1}\{||X_j - X_i|| \leq d_M^{\omega}(i)\} = M \tag{26}$$

where ω_j contains information about the probability of worker j being sampled from the underlying population.[7] $d_M^{\omega}(i)$ is the distance from the covariates of person i, X_i, to the Mth nearest match in the opposite treatment group while accounting for the sampling weight. The vector norm $||\cdot||$, which measures the distance between two covariates vectors, takes the form:

$$||X_j - X_i|| = (X_j - X_i)'\text{diag}(\Sigma_X^{-1})(X_j - X_i)$$

where Σ_X is the covariance matrix of the covariates.

Further, let $\mathcal{J}_M^{\omega}(i)$ denote the set of all indices satisfying (26). Under (A-1) and (A-2), ATT can be estimated by DID matching as follows:[8]

$$\text{ATT} = \frac{1}{\sum_{i=1}^{N_1} \omega_i} \sum_{i=1}^{N} \omega_i [(w_{it} - w_{it'}) - \frac{1}{\sum_{j \in \mathcal{J}_M^{\omega}(i)} w_j} \sum_{l \in \mathcal{J}_M^{\omega}(i)} \omega_l (w_{lt} - w_{lt'})] \tag{27}$$

where N is the total number of workers in the treatment group.

[7]The original sampling weight is normalized to sum to sample size N to form ω here.
[8]The analytical variance of the estimator is given in Abadie and Imbens (2006).

2.5.4 Bias Correction

According to Abadie and Imbens (2006), the estimator in (27) will have a term corresponding to the matching discrepancies that will be of the order $O_p(N^{-1/k})$, where k is the number of continuous covariates. To remove some of this bias term, I adopt a bias-correction procedure proposed by Abadie et al. (2004).

For brevity of exposition, I denote the outcome variable, $w_{it} - w_{it'}$, as Δw_i. Further, I denote the matches for the treated person i as $\ell_M(i)$. The correction is motivated by the fact that the covariates X_i and $X_{\ell_M(i)}$ are not equal, although they are close after matching. Let $\mu_0(x) = E[\Delta w | D = 0, X = x]$, the imputed value $\widehat{\Delta w_{0i}}$ is unbiased for $\mu_0(X_{\ell_M(i)})$, but not necessarily for $\mu_o(X_i)$. To adjust the difference $\mu_o(X_i) - \mu_0(X_{\ell_M(i)})$, a linear regression function for $\mu_o(x)$ is estimated using observations that are used as matches for the treated unites, weighted by $K_M^\omega(i)$. $K_M^\omega(i)$ is the number of times i is used as a match for all observations l of the opposite treatment group, each time weighted by the total number of matches for l:

$$K_M^\omega(i) = \sum_{l=1}^{N} \omega_l \mathbf{1}\{i \in \mathcal{J}_M^\omega(l)\} \frac{\omega_i}{\sum_{j \in \mathcal{J}_M^\omega(l)} \omega_j}$$

The matching estimator is then modified to:

$$\text{ATT} = \frac{1}{\sum_{i=1}^{N_1} \omega_i} \sum_{i=1}^{N} \omega_i [\Delta w_i - \frac{1}{\sum_{j \in \mathcal{J}_M^\omega(i)} w_j} \sum_{l \in \mathcal{J}_M^\omega(i)} \omega_l (\Delta w_l + \hat{\mu}_0(X_i) - \hat{\mu}_0(X_l))] \tag{28}$$

2.6 Results

In this section, I present DID results both before and after matching. Throughout the analysis, I differentiate two types of job move: complex job change and simple job change. A complex job change occurs when workers not only change employers but also change tasks;

a worker at this stage is more concerned about how well she is suited to the type of work she is doing. A simple job change refers to firm-level separations when workers change employers but continue doing the same line of work; the main concern of worker at this stage is how well she is suited to the work environment created by the firm (Neal, 1999). Again, the emphasis is the the effect of job mobility on the well-being of movers who *voluntarily* change jobs; as mentioned in Section 3, these individuals are identified by job changes without any intervening joblessness.

2.6.1 Difference in Differences before Matching

Panel 1 of Table 8 summarizes the levels and changes in worker well-being from 1982 to 1988 for stayers, movers with complex change and movers with simple change. The first two rows of column (3) demonstrate the differences in well-being between complex job changers and stayers in 1982 and 1988, respectively. Complex job changers were initially less satisfied with their jobs whereas in 1988 they became happier than their staying counterparts. Inspection of row (3) reveals that the well-being of stayers dropped by 0.0211 from 1982 to 1988, while that of complex job changers increased by 0.0515. Row (3) of Column (3) presents the DID of the changes in well-being; the relative gain for those who took complex job moves was 0.0726. The well-being of workers with simple job change was also rising from 1982 to 1988, albeit moderately with a relative gain of 0.0382. To check if the results are robust to other outcomes related to worker well-being, I repeat the same analysis in Panel 2 and Panel 3 for *hourly rate of pay* and *proportion of highly satisfied*[9]. Panel 2 shows that workers with simple

[9]Respondents in NLSY79 were asked to rate their overall satisfaction using four scales towards the current job in each survey round: highly dissatisfied, somewhat dissatisfied, somewhat satisfied, high satisfied. The variable *highly satisfied* is defined as the proportion of respondent who evaluate their current job as "highly satisfied".

job change experienced the highest wage growth during that period, followed by workers with complex job change and stayers, respectively. The results for *proportion of highly satisfied* mimic that of worker well-being in that it is essentially another way of measuring the degree of happiness at workplace.

Despite their intuitive appeal, the above direct DID results should not be treated as the average treatment effect of voluntary job mobility on the movers because it is likely that the trends of worker well-being for movers in the absence of treatment would be different from that of stayers. One reason is that stayers were on average older and better paid in 1982 than movers; thus I suspect that a significant part of the stayers already had a good job match by the time they entered the sample. I investigate this possibility using the data of hourly rate of pay on multiple years. Figure 1 shows that the wage growth of stayers was steeper than that of both types of movers at first, perhaps a resultant of greater investment in firm-specific skills by those stayers who planned to stay where they were. The wage of both types of movers then caught up quickly when they found their match, which is a further evidence that movers found their match later than most stayers.

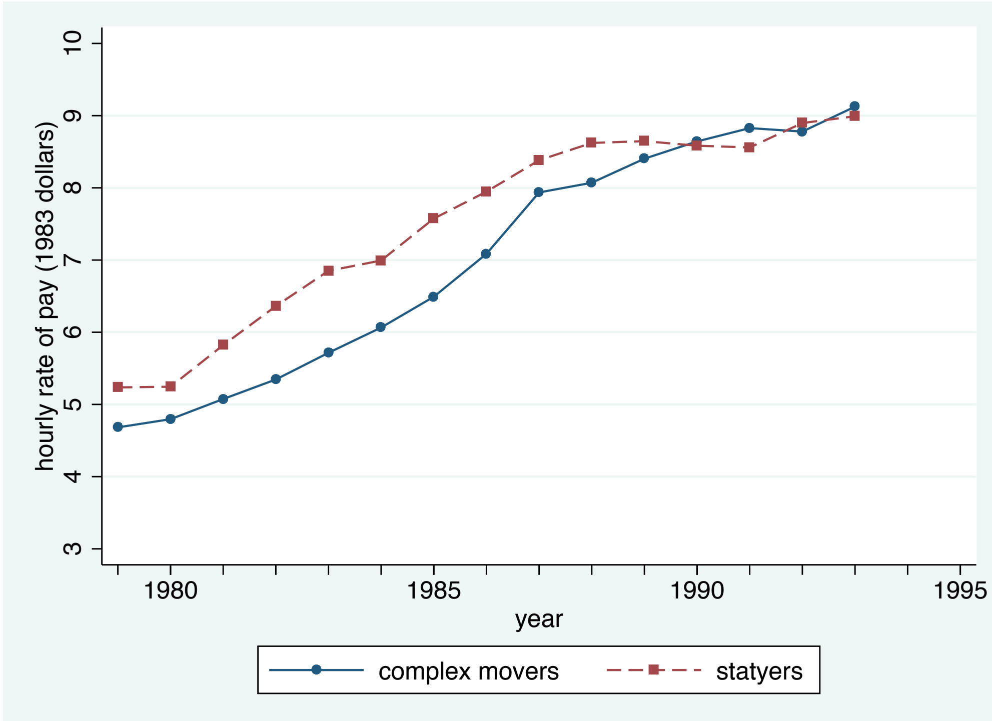
2.6.2 Choosing Conditional Variables

Choosing conditional covariates is the key step to ensure the quality of matching. Omitting important variables can seriously increase bias in resulting estimates, but overflow of variables should also be avoided for considerations of efficiency loss. Only variables that influence simultaneously the job change decision and the outcome variable should be included (Caliendo and Kopeinig, 2008). I start with a logit model of switching jobs. Rather than allowing a mechanical algorithm to determine the specification of the job change model, I

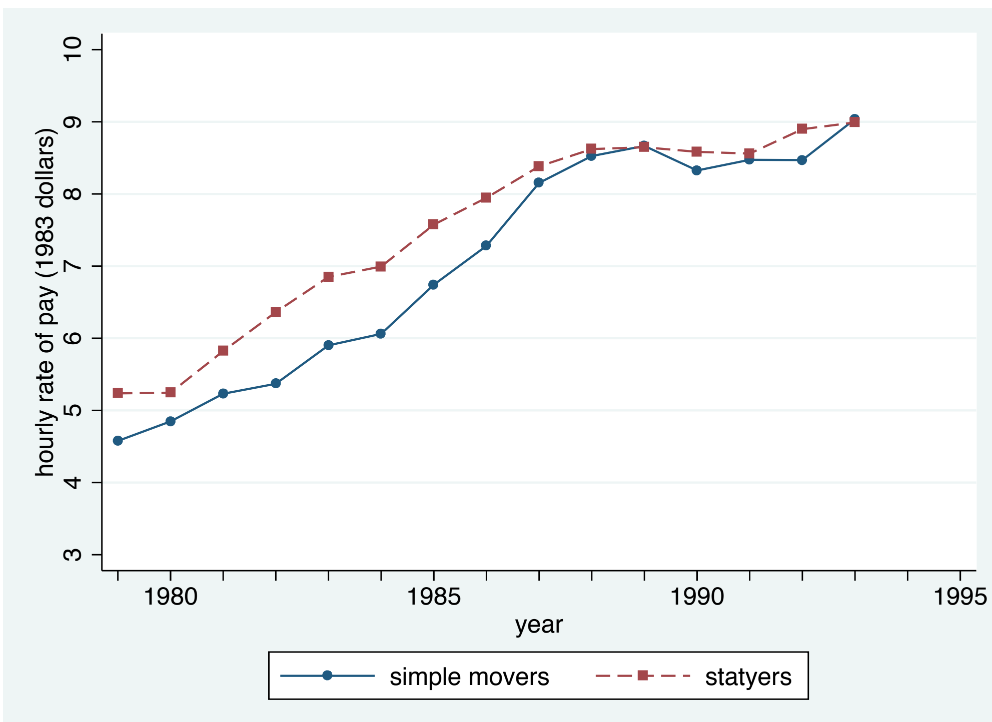
Table 8: DID Analysis Before Matching

		Movers			
	Stayers	Complex change	Difference[a]	Simple change	Difference[b]
Variable	(1)	(2)	(3)	(4)	(5)
1. *Worker well-being*					
(1) 1982	0.7234	0.6764	-0.0471	0.7095	-0.0139
	(0.0082)	(0.0090)	(0.0122)	(0.0104)	(0.0133)
(2) 1988	0.7023	0.7279	0.0256	0.7266	0.0242
	(0.0089)	(0.0082)	(0.0121)	(0.0094)	(0.0129)
(3) Change	-0.0211	0.0515	0.0726	0.0171	0.0382
	(0.0121)	(0.0122)	(0.0172)	(0.0140)	(0.0185)
2. *Hourly rate of pay*					
(4) 1982	6.3621	5.3456	-1.0165	5.3687	-0.9934
	(0.1297)	(0.1056)	(0.1674)	(0.1225)	(0.1785)
(5) 1988	8.6216	8.0712	-0.5515	8.5253	-0.0964
	(0.1721)	(0.1606)	(0.2355)	(0.2393)	(0.2949)
(6) Change	2.2596	2.7246	0.4650	3.1565	0.8970
	(0.2156)	(0.1923)	(0.2890)	(0.2690)	(0.3447)
3. *Proportion of highly satisfied*					
(7) 1982	0.4959	0.3547	-0.1412	0.4231	-0.0729
	(0.0251)	(0.0248)	(0.0353)	(0.0289)	(0.0383)
(8) 1988	0.4063	0.4544	0.0481	0.4480	0.0417
	(0.0249)	(0.0257)	(0.0358)	(0.0292)	(0.0384)
(9) Change	-0.0897	-0.0996	0.1893	0.0249	0.1145
	(0.0354)	(0.0357)	(0.0503)	(0.0411)	(0.0542)

Notes: Standard errors are given in parentheses. Observations are weighted by sampling weights.
[a] The difference between movers with complex job change and stayers. Complex job change occurs when workers not only change employers but also change tasks.
[b] The difference between movers with simple job change and stayers. Simple job change refers to the case when workers change employers but continue doing the same line of work.



(a) Comparison between Stayers and Complex Movers



(b) Comparison between Stayers and Simple Movers

Figure 1: Average Hourly Rate of Pay from 1979 to 1993: (a) comparison between stayers and complex movers; (b) Comparison between stayers and simple movers.

combine economic theory and previous empirical evidence in building up the model. Further, to keep the model as parsimonious as possible, I drop variables that are either a priori unrelated to $w_{0t} - w_{0t'}$ or are highly insignificant in the logit model, contributing nothing to its predictive power.

In the logit model, I include information about the respondent's labor market experience, demographic characteristics, and the key characteristics of the current job at the time of interview such as geographical area and hourly rate of pay. Those variables are expected to affect both of job change decisions and the growth of worker well-being at the workplace. For example, workers with longer tenure may be less inclined to leave the job and the closer tie to the job will accordingly affect the degree of happiness at the workplace.

Table 9 presents the resulting logit model using the above-mentioned selection rule. The correct prediction rates for both groups of movers are quite satisfactory. Specifically, I set an observation as 1 if the estimated probability is greater than 0.5, and 0 otherwise. The percentage of correctly specified values for the case of complex movers is 66.08; for the case of simple movers, it is 65.24. It is noted that marital status, education and experience are excluded from the model. The main reason is that their coefficients are highly insignificant and including them in the model does not improve the predictive power by a substantial amount. I suspect that the reason for the insignificance of experience is that its effect has been absorbed by age and tenure, i.e., experience prior to 1982 does not vary much in this sample of individuals who are continuously employed between 1982 and 1988 once age and tenure are controlled for. As for the education, since most of the workers in the sample have relatively low levels of education, the differences between mover and stayers are only marginal and do not contribute much to explaining mobility.

Table 9: Logit Model for Switching Jobs

	Complex change		Simple change	
Variable	Coefficient	AME [a]	Coefficient	AME
Age	-0.1839	-0.0370**	-0.1760	-0.0367**
	(0.0446)	(0.0086)	(0.0474)	(0.0095)
Tenure at current job	-0.2992	-0.0602**	-0.2836	-0.0590**
	(0.0605)	(0.0115)	(0.0657)	(0.0130)
Male	0.5326	0.1079**	0.0692	0.0144
	(0.1767)	(0.0354)	(0.1803)	(0.0375)
Black	-0.4276	-0.0861*	-0.4978	-0.1022*
	(0.1925)	(0.0384)	(0.2098)	(0.0419)
Hispanic	-0.5325	-0.1064*	-0.9586	-0.1885**
	(0.2465)	(0.0481)	(0.2664)	(0.0471)
Living in urban area	0.0830	0.0167	0.3305	0.0683
	(0.2198)	(0.0442)	(0.2254)	(0.0459)
Living in the Northeast	0.5195	0.1038*	0.3191	0.0665
	(0.2477)	(0.0483)	(0.2607)	(0.0540)
Living in the South	0.2639	0.0528	0.2485	-0.0517
	(0.2199)	(0.0435)	(0.2297)	(0.0475)
Living in the West	0.6379	0.1271	0.6614	0.1375*
	(0.2841)	(0.0549)	(0.3151)	(0.0637)
Highly satisfied with current job	-0.4970	-0.1014*	-0.5211	-0.1084*
	(0.2005)	(0.0411)	(0.2123)	(0.0433)
Log hourly wage	-0.5884	-0.1184*	-0.5569	-0.1160**
	(0.2336)	(0.0463)	(0.2062)	(0.0.0422)
Employment-based health insurance	-0.4557	-0.0930*	-0.5264	-0.1118*
	(0.2186)	(0.0447)	(0.2287)	(0.0487)
Learning general skills at current job	0.1207	0.0243	0.2468	0.0514
	(0.1779)	(0.0357)	(0.1868)	(0.0388)
Propensity of taking another type of job	0.3678	0.0749†	0.2058	0.0429
	(0.2018)	(0.0.0412)	(0.2083)	(0.0435)
Work in managerial professions	0.8522	0.1679**	1.0191	0.2109**
	(0.2602)	(0.0485)	(0.2889)	(0.0563)
Intercept	5.3052	–	5.0989	–
	(1.8062)	–	(1.0417)	–
Pseudo R^2	0.1513		0.1268	

Notes: Robust standard errors are given in parentheses. Significance levels: † : 10% * : 5% ** : 1% .
[a] AME denotes the average marginal effect on $Prob(mover = 1)$.

Most of the coefficient estimates in the model are in the expected direction, given the differences observed in Table 4. For example, younger workers are more likely to switch jobs. One additional year of tenure at the current job decreases the probability of leaving by 0.06. The tendency to leave drops by almost 0.12 when hourly rate of pay increases by 1%. Workers in managerial professions have a higher level of mobility than the other professions, perhaps because the managerial skills are general and thus transferable across employers. Related is a variable directly identifying general skills at work.[10] Workers acquiring firm-specific skills should be less likely to leave, which affects the way they form the expectation for a job, the key determinant of well-being at the workplace. Therefore, although the coefficient for this dummy variable is insignificant, I keep it in the model given its conceptual importance.

Employment-based health insurance and job mobility has been linked through a phenomenon called "job-lock". Workers are "locked" into their jobs because preexisting conditions exclusions make it expensive for individuals with medical problems to relinquish their current health insurance (Madrian, 1994). The coefficients of insurance in both samples are significant at 5% level; having an employment-based health insurance reduces the probability of voluntary turnover by approximately 10 percentage points. On the other hand, being insured or not will surely make a difference in terms of worker well-being over time.

The dummy variable *highly satisfied with current job* is also an important predictor of voluntary turnover. In one sense, overall satisfaction captures the intention to stay or quit; in another, it doubles as an indicator for the unobservable alternative job opportunities given that those with good opportunities are less satisfied than those with poor opportunities

[10] Responses to the question "learning general skills at work" and coded 1 if the answer is "highly agree".

(Freeman, 1978). For both samples, the effects are significant and the probability of leaving current job is about 0.10 lower for those who are highly satisfied. Lastly, the dummy variable *propensity of taking another type of job* is measured by responses to the question: “if you were free to go into any type of job you wanted, what would you do? would you take another job or keep the same job as you have now?” The coefficient for this variable is significant for the sample of complex movers at 10% level, indicating that workers taking complex job changes later on are generally highly self-motivated to try other types of job in 1982.

2.6.3 The Difference in Differences Matching Estimates

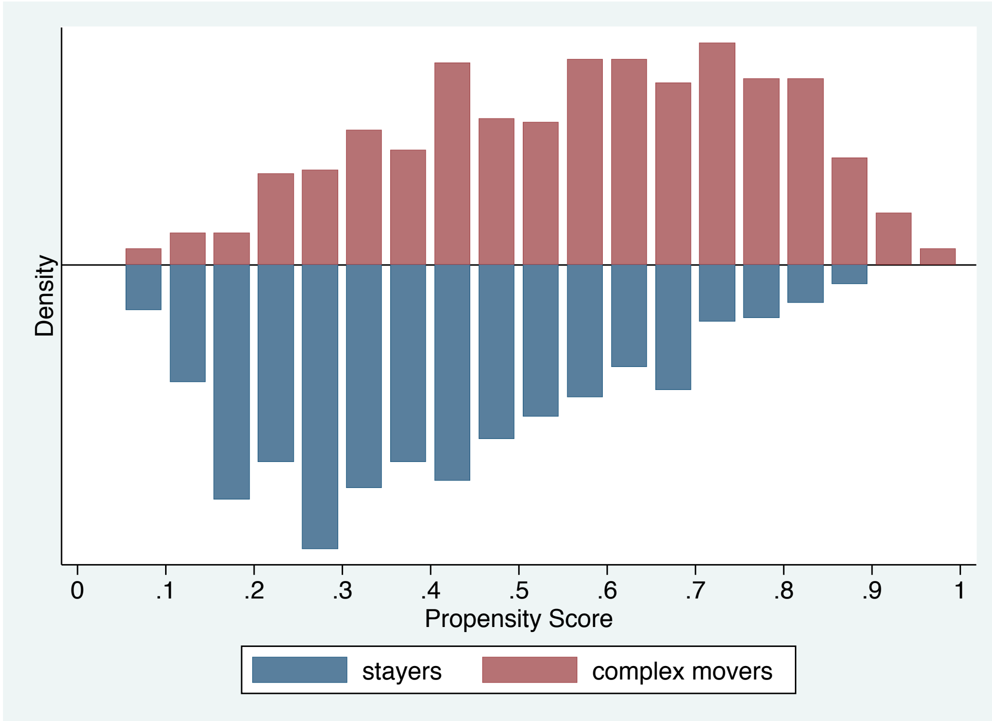
Following the literature of matching, I define the probability of treatment given X as propensity score:

$$e(X) = Pr(D = 1|X).$$

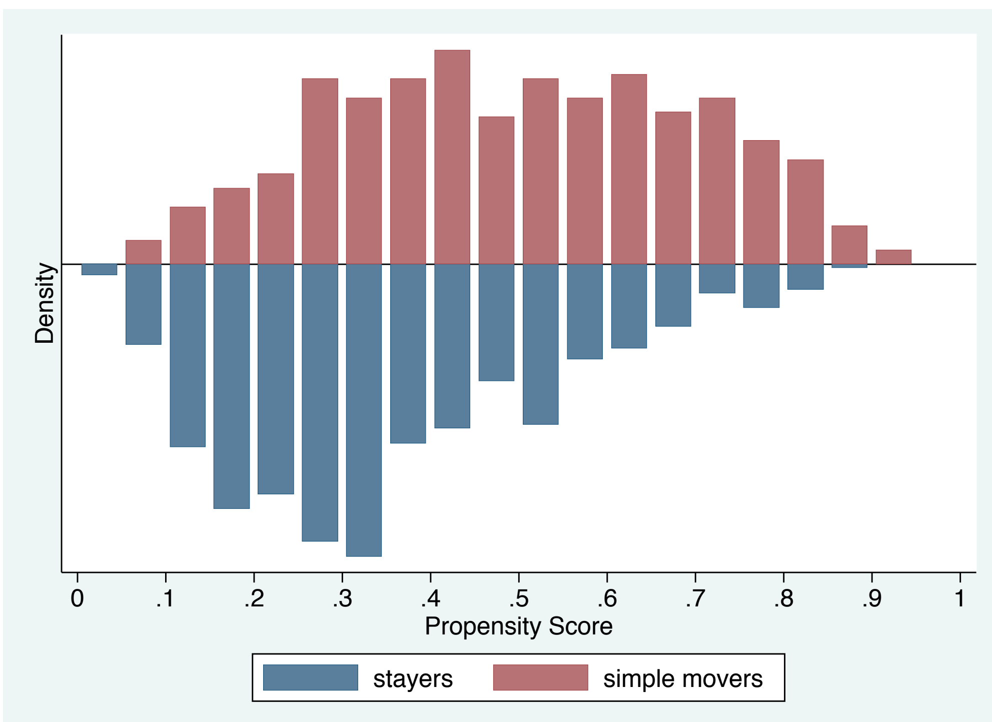
Figure 2 presents a simple graphical assessment of the extent of any support problem by plotting histograms of the propensity score for both groups of movers. It is apparent that mass of the density for the group of stayers lies well to the left of that of both groups of movers, indicating that most stayers have low predicted probabilities of turnover. Since there are few observations of opposite treatment at the right end of the distribution for movers, a subsample with overlap needs to be selected. I apply the following procedure to determine the subsample with overlap:

$$\mathbb{A} = \{x \in \mathbb{X} \mid \alpha \leq e(X) \leq 1 - \alpha\}$$

In principle, the cutoff value α is determined on the basis of the joint distribution of the treatment indicators and the covariates. However, calculations for Beta distributions for the



(a) Comparison between Stayers and Complex Movers



(b) Comparison between Stayers and Simple Movers

Figure 2: Histograms of the Estimated Probability of Changing Jobs: (a) comparison between stayers and complex movers; (b) comparison between stayers and simple movers.

Table 10: Test of Match Quality: Pre-treatment Characteristics–Complex change

	Movers	Matched controls	All stayers
Age	21.4812	22.1309	22.5532
Male	0.5519	0.5329	0.4563
Black	0.0849	0.0807	0.1141
Hispanic	0.0564	0.0459	0.0609
Living in urban area	0.7948	0.8037	0.7900
Living in the Northeast	0.2524	0.2618	0.2075
Living in the South	0.3548	0.3558	0.3686
Living in the West	0.1437	0.1247	0.1010
Highly satisfied with current job	0.3771	0.3901	0.4949
Log hourly wage	1.6125	1.7225	1.7632
Tenure at current job	1.5676	1.8789	2.3059
Employment-based health insurance	0.7093	0.8092	0.8550
Learning general skills at current job	0.4813	0.4719	0.4867
Propensity of taking another type of job	0.7033	0.6843	0.5947
Work in managerial professions	0.1211	0.0938	0.0705

propensity score suggest that $\alpha = 0.1$ approximates the optimal set well in practice (Imbens, 2007).

The next step is to apply the covariate matching algorithm described in the previous section to find the best match for each mover on the common support. To assess the quality of matching, I compare the before- and after-matching characteristics at the baseline for complex movers and simple movers and present the results separately in Table 10 and Table 11. The differences between both types of movers and their matched controls are much smaller compared to the differences between movers and all stayers. Those matched controls are then effectively used for the estimation of ATT.

Table 12 presents the DID matching estimates of ATT. The first row contains the baseline matching estimates. The ATT for complex movers is estimated to be equal to 0.0572, a little lower than the direct DID estimates of 0.0726. Further, since *propensity of taking another type of job* and *highly satisfied with current job* directly captures the intention of turnover and

Table 11: Test of Match Quality: Pre-treatment Characteristics–Simple change

	Movers	Matched controls	All stayers
Age	21.6063	22.1636	22.5532
Male	0.4554	0.4400	0.4563
Black	0.0778	0.0826	0.1141
Hispanic	0.0403	0.0378	0.0609
Living in urban area	0.8107	0.8863	0.7900
Living in the Northeast	0.2214	0.2422	0.2075
Living in the South	0.3743	0.3423	0.3686
Living in the West	0.1350	0.1251	0.1010
Highly satisfied with current job	0.4225	0.4564	0.4949
Log hourly wage	1.5869	1.7113	1.7632
Tenure at current job	1.5767	1.8942	2.3059
Employment-based health insurance	0.7358	0.7658	0.8550
Learning general skills at current job	0.5024	0.4969	0.4867
Propensity of taking another type of job	0.6732	0.6434	0.5947
Work in managerial professions	0.1451	0.1444	0.0705

are also closely tied to the outcome $w_{0t} - w_{0t'}$, I overweigh these two variables in matching and the resulting estimated ATT increases to 0.0629 as is shown in row (2). The third specification adds the bias-correction procedure into matching and the variables used for correction are the same as conditional covariates. The ATT estimate increases to 0.0673 as a result. Finally, the excluded variable marital status is an important mediating factor when workers report their job satisfaction and changes in marital status are suspected to affect changes in worker well-being. To disentangle its effect from that of job mobility, I include changes in marital status in the bias-correction step in the fourth specification. Row (4) shows that the estimate is 0.0673, with a standard error of 0.0102. According to those results, worker well-being is indeed boosted through complex job mobility in the sense that those movers are significantly happier at the workplace than they otherwise would be if they stayed with their previous employers. However, the positive effect of job mobility is insignificant and much smaller for movers taking simple job changes. Row (4) for ATT for simple change shows that the effect is only 0.0136 and insignificant. This is in contrast with the fact that movers with simple change actually experienced higher wage growth according to Panel 2 of Table 8, implying that the source of increase in well-being is mainly non-pecuniary aspects of work. Moreover, when I modify the DID results for hourly rate of pay using the third specification of matching, it turns out that the wage gains for complex job changers are insignificant, reinforcing the role of non-pecuniary rewards in understanding the behavior of turnover in the labor market. Finally, discarding observations outside the common support raises the concern over whether the remaining workers can be viewed as representative. Therefore, I repeat the matching using all observations to test the robustness of ATT estimates. Inspection of row 4 and row 5 shows that the proportion of discarded observations is relatively small so it poses few problems (Bryson et al., 2002). It also shows that excluding observations outside the common support is important given the relatively large variations in ATT estimates between row 4 and row 5.

Since it is argued that worker well-being should only measure satisfaction regarding non-monetary aspects of the job, I repeat the analysis excluding satisfaction over pay. The results are consistent with those in Table 12 and there are only minor changes in the estimates. Taking specification 4 as an example, the DID result for complex job changes is 0.0642 and significant, while that for simple job changes is 0.0214 and insignificant. As a comparison, the DID results in terms of hourly rate of pay for both simple and complex job changers are insignificant, showing that the well-being increase for complex job changers originates from non-pecuniary aspects of the job match.

Also, it is interesting to compare the results in Table 12 with the simple DID results shown in Table 8. In the case of worker well-being, the simple DID result for complex

Table 12: DID Matching Estimates of ATT

Specification	Complex Change			Simple Change		
	ATT	Treated	Controls	ATT	Treated	Controls
1. *Worker wellbeing*						
(1)	0.0572**	552	288	0.0265*	463	260
	(0.0101)			(0.0118)		
(2)	0.0629**	552	285	0.0169	463	269
	(0.0102)			(0.0115)		
(3)	0.0671**	552	285	0.0134	463	269
	(0.0102)			(0.0115)		
(4)	0.0673**	552	285	0.0136	463	269
	(0.0102)			(0.0115)		
(5)	0.0576**	573	294	0.0173	471	272
	(0.0106)			(0.0115)		
2. *Hourly rate of pay*						
(3)	0.0384	571	292	0.3674*	453	261
	(0.1750)			(0.1852)		

Notes: Specification (1): base specification, matching on specified variables only; specification (2): with exact matching on *propensity of taking another type of job* and *highly satisfied with current job*; specification(3): with exact matching and bias-correction; specification(4): with exact matching, bias-correction and additional correction for *married*; specification (5): test robustness of specification(4) when observations outside the common support are included.

changes is 0.0726, while the matching DID ranges from 0.05 to 0.07 in Table 12. In Table 8, the DID for simple changes is 0.0382 while the matching DID ranges from 0.01 to 0.03 in Table 12. Matching adjusts the effects of job mobility downward a little bit for both types of job change, but with only a small amount. On the other hand, matching does make a bigger difference in the case of the wage change. For complex job changes, it falls from 0.4650 to 0.0384. Form simple changes it falls from 0.8970 to 0.3674. Further research needs to be done in order to explain why the matching does not make big differences in terms of worker well-being, especially when matching is a theoretically important step towards estimating the true effect of job mobility.

Last but not least, all the results in Table 8 and Table 12 are essentially DID estimates from two-by-two cases. The main problem of the case is that the standard error of the estimate is subject to the criticism raised by Donald and Lang (2007). Specifically, the reported variance of the DID estimates from a two-by-two case includes only the sampling variance but not the part of the variance due to the common error. So if there are shocks that are correlated within year/group cells, the resulting t-statistic will be too high. Therefore, when interpreting the significance of the estimates, some caution should be exerted.

2.7 Conclusion

The study has attempted to identify the effect of voluntary job mobility on worker well-being, using a measure of worker well-being that encapsulates both pecuniary and non-pecuniary aspects of work. The results show that voluntary turnover increases well-being at the workplace for movers who are in the early stage of their career and conduct complex job changes involving different types of job.

The empirical results in this paper also provide insights into the process of matching for young workers. Individuals face different career options at the outset of their professional life, but they will gradually sort themselves into the types of job that are in accord with their abilities, values and interests through job mobility. Voluntary turnover is not only triggered by financial motives. An employee may, for example, prefer a pleasant work to decent pay, and become anxious to switch when dissatisfied with peer relationships at the current workplace even if no wage gains are expected as a result of the job change. This explains why complex movers were happier after a period of career matching even if they would have realized a wage growth of the same magnitude if they had chosen to stay.

CHAPTER III

THE WAGE EFFECTS OF EDUCATIONAL MISMATCH IN THE COLLEGE GRADUATE LABOR MARKET

3.1 Introduction

As the investment on higher education has risen dramatically over the past several decades, educational mismatch becomes a growing concern for policy makers. Educational mismatch is a complex and dynamic phenomenon occurring at different dimensions. Graduates who possess more schooling than their job requires are overeducated, while those with less schooling than required are undereducated. Both over- and under-education refer to the differences between attained and required level of schooling, and therefore can be generalized as vertical mismatch. Another dimension of mismatch, which is linked to the heterogeneity of skills generated by different types of education among individuals having the same educational level, is largely ignored in the literature. As noted by Sloane (2003), workers may be mismatched if the level of schooling is appropriate but the type of schooling is not, which can be denoted as horizontal mismatch.

Skill heterogeneity and job match theory cast some light on the existence and wage implications of horizontal mismatch. On the demand side of the labor market, jobs differ in different skill requirements; on the supply side, job seekers differ in skills that were acquired by means of education and training. Job match theory claims that the quality of match, i.e., the alignment between required and acquired skills, determines the productivity and thus wages at the job. The higher the match quality, the higher the realized productivity and earnings. Hence the economy faces an allocation problem. Optical allocation results if the individual worker can not increase his productivity by switching to another job. In practice, however, the optical allocation is rarely realized due to market failure. Students entering higher education choose a field of specialization such as mathematics or economics. But after graduation, they are recruited for a job with skill requirements that are not accentuated in their field of study. Therefore, they can not reach their optimal levels of productivity and wage. From this perspective, the wage prospects of those horizontally mismatched graduates are dampened relative to their well-matched peers.

On the other hand, the disparity between required skills at work and acquired skills from school could be treated as a special type of occupational mobility, especially among the highly-educated. Choosing a college major is no doubt a choice under uncertainty, even if students may have a significant amount of knowledge about occupations related to their selected majors. Intertemporal-utility-maximizing individuals in different fields will keep adjusting their intended career path through self-selection into dissimilar jobs, including those different from their initial choices when they are in college. Graduates entering jobs requiring different sets of skills from what they have learnt at school will incur wage loss because some of their obtained human capital can not be used in the new occupation.

Under this background, this paper focuses on the horizontal aspect of mismatch and examines the wage impacts of this type of mismatch at the individual level. However, it is important to note that examination of mismatch effects may also be extended to the macroeconomic level because macro-level imbalance between skill needs and skill supply may have far-reaching adverse effects on economic growth and innovation capacity.

This paper proceeds as follows. Section 2 begins with a discussion of the existing research on educational mismatch. Section 3 describes data, variable definitions and descriptive statistics. Section 4 then lays out the empirical strategy and presents results. Section 5 concludes and discusses some possible extensions for future research.

3.2 Existing Research on Educational Mismatch

The notion of horizontal mismatch is closely related to the vast literature of vertical mismatch. Ever since Duncan and Hoffman (1981) started the literature of mismatch in terms of over- and under-education, as well as wage effects of this type of mismatch, vertical mismatch has been extensively investigated (e.g., Hersch, 1991; Dolton and Vignoles, 2000; Bauer, 2002; McGuinness and Bennett, 2007; Dolton and Silles, 2008). Hartog (2000) provides a comprehensive overview of the questions addressed in the literature of over-education and earnings.

However, the issue of horizontal mismatch remains largely unexplored, although recent contributions are growing. For instance, the fact of horizontal mismatch has been documented in several studies (e.g. Farber, 1999). Using the 1993 wave of National Survey of College Graduates (NSCG), Robst (2007) focuses on the incidence and effects of horizontal mismatch and examines which degree fields lead to greater mismatch and mismatch effects. Bender and Heywood (2006) use the Survey of Doctoral Recipients to estimate the magnitude and consequences of horizontal job mismatch among Ph.D. graduates in science. Their study shows that mismatch is associated with substantially lower earnings, lower job satisfaction and a higher rate of turnover for both academics and nonacademics.

Several studies combine both horizontal and vertical mismatch in the discussion of wage effects. For example, Allen (2001) shows that when over-education, under-education and horizontal mismatch are all taken into consideration, horizontal mismatch has no significant effects on wage, job satisfaction and on-the-job search. In contrast, Kucel (2010) uses seemingly unrelated equations to show that when both dimensions of mismatch are built into the model, in most cases being horizontally matched has a wage premium and being over-educated does not affect wage. Robst (2008) also expands the concept of educational mismatch to incorporate both the quantity and type of schooling and concludes that the wage effects of over-education are larger when workers are also mismatched based on degree field.

3.3 Data, Variable Definitions and Descriptive Statistics

This section provides details about data, variable definitions and descriptive statistics. The data are from the 2003 wave of the National Survey of College Graduates (NSCG 2003), which consists of a sample of 100,402 US residents who hold at least a bachelor's degree (completed by April 1, 2000), who lived in the US or its territories in both 2000 and 2003, and who were age 76 and under as of October 1, 2003. NSCG 2003 captures rich information about the occupational and educational histories of these college graduates. The original raw data are available from the Scientist and Engineers Statistical Data System (SESTAT) website. Given my focus on the educational mismatch, I restrict the sample to those who have responded to the main question about the relatedness of the highest degree and work, who account for 83,024 of the original sample.

To select a more homogeneous sample, I exclude respondents from the sample if (1) they obtained the first bachelor's degree outside the US, (2) they worked 10 or fewer hours per week, (3) the computed hourly wage was less than \$6 or greater than \$300. After imposition of these sample selection rules, the final sample size is 67,746.

The reported wages are measured by basic annual salary on the principal job as of the week of October 1, 2003, excluding bonuses, overtime or additional compensation.[1] The salary variable used in the study is defined as the log annual wage.

The measure of educational mismatch is based on respondent's self-evaluation. Specifically, during the survey respondent was asked to what extent his or her highest degree was related to the principal job. The answers were framed on three scales, from *closely related* and *somewhat related* to *not related.* Respondents answering "not related" were further asked to specify the most important reasons for working in an area outside the field of highest degree. I further create a categorical variable summarizing reasons for working outside the highest degree. The seven categories are: (1) pay, promotion opportunities, (2) working condition, (3) job location, (4) change in career or professional interests, (5) family-related reasons, (6) job in highest degree field not available, and (7) other reasons, respectively.

For each respondent I observe highest degree type attained, grouped by bachelor, master, doctorate or professional degree. Based on field of study for the highest degree, I create a dummy variable to distinguish science and engineering (SE) fields with other specializations.[2] I also consider informal human capital acquired through on/off-the-job training. This information is captured by the survey question, "During the past year, did you take any

[1] Self-employed or other non-salaried respondents reported their estimated earned income, excluding business expenses. NSCG 2003 instructed respondents to do above-mentioned deductions before reporting their annual salary. However, it would be interesting to see how the wage effects will change if information on those additional parts of compensation is available.

[2] Roughly speaking, SE fields include computer and mathematical sciences, life and related sciences, physical and related sciences, social and related sciences, engineering, and other SE-related fields.

work-related training, such as workshops or seminars?". Another dummy variable is created to indicate whether or not the respondent received any work-related training in 2002.

The major work activity on the principal job considers work activity the respondent spent most hours during a typical week. There are 14 categories of work activities in this study: (1) accounting, finance and contracts, (2) basic research, (3) applied research, (4) development, (5) design of equipment, processes, structures, models, (6) computer applications, programming, system development, (7) employee relations, (8) managing or supervising people or projects, (9) production, operations, maintenance, (10) professional services, (11) sales, purchasing, marketing, customer service, public relations, (12) quality or productivity management, (13) teaching, (14) other activities.

I also take into account employer type of the principal job. I use six categories to characterize employer type. I group private-for-profit and private-for-non-profit non-educational employers with more than 1,000 employees into the first category *large private*, private-for-profit and private-for-non-profit non-educational employers with fewer than 1,000 employees into the second category *small private*, all education institutions into the third category *school*, local, state, federal government and military into the fourth category *government*, incorporated and non-incorporated self-employed into the fifth category *self-employed*, and the rest into the sixth category *others*. In addition, I define a dummy variable to indicate whether the respondent supervised the work of others as part of the principal job.

Table 13: NSCG Summary Statistics by Degrees of Educational Mismatch

	All	Closely related	Somewhat related	Not related
a. Salary				
Salary ($000)	67.1	71.8	67.4	52.6

Continued on next page

Table 13 – *Continued from previous page*

	All	Closely Related	Somewhat related	Not related
b. Human capital variables				
Tenure	8.58	9.28	7.84	7.52
Highest degree type (excluded: bachelor, %)				
Master	26.3	32.1	22.1	14.7
Doctorate	3.28	4.56	1.99	1.18
Professional	6.72	10.93	1.52	1.14
Field of study for highest degree (%)				
SE	38.9	40.1	37.9	36.5
Attend Training (%)	65.4	72.9	60.5	49.4
c. Demographic variables				
Age (years)	44.2	44.0	44.1	45.0
Female (%)	45.8	49.3	39.4	43.9
White (%)	83.6	83.1	84.6	83.8
Asian (%)	3.06	3.10	3.16	2.82
Black (%)	6.63	6.53	6.45	7.17
Hispanic (%)	4.78	5.40	3.79	4.27
Non-citizen (%)	0.862	0.909	0.836	0.756
Never married (%)	14.0	12.9	14.1	16.9
d. Job characteristics				
Major work activity (excluded: other, %)				
Accounting	9.38	8.88	10.03	9.97
Basic research	1.33	1.528	1.228	0.893
Applied research	2.56	2.89	2.77	1.28
Development	2.41	2.64	2.46	1.69
Design	5.72	4.41	7.31	7.48
Computer	2.48	2.56	2.77	1.28
Employee relations	3.05	2.43	3.93	3.73
Managing projects	1.93	1.27	2.60	3.00
Production	3.47	1.80	3.69	8.13
Service	17.3	24.28	9.02	7.56
Sales	13.6	6.44	20.80	25.17
Management	17.2	14.8	22.5	17.0
Teaching	15.8	23.54	7.11	4.20
Employer type (excluded: other, %)				
Large private	23.7	19.5	30.3	27.3
Small private	25.5	22.5	29.0	29.7
School	22.5	32.40	11.16	8.13
Government	11.5	10.8	13.0	11.6
Self-employed	16.7	14.7	16.3	23.1
Supervise others (%)	47.8	49.8	50.9	38.0

Table 13 reports the summary statistics both for the sample overall and separately for the three subgroups defined by the relatedness of highest degree and principal job. In this and all of the following analysis, observations are weighted to create a nationally representative sample. The top panel of Table 13 reports the average annual salary for each group, the second panel reports summary statistics of human capital variables, the third panel demographic variables and the bottom panel job characteristics variables such as major job activity and employer type. Notice that average observed annual salary decreases gradually from closely-related group to not-related group. In some instances this could speak to composition effect instead of negative wage effects of educational mismatch. For example, the closely-related group composes of a higher proportion of graduates with advanced degree, especially professional degree which is always rewarded with higher earnings. The mismatched group also has a over-representation of African American so that black-white pay gap might play a role in the wage differences between well-matched and mismatched groups. Also, compared to the other two groups, the mismatched group has a lower percentage of graduates who supervise others at work. If supervising others signals higher job rank, then promotion and accompanying wage raise might explain at least a portion of the wage differences. Therefore, it is likely that the observed wage variations across the groups are due to heterogeneity in observed characteristics or even variation in unobserved factors. The following analysis thus aims to separate these sources of variation and establish the link between educational mismatch and wage.

3.4 Methodology and Results

3.4.1 Baseline Regression

To assess the wage effect of educational mismatch among the highly educated, I begin with a modified version of the standard Mincer equation which has been widely adopted within the educational mismatch literature:

$$lnW = \beta_0 + \mathbf{MM}\boldsymbol{\beta_1} + \mathbf{HC}\boldsymbol{\beta_2} + \mathbf{JC}\boldsymbol{\beta_3} + \mathbf{X}\boldsymbol{\beta_4} + \mu \tag{29}$$

where MM denotes different levels of educational mismatch, HC human capital variables, JC job characteristics variables and X family background and personal characteristics. I first run the regression with only MM and HC, the traditional Mincer equation specification, then add JC, followed by X_i. The focus is on the mismatch coefficients, $\boldsymbol{\beta_1}$, which reflect the level difference in wage between mismatched groups and matched group, the excluded category.

From educational mismatch, I construct two dummy variables, *somewhat related* and *not related*, to measure the extent of educational mismatch. These two variables are included in MM. The human capital variables included in HC are highest degree type, the SE dummy variable for the field of study of the highest degree, dummy variable for attending work-related training in the past year, age, age square, job tenure and tenure square. Work-related training captures the part of human capital acquired after starting work. Work experience measures human capital acquired on the job, while job tenure measures firm specific human capital. Since work experience is not observed in the data, I use age as a proxy for work experience. JC include variables related to the job, such as major job activity, employer type and a dummy variable that equals one if the respondent supervised others as part of the work. Finally, I include controls that are likely to pick up elements of personal characteristics and family background. These include gender, race and ethnicity, citizenship, marital status, physical disability indicator, a dummy variable that equals to one if father has a bachelor's degree or higher, a dummy variable that equals to one if mother has a bachelor's degree or higher and a dummy variable that equals to one if the respondent attended a private institution for the first bachelor's degree. The latter three are proxies for family background because there is no detailed measure of family background in the data set.[3]

Consider the importance of adding controls JC and X to the traditional Mincer specification by assuming that the true causal effect of educational mismatch on wage is zero. If professions with lower requirements on matching educational background tend to pay less than those with higher requirements, then the coefficient of mismatch will be negative if professions are not controlled for. On the other hand, personal characteristics and family background will pick up unobserved elements of taste or constraints that may affect both education mismatch and income prospect, so failing to include X will also drive the coefficient away from zero.

Table 14 reports the differences in wage effect of mismatch by degrees of mismatch. Line (1) reports the result of standard Mincer specification, Line (2) the results after including only JC, and Line (3) the fully-controlled specification with both JC and X. The last column shows that including more controls appreciably raises the adjusted-R^2. For all specifications,

[3]Ideally, I would like to include measures of parental income and intelligence, but such measures are unavailable in NSCG 2003.

Table 14: Differences in the Wage Effects of Mismatch by Degrees of Mismatch

	Somewhat related	Not related	Adjusted R^2
(1) Mincer Specification	0.0446 $**$	-0.2199 $**$	0.1498
	(0.0085)	(0.0104)	
(2) + JC	-0.0659 $**$	-0.2800 $**$	0.2934
	(0.0081)	(0.0102)	
(3) + X	-0.0738 $**$	-0.2690 $**$	0.3405
	(0.0077)	(0.0098)	

Notes: Each line reports the coefficients associated with two dummy variables of educational mismatch, with matched group as the excluded category. Heteroskedastic-consistent standard errors are in the parenthesis. Observations are weighted to create a nationally representative sample. Line (1) reports the result of standard Mincer specification, Line (2) the results after including only JC, and Line (3) the fully-controlled specification with both JC and X. The last column reports the adjusted R^2. Significance levels: $\dagger$: 10% $*$: 5% $**$: 1% .

I find that the coefficients on *somewhat related* are less negative than that on *not related*, which indicates that wage declines, if any, are smaller for graduates when more skills transfer to the current job. In addition, controlling for JC dramatically changes the coefficient on *somewhat related* from positive to negative. Because people who are promoted along the career ladder are most likely to do the work that is not exactly but somewhat related to their highest degree and job rank is in part picked up by variables in JC, this helps to explain the changes in the coefficient. Finally, a comparison of Line (1), (2) and (3) shows that persistent and significant wage differences remain, even after controlling for these rich sets of job and personal characteristics.

Table 15: The Wage Effects of Educational Mismatch

	Estimate	Std. Error	Pr(>|t|)
Educational mismatch			
Somewhat related	-0.0738	0.0077	0.0000$**$
Not related	-0.2690	0.0098	0.0000$**$
Human capital			
Highest degree type (excluded: bachelor)			
Master	0.1508	0.0070	0.0000 $**$

Continued on next page

Table 15 – *Continued from previous page*

	Estimate	Std. Error	Pr(>|t|)
Doctorate	0.3188	0.0139	0.0000 ∗∗
Professional	0.5318	0.0149	0.0000 ∗∗
SE major	0.0633	0.0065	0.0000 ∗∗
Attend work-related training	0.0702	0.0071	0.0000 ∗∗
Age	0.0597	0.0025	0.0000 ∗∗
Age square	-0.0007	0.0000	0.0000 ∗∗
Tenure	0.0173	0.0011	0.0000 ∗∗
Tenure square	-0.0002	0.0000	0.0000 ∗∗
Personal background			
Female	-0.2897	0.0065	0.0000 ∗∗
Asian	0.0787	0.0122	0.0000 ∗∗
Black	-0.0278	0.0091	0.0023 ∗∗
Hispanic	-0.0702	0.0109	0.0000 ∗∗
Non US citizen	-0.0205	0.0245	0.4014
Never married	-0.0626	0.0091	0.0000 ∗∗
Father has bachelor's degree	0.0414	0.0073	0.0000 ∗∗
Mother has bachelor's degree	0.0168	0.0080	0.0344 ∗
Attend private college	0.0300	0.0065	0.0000 ∗∗
Physical disability indicator	-0.0800	0.0120	0.0000 ∗∗
Job characteristics			
Major work activity (excluded: accounting and finance)			
Basic research	-0.1666	0.0247	0.0000 ∗∗
Applied research	-0.0377	0.0191	0.0487 ∗
Computer application	0.0054	0.0195	0.7814
Development	0.0072	0.0173	0.6765
Design	0.0062	0.0159	0.6953
Employee relation	0.0098	0.0189	0.6019
Management	0.0373	0.0134	0.0054 ∗∗
Production	-0.2476	0.0200	0.0000 ∗∗
Service	-0.1058	0.0143	0.0000 ∗∗
Sales	-0.0367	0.0153	0.0166 ∗
Quality management	-0.0645	0.0268	0.0162 ∗
Teaching	-0.2012	0.0153	0.0000 ∗∗
Other	-0.2495	0.0208	0.0000 ∗∗
Employer type (excluded: government)			
Large private	0.2176	0.0087	0.0000 ∗∗
School	-0.1094	0.0105	0.0000 ∗∗
Self-employed	0.0244	0.0115	0.0343 ∗
Small private	0.0404	0.0094	0.0000 ∗∗
Other	0.1822	0.0497	0.0002 ∗∗
Supervising others	0.2171	0.0066	0.0000 ∗∗

Continued on next page

Table 15 – *Continued from previous page*

	Estimate	Std. Error	Pr(>|t|)
Intercept	9.4381	0.0564	0.0000 **

Table 15 reports complete results from the full-controlled specification. While not the focus of this paper, I briefly discuss some of the other coefficients. Obtaining an advanced degree or attending work-related training boosts the earnings. Graduates with science and engineering degree earn more than those from other fields. Female graduates earned approximately 29% less than male graduates. Other things equal, respondents coming from a better family background have higher earnings on average than others while citizenship does not make any significant difference. Professions of *research, production, service, sales, quality management* and *teaching* have lower average earnings than the reference profession category *accounting and finance*, while people in *management* earn 4% higher than those in *accounting and finance*. I also find that both self-employed and private company employees earn more than government employees, while school on average pays less than the government sector.

3.4.2 Hypotheses about the Wage Effects of Mismatch

In this section, I use variations of model (29) to test three hypotheses related to the wage effects of educational mismatch. The first hypothesis concerns the impact persistency of educational mismatch. It is argued that investment in training, adapting jobs and gaining experience will eventually make up for the mismatch between highest degree and job. If this statement is true, then the negative impacts of mismatch reflect temporary frictions occurring at the first stage of careers and will gradually disappear when tenure or experience

Table 16: The Wage Effects of Mismatch by Tenure, Reason and Degree Type

	Estimate	Std. Error	Pr(>|t|)
Specification #1			
Mismatch by tenure			
Somewhat related	-0.0824	0.0109	0.0000 $**$
Somewhat related × Tenure	0.0010	0.0009	0.2773
Not related	-0.3012	0.0132	0.0000 $**$
Not related × Tenure	0.0041	0.0011	0.0003 $**$
Adjusted R^2		0.3408	
Specification #2			
Mismatch by reasons			
Pay, promotion opportunities	0.0001	0.0141	0.9963
Working conditions	-0.4222	0.0261	0.0000 $**$
Job location	-0.3582	0.0325	0.0000 $**$
Career change	-0.2006	0.0181	0.0000 $**$
Family-related reasons	-0.5137	0.0244	0.0000 $**$
Job in degree field not available	-0.3678	0.0198	0.0000 $**$
Other	-0.3398	0.0358	0.0000 $**$
Adjusted R^2		0.3473	
Specification #3			
Mismatch × degree type			
Bachelor	-0.2221	0.0104	0.0000 $**$
Master	-0.3042	0.0218	0.0000 $**$
Doctorate	-0.2611	0.0862	0.0024 $**$
Professional	-0.6207	0.0840	0.0000 $**$
Adjusted R^2		0.3339	

Notes: Heteroskedastic-consistent standard errors are used for statistical inferences. Observations are weighted to create a nationally representative sample. Significance levels: $\dagger$: 10% $*$: 5% $**$: 1% .

accumulates. On the other hand, along the career people will get promoted and oversee others doing that work or they will be assigned a job that utilizes their intelligence in other ways, so educational mismatch, especially the *somewhat related* category, is not as harmful later on in their career as it is right after graduation. To test the hypothesis, I add an interaction term $Mismatch \times Tenure$ to model (29). The top panel of Table 16 reports the coefficients on both mismatch dummies and the two interaction terms in the new regression. The results only partly support the hypothesis. It is true that the wage penalty for complete mismatch is lower later in the career because the coefficient on the interaction of *not related* and *tenure* is positive, but the coefficient on the other interaction term is not significant. Further, the negative effects of complete mismatch attenuate so slowly that eliminating the negative impacts requires approximately 73 years of tenure.[4]

Therefore, the impact of mismatch lasts far beyond the initial transition phase between education and work. It also implies that the mismatch problem can not be fully resolved through current adjustment mechanisms, which might to some extent justify specific actions from policymakers.

The second hypothesis concerns underlying reasons of mismatch. Human capital theory suggests that choosing a field of specialization is a choice under uncertainty because individuals do not have perfect foresight of their career prospects when they are at school. Therefore, they will keep adjusting their intended career path through self-selection into dissimilar jobs, including those different from their initial choices at school. This dynamic career adjustment

[4]Another way to measure the persistency of mismatch is to use the interaction between mismatch and experience instead of tenure. However, there is no way to infer years of experience from the data set. The conventional way of Age-6-Years of Schooling can not be applied here because years of schooling is also unavailable.

process explains why graduates enter jobs requiring different sets of skills from what they have learnt at school. But misalignment of workers with occupations could occur for various involuntary reasons as well. For example, oversupply of individuals with specific sets of skills might force some graduates to branch out into other occupations because the cost or even risk of waiting for a job is too high. Skill obsolescence, such as the declining value of human capital acquired at school, is another crucial reason for educational mismatch given the increasing changes in work and organizations. Individuals could also be mismatched due to family or other external constraints. Accordingly, it is natural to hypothesize that the incurred wage loss of involuntary mismatch is higher than that of voluntary mismatch.

To test the second hypothesis, I utilize information contained in the data set regarding the reasons for working outside the field of highest degree. Notice that from now on I include only one mismatch indicator *mismatch* that equals to one if the principal job is *not related* with highest degree. The major reason to combine the category *somewhat related* and *closely related* is that only respondents who claimed "not related" were asked to provide further information about the mismatch, such the reasons of mismatch. Also, psychological study shows that respondents tend to choose the middle category if they are indecisive, so complete mismatch is a more accurate measure of mismatch given its subjective feature. I define voluntary mismatch as those resulting from the self-motivated dynamic career adjustment process when uncertainty is gradually resolved. According to this standard, I assign *pay, promotion opportunities* and *career change* into the voluntary category, *family-related reasons* and *job in highest degree field not available* into the involuntary category, respectively. But it is not clear how to distinguish between voluntary and involuntary mismatch for *working condition* and *job location*, so I focus on the four reasons that can be classified. I modify model (29) by splitting *mismatch* into different categories based on the reported reasons of mismatch.

Results from the second panel of Table 16 show that the coefficient on *career change* is significantly lower in absolute value than that of *family-related reasons* and *job in highest degree not available.* The coefficient on *pay, promotion opportunities* is positive but insignificant, indicating that there is no significant wage difference between matched graduates and those who choose to work outside the highest degree for better pay or promotion opportunities.[5] Overall, the result is consistent with the second hypothesis that the negative impacts are smaller for voluntary mismatch.

Finally, I test the hypothesis that the wage penalty is greater among graduates with advanced degree because they have substantial occupation specific skills that are not easily transferable to other occupations. There are various factors that may affect the magnitude of wage loss, among which is the specificity of human capital. General skills are by definition transferable, while occupation specific skills differ by degrees of transferability. The bottom panel of Table 16 shows that the well-matched graduates earn 22% more than their mismatched counterparts among bachelors, compared to 62.7% among graduates with professional degrees such as MBA or JD. This result justifies the hypothesis.

3.4.3 The Quantile Process of Mismatch Impacts

To further examine the impact of mismatch on earnings distribution besides conditional mean, I use quantile regression in this section to estimate the entire QR process of mismatch

[5]It is possible that some of its effects are captured by the dummy variable *supervising others* and the work activity dummy of *management.* For a robustness check, I omit these two variables and run the regression again. It turns out that the coefficient on *pay and promotion opportunities* is still insignificant.

Table 17: The Quantile Process of Mismatch Impacts

Quantile	Estimate	Std. Error	Quantile	Estimate	Std. Error
0.10	-0.3031	0.0211	0.15	-0.2865	0.0194
0.20	-0.2843	0.0134	0.25	-0.2841	0.0126
0.30	-0.2700	0.0127	0.35	-0.2607	0.0112
0.40	-0.2480	0.0103	0.45	-0.2427	0.0121
0.50	-0.2332	0.0116	0.55	-0.2177	0.0102
0.60	-0.2074	0.0104	0.65	-0.1957	0.0101
0.70	-0.1796	0.0108	0.75	-0.1697	0.0106
0.80	-0.1571	0.0111	0.85	-0.1424	0.0125
0.90	-0.1351	0.0127			

Notes: Observations are weighted to create a nationally representative sample.

effect. Consider the following linear conditional quantile function:

$$Q_\tau(lnW|X) = \mathbf{X}\boldsymbol{\beta(\tau)} \tag{30}$$

where $Q_\tau(lnW|X)$ denotes the τ-quantile of lnW given X. For ease of display, I use $\mathbf{X}$ to denote all the variables included in the model, including a mismatch dummy and other controls in model (29). The coefficients, $\boldsymbol{\beta}$, are allowed to depend on τ. The quantile regression process $\beta(\tau)$ are defined as:

$$\hat{\beta}(\tau) = argmin_{b\in\mathbb{R}^d} E[\rho_\tau(lnW_i - \mathbf{X_i}'\mathbf{b})] \tag{31}$$

where $\rho_\tau(u) = u(\tau - I(u < 0))$.

The entire QR process is of interest here because I would like to know how the wage effects vary across different quantiles of wage distribution. Under certain conditions, the entire QR process has joint asymptotic normality and the tests are formulated like Kolmogorov-type statistics, which have a well-behaved limit distribution. However, the limit distribution of the QR process will generally be affected by misspecification, which jeopardizes the "distribution-free" feature for confidence regions and tests, making the inference of the QR process difficult in empirical work. Here I adopt a subsampling procedure introduced by Chernozhukov and Fernández-Val (2005) and Angrist, Chernozhukov and Fernández-Val (2006) to obtain a consistent estimate of the critical value, which is not distribution-free under misspecification.

Table 17 reports the coefficients on mismatch from quantile regressions at different points. The 0.5 quantile coefficient of -0.2332, for the conditional median, is close to the OLS coefficient of -0.2366. Also noteworthy is the fact that there are a considerable amount of variation in the distributional impacts. At lower quantiles, the estimates suggest that educational mismatch is associated with 24-30% lower annual salary, while the wage penalty of mismatch drops straightly to 13.51% in the 0.90 quantile.

Figure 1 provides a graphical representation of the variation in mismatch penalty across quantiles. In particular, the figure shows the absolute value of the estimates of the mismatch coefficient quantile process, along with a robust simultaneous 95% confidence band. The horizontal line is the absolute value of the OLS point estimate and its confidence interval, which do not vary by quantile. Notice that the simultaneous confidence band does not contain a horizontal line, so I reject the hypothesis that the wage effect is constant across most of the distribution. In addition, the value of zero falls outside the confidence band, indicating that the wage effect is significant across this portion of distribution. Also, the wage penalty is quite large at lower quantiles and decreases sharply toward higher quantiles. Furthermore, if it is assumed that mismatch preserves an individual's rank in the wage distribution, i.e., a mismatched individual's earning would stay at the same quantile even if he or she were well-matched, then the above results lend some support to the hypothesis that low-income individuals face even more disadvantages if they are mismatched in the labor market, relatively to their high-income counterparts.

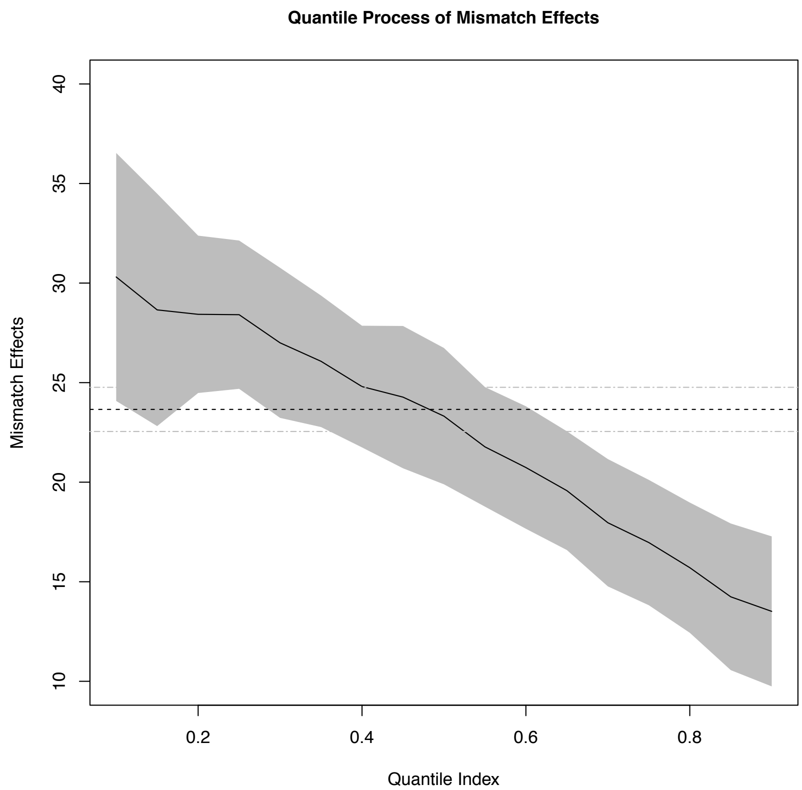


Figure 3: QR and OLS Coefficients and Confidence Regions on *Mismatch*

Ideally, if these graduates have homogeneous human capital profiles, then the relative position in the wage distribution will approximately reflect the unobservable individual characteristics such as ability. If this is the case, it may be reasonable to assume that mismatch is rank-preserving, and the above results suggest that the mismatch penalty is much higher among low-ability people. But NSCG 2003 is an economy-wide data set that contains large variations in human capital levels, so it is impossible to pinpoint the exact explanations of the variations in the wage effects of mismatch within current context.

3.5 Conclusion

The results in this paper provide insights into the wage impacts of horizontal mismatch. I find that the average wage loss associated with educational mismatch is significant and persistent. Graduates who are mismatched for involuntary reasons incur greater wage penalty compared to those for voluntary reasons. In addition, graduates with advanced degree suffer more from mismatch relative to those with only bachelor's degree. Lastly, there are considerable amounts of variations in the distributional impacts. The wage penalty is quite large at lower quantiles and decrease sharply towards higher quantiles. Given the magnitude and persistency of negative wage effects, efforts should be undertaken to facilitate early identification of skill needs and to provide a better mechanism of aligning skill needs and skill supply.

There are a number of issues that may be addressed in future research. The first issue concerns the measurement of horizontal mismatch. Objective measures of mismatch are useful to supplement the findings using subjective measures such as self-assessment. According to Groot and Maassen van den Brink (2000), measures based on self-reports result in higher incidence of mismatch than objective measures. There are at least two ways to tackle this issue. The first approach is to couple the data with systematic job and education analysis conducted by experts, and to parameterize the distance between skills obtained through education and that required in the job. The extent to which this method can be implemented depends on the availability and reliability of job and education analysis, which are always too expensive to carry out on a large scale. Freeman and Hirsch (2007) demonstrates a method of this type by linking a census of U.S. degrees and fields of study with measures of the knowledge content of jobs. The knowledge contents of jobs in their paper come from Occupational Information Network(O*NET), while the degree data are from the Digest of Education Statistics. The other way to measure horizontal mismatch is similar to the statistical measurement of vertical mismatch, only more complicated. Provided that skill contents of a job could be parameterized in the first stage, graduates are then considered to be mismatched if their skill contents are quite different compared to the mean or modal graduates in the same degree field. Notice that the statistical measurement is based on realized outcomes with some endogeneity concerns.

The second issue involves omitted variable bias. Because the data do not provide us any information on the respondent's ability or some other unobserved characteristics such as preference of non-pecuniary benefits above wage, it is possible that the estimated wage penalty is caused by those omitted variables instead of mismatch. A natural experiment or a good instrument for mismatch is the key to a causality interpretation.

REFERENCES

[1] A. Abadie, D. Drukker, J.L. Herr, and G.W. Imbens. Implementing matching estimators for average treatment effects in stata. *Stata journal*, 4:290–311, 2004.

[2] A. Abadie and G.W. Imbens. Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1):235–267, 2006.

[3] G.A. Akerlof, A.K. Rose, J.L. Yellen, L. Ball, and R.E. Hall. Job switching and job satisfaction in the us labor market. *Brookings Papers on Economic Activity*, 1988(2):495–594, 1988.

[4] J. Allen and E. De Weert. What do educational mismatches tell us about skill mismatches? a cross-country analysis. *European Journal of Education*, 42(1):59–73, 2007.

[5] J. Allen and R. Van der Velden. Educational mismatches versus skill mismatches: effects on wages, job satisfaction, and on-the-job search. *Oxford Economic Papers*, 53(3):434–452, 2001.

[6] J. Allen and R. van der Velden. The role of self-assessment in measuring skills. In *Transition in Youth Workshop, Valencia, Spain*, 2005.

[7] Joseph G. Altonji, Anthony Smith, and Ivan Vidangos. Modeling earnings dynamics. Working Paper 14743, National Bureau of Economic Research, February 2009.

[8] J. Angrist and V. Chernozhukov. Quantile regression under misspecification, with an application to the us wage structure. *Econometrica*, 74(2):539–563, 2006.

[9] J.D. Angrist and J.S. Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press, 2009.

[10] A.P. Bartel. Earnings growth on the job and between jobs. *Economic Inquiry*, 18(1):123–137, 1980.

[11] A.P. Bartel and G.J. Borjas. Wage growth and job turnover: An empirical analysis. *Studies in labor markets*, 1981.

[12] T.K. Bauer. Educational mismatch and wages: a panel analysis. *Economics of Education Review*, 21(3):221–229, 2002.

[13] Keith A. Bender and John S. Heywood. Educational mismatch among phds: Determinants and consequences. Working Paper 12693, National Bureau of Economic Research, November 2006.

[14] P.M. Bentler. Comparative fit indexes in structural models. *Psychological bulletin*, 107(3):238–246, 1990.

[15] G. Betti, A. DAgostino, and L. Neri. Educational mismatch of graduates: a multidimensional and fuzzy indicator. *Social Indicators Research*, 103(3):465–480, 2011.

[16] O.J. Blanchard and P. Diamond. Ranking, unemployment duration, and wages. *The Review of Economic Studies*, 61(3):417–434, 1994.

[17] R. Blundell and I. Preston. Consumption Inequality and Income Uncertainty. *Quarterly Journal of Economics*, 113(2):603–640, 1998.

[18] K.A. Bollen and K.A. Bollen. *Structural equations with latent variables*. Wiley New York, 1989.

[19] G.J. Borjas. Race, turnover, and male earnings. *Industrial Relations: A Journal of Economy and Society*, 23(1):73–89, 1984.

[20] G.J. Borjas and S. Rosen. Income prospects and job mobility of younger men. *Research in Labor Economics*, 3:159–181, 1980.

[21] T.A. Brown. *Confirmatory factor analysis for applied research*. The Guilford Press, 2006.

[22] M.W. Browne and R. Cudeck. Alternative ways of assessing model fit. *Testing structural equation models*, 154:136–162, 1993.

[23] A. Bryson, R. Dorsett, and S. Purdon. The use of propensity score matching in the evaluation of active labour market policies. *Working Paper No. 4, Department for Work and Pensions*, 2002.

[24] K. Burdett. A theory of employee job search and quit rates. *The American Economic Review*, 68(1):212–220, 1978.

[25] M. Caliendo and S. Kopeinig. Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1):31–72, 2008.

[26] D. Chan. The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal mean and covariance structures analysis (lmacs) and multiple indicator latent growth modeling (mlgm). *Organizational Research Methods*, 1(4):421–483, 1998.

[27] V. Chernozhukov and I. Fernández-Val. Subsampling inference on quantile regression processes. *Sankhya: The Indian Journal of Statistics*, 67(2):253–276, 2005.

[28] A.E. Clark. What really matters in a job? hedonic measurement using quit data. *Labour Economics*, 8(2):223–242, 2001.

[29] M.G. Coles and E. Smith. Marketplaces and matching. *International Economic Review*, 39(1):239–254, 1998.

[30] F. Cunha, J. Heckman, and S. Navarro. Separating uncertainty from heterogeneity in life cycle earnings. *Oxford Economic Papers*, 57(2):191, 2005.

[31] A. Deaton and C. Paxson. Intertemporal choice and inequality. *Journal of Political Economy*, 102(3):437–467, 1994.

[32] P. Dolton and A. Vignoles. The incidence and effects of overeducation in the uk graduate labour market. *Economics of Education Review*, 19(2):179–198, 2000.

[33] P.J. Dolton and M.A. Silles. The effects of over-education on earnings in the graduate labour market. *Economics of Education Review*, 27(2):125–139, 2008.

[34] S.G. Donald and K. Lang. Inference with difference-in-differences and other panel data. *The Review of Economics and Statistics*, 89(2):221–233, 2007.

[35] G.J. Duncan and S.D. Hoffman. The incidence and wage effects of overeducation. *Economics of Education Review*, 1(1):75–86, 1982.

[36] Z. Eckstein and K.I. Wolpin. Estimating a market equilibrium search model from panel data on individuals. *Econometrica*, 58(4):783–808, 1990.

[37] H.S. Farber. Mobility and stability: The dynamics of job change in labor markets. In O. Ashenfelter, R. Layard, and D. Card, editors, *Handbook of Labor Economics*, volume 3, pages 2439–2483. Elsevier, 1999.

[38] A. Fasang, S. Geerdes, Schömann K., and L. Siarov. Job satisfaction and labour market mobility. *European Foundation for the Improvement of Living and Working Conditions*, 2007.

[39] F.A. Flyer. The influence of higher moments of earnings distributions on career decisions. *Journal of Labor Economics*, 15(4):689–713, 1997.

[40] J.A. Freeman and B.T. Hirsch. College majors and the knowledge content of jobs. *Economics of Education Review*, 27(5):517–535, 2008.

[41] R.B. Freeman. Job satisfaction as an economic variable. *The American Economic Review*, 68(2):135–141, 1978.

[42] A. Grip and J. Van Loo. The economics of skills obsolescence: a review, the economics of skills obsolescence: theoretical innovations and empirical applications. *Research in Labor Economics*, 21:1–26, 2002.

[43] W. Groot and H. Maassen van den Brink. Overeducation in the labor market: a meta-analysis. *Economics of Education Review*, 19(2):149–158, 2000.

[44] Fatih Guvenen and Anthony Smith. Inferring labor income risk from economic choices: An indirect inference approach. Working Paper 16327, National Bureau of Economic Research, September 2010.

[45] R.E. Hall and F.S. Mishkin. The sensitivity of consumption to transitory income: Estimates from panel data on households. *Econometrica*, 50(2):461–482, 1982.

[46] D.S. Hamermesh. Economic aspects of job satisfaction. *Essays in Labor Market Analysis. New York: John Wiley*, 1977.

[47] J. Hartog. Over-education and earnings: where are we, where should we go? *Economics of Education Review*, 19(2):131–147, 2000.

[48] J. Heckman, R. LaLonde, and J. Smith. The economics and econometrics of active labor market policies. In O. Ashenfelter, R. Layard, and D. Card, editors, *Handbook of Labor Economics*, volume 3, pages 1865–2097. Elsevier, 1999.

[49] J. J. Heckman, H. Ichimura, and P.E. Todd. Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, 64(4):605–654, 1997.

[50] J.J. Heckman, H. Ichimura, J. Smith, and P.E. Todd. Characterizing selection bias using experimental data. *Econometrica*, 66(5):1017–1098, 1998b.

[51] J.J. Heckman, H. Ichimura, and P. E. Todd. Matching as an econometric evaluation estimator. *The Review of Economic Studies*, 65(2):261–294, 1998a.

[52] J. Hersch. Education match and job match. *The Review of Economics and Statistics*, 73(1):140–144, 1991.

[53] G.W. Imbens. Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics*, 86(1):4–29, 2004.

[54] G.W. Imbens. Implementing matching estimators for average treatment effects in stata. In *West Coast Stata Users Group Meeting*, October, 2007.

[55] Danish Technological Institute. Job mobility in the european union: Optimising its social and economic benefits. 2008.

[56] K.G. Jöreskog. Some contributions to maximum likelihood factor analysis. *Psychometrika*, 32(4):443–482, 1967.

[57] B. Jovanovic. Firm-specific capital and turnover. *The Journal of Political Economy*, 87(6):1246–1260, 1979.

[58] B. Jovanovic. Job matching and the theory of turnover. *Journal of Political Economy*, 87(5):972–990, 1979.

[59] S. Kolenikov. Confirmatory factor analysis using confa. *Stata Journal*, 9(3):329–373, 2009.

[60] N. Kristensen and N. Westergård-Nielsen. Does low job satisfaction lead to job mobility? *IZA Discussion Paper No. 1026*, 2004.

[61] P. Krugman. Viagra and the wealth of nations. *The New York Times Magazine*, August 23, 1998.

[62] A. Kucel, M. Vilalta-Bufí, et al. Is there a wage penalty for horizontal and vertical mismatch? *DemoSoc Working Papers*, 2010.

[63] A. Light and K. McGarry. Job change patterns and the wages of young men. *Review of Economics and Statistics*, 80(2):276–286, 1998.

[64] L.A. Lillard. Job turnover heterogeneity and person-job-specific time-series wages. *Annales d'Économie et de Statistique*, 55-56:183–210, 1999.

[65] H. Low, C. Meghir, and L. Pistaferri. Employment risk and wage risk over the life cycle. *American Economic Review*, 100(4):1432–1467, 2010.

[66] J. Lukasiewicz. The dynamics of science and engineering education. *Journal of Engineering Education*, 61(8):880–882, 1971.

[67] B.C. Madrian. Employment-based health insurance and job mobility: Is there evidence of job-lock? *The Quarterly Journal of Economics*, 109(1):27–54, 1994.

[68] S. McGuinness and J. Bennett. Overeducation in the graduate labour market: A quantile regression approach. *Economics of Education Review*, 26(5):521–531, 2007.

[69] E. M. Morin. The meaning of work in modern times. In *10th World Congress on Human Resources Management*, 2004.

[70] D.T. Mortensen and C.A. Pissarides. New developments in models of search in the labor market. In O. Ashenfelter, R. Layard, and D. Card, editors, *Handbook of Labor Economics*, volume 3, pages 2567–2627. Elsevier, 1999.

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

[72] P. Nelson. Information and consumer behavior. *The Journal of Political Economy*, 78(2):311–329, 1970.

[73] B. Petrongolo and C.A. Pissarides. Looking into the black box: A survey of the matching function. *Journal of Economic Literature*, 39(2):390–431, 2001.

[74] C.A. Pissarides. *Equilibrium unemployment theory*. the MIT press, 2000.

[75] F. Postel-Vinay and J.M. Robin. Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*, 70(6):2295–2350, 2002.

[76] B. Richard, L. Pistaferri, and I. Preston. Consumption inequality and partial insurance. *American Economic Review*, 98(5):1887–1921, 2008.

[77] J. Robst. Education and job match: The relatedness of college major and work. *Economics of Education Review*, 26(4):397–407, 2007.

[78] J. Robst. Overeducation and college major: expanding the definition of mismatch between schooling and jobs. *The Manchester School*, 76(4):349–368, 2008.

[79] D. Rodríguez and M. Angeles. Job mobility and wage growth at the beginning of the professional career in spain. *Revista de economía aplicada*, 18(52):5–34, 2010.

[80] M.A. Shields and M. Ward. Improving nurse retention in the national health service in england: the impact of job satisfaction on intentions to quit. *Journal of Health Economics*, 20(5):677–701, 2001.

[81] J.H. Steiger and J.C. Lind. Statistically based tests for the number of common factors. In *annual meeting of the Psychometric Society, Iowa City, IA*, 1980.

[82] R.H. Topel and M.P. Ward. Job mobility and the careers of young men. *The Quarterly Journal of Economics*, 107(2):439–479, 1992.

[83] G.J. van den Berg and G. Ridder. An empirical equilibrium search model of the labor market. *Econometrica*, 66(5):1183–1222, 1998.

[84] W.P.M. Vijverberg. Early human capital skills and labor market entry. In *Third IZA/SOLE Transatlantic Meeting of Labor Economics*, 2004.